GRANT journal
ISSN 1805-062X, 1805-0638 (online), ETTN 072-11-00002-09-4

EUROPEAN GRANT PROJECTS | RESULTS | RESEARCH & DEVELOPMENT | SCIENCE

Genetic polymorphism at the milk protein genes (*CSN1S1*, *CSN2*, and *CSN3*) in the Czech Sumava and Walachian sheep breeds

Zuzana Sztankóová[1]
Věra Mátlová[2]
Jana Rychtářová[3]

[1] Výzkummný ústav živočišné výroby, Přátelství 815, 104 00 Praha - Uhříněves; sztankoova@seznam.cz
[2] Výzkummný ústav živočišné výroby, Přátelství 815, 104 00 Praha - Uhříněves; matlova.vera@vuzv.cz
[3] Výzkummný ústav živočišné výroby, Přátelství 815, 104 00 Praha - Uhříněves; rychtarova.jana@vuzv.cz

Grant: MZE0002701404
Název grantu: Udržitelný rozvoj chovu hospodářských zvířat v evropském modelu multifunkčního zemědělství
Oborové zaměření: Genetika a molekulární biologie

Abstrakt The aim of this work was to determine allele and genotype frequencies at the casein loci, alphaS1-casein (*CSN1S1*), beta-casein (*CSN2*) and kappa-casein (*CSN3*) in two endangered Czech sheep breeds. The study was carried out on 265 sheep (133 Sumava, 132 Walachian) by means of PCR-RFLP and Light Cycler Analysis. These breeds are kept mainly for their good combined efficiency (milk, meat, wool). In particular, the genetic variant C at the *CSN1S1* locus occurred with high frequency in both breeds Sumava (0.981) and Walachian (0.992). Variants A and D were either "absent" in Walachian (A=0.008, D=0), or present with a very low frequency in Sumava (A=0.008, D=0.011). Molecular analysis of the *CSN2* locus showed that the genetic variant A were predominant, had higher frequency (Sumava, 0.778; Walachian, 0.829) than did variant G (Sumava, 0.222; Walachian, 0.171) in both sheep populations. The *CSN3* locus was found to be monomorphic, with no polymorphism typed in either population. According to Hardy-Weinberg equilibrium, both breeds were in genetic equilibrium at the loci *CSN1S1* and *CSN2* ($P<0.05$). The information on the aggregate genotype variability in both breeds could be exploited in the future using specific breeding programs aimed at preserving biodiversity or select animals for the production of both unusual and typical milk products in the Czech Republic.

Klíčová slova milk protein genes, genetic polymorphism, sheep

1. INTRODUCTION

Casein polymorphisms are important and well known due to their effect on qualitative and quantitative traits and technological properties of milk (1), (2), (3), (4). Though studying the genetic polymorphism of milk proteins have raised considerable research interest for goats and cows species, at the present time, there are only a few description of polymorphism in ovine milk, which is addressing genetic control of the variation (5), (6), (7).
Currently, at the *CSN1S1* locus, eight genetic variants (A, B, C, D, E, F, H and I) have been identified by the electrophoretic technique (8), (9), (10) (11) with the nomenclature used for cow and goat milk caseins and genotyped at the DNA level by AS-PCR (12) or PCR-RFLP (13). The primary structure of ovine *CSN1S1* was deduced by Mercier et al. (14) from cDNA and confirmed for the ovine casein variants A, C and D from protein sequencing by (5).
The primary sequence of the *CSN2* was described by Ferranti et al. (15) and the complete sequence confirmed by Provot et al. (16). At present, three genetic variants A, B and C are described (17), (1) and only the sequence difference found between A and C was the substitution of the amino acid Glu in position 2 in variant A for Gln in variant C. Sequence data for the B variant are lacking (7).
CSN3 plays an important role in the formation, stabilization, and aggregation of casein micelles (18), (19) as well as milk production parameters (20), (3). The complete nucleotide sequence of *CSN3* was described byFurret et al. (21). At present, a single nucleotide polymorphism (SNP) at position 237 (GeneBank accession number X51822) of the sheep κ-casein mRNA (22), is described where a thymine was substituted by a cytosine, while the corresponding amino acid remained (6), (7).

The aim of this work was the first time attempt to investigate allele frequencies at casein loci and the frequencies of different allelic combinations of milk protein genes (*CSN1S1*, *CSN2* and *CSN3*) in two Czech sheep breeds, Sumava (S) and Walachian (Wa) and to provide information for a program of selection and conservation.

2. MATERIAL AND METHODOLOGY

A total of 265 individuals from two sheep breeds – Sumava (S, n=133) and Walachian (Wa, n=132) – were used in this study to evaluate genetic polymorphism at the calcium-sensitive caseins (*αS1*, *β*-) and kappa casein. The Sumava and Walachian breeds are bred for combined efficiency (milk, meat, wool). Both sheep breeds are playing a crucial role in the regeneration of the environmental system in the Czech Republic. Currently, both the Sumava and Walachian breeds are entered in the World Genetic Resources Pool and the National Program on Genetic Resources.

Genomic DNA was extracted from blood using ABI PRISM 6100 Analysis (Nucleic Acid PrepStation, Applied Biosystem Co., Foster City, CA, USA) by standard protocol.

The locus *CSN1S1* was determined using PCR-RFLP analysis (13). We typed three alleles: A and C, associated with normal level of the corresponding milk proteins and the allele D associated with a

low level of corresponding protein in milk. The fragments were separated in 4% Metaphor Agarose gel (Cambrex, Rockland, USA).

The amplification products and the restriction patterns were visualized on agarose gel (PCR-agarose, Top-Bio, Ltd., CR) in TBE buffer stained with Ethidium Bromide. Analysis at the *CSN2* and *CSN3* loci were performed with PCR followed by the LightCycler Analysis (22).

PowerMarker data analysis software (version 3.25) (23) was used to estimate allele frequencies, to verify the Hardy-Weinberg equilibrium ($P<0.05$), heterozygosity both observed (Hobs) and expected (Hexp), and polymorphism information content (PIC). Allele frequencies are estimated by simple counting. The differences between heterozygosity observed and expected in accordance with the Hardy-Weinberg equation were tested by $\chi 2$ –analysis (Weir, 1996). The polymorphism information content (PIC) was estimated by (24).

3. RESULTS AND DISCUSSION

For our molecular analysis we typed three genetic variants: A, C and D which are associated with a normal content of protein. Variants A and C correspond to a normal level of the protein content in the milk, compare to with variant D (welsh allele) that is associated with a low level of protein content in the milk of the *CSN1S1*. Results of the genetic polymorphism at the *CSN1S1* locus in sheep populations (Sumava and Walachia) are reported in Table 1. Analysis of the *CSN1S1* locus showed a prevalence of the C allele. The A variant were identified with a very low frequency in both sheep populations. The genetic variant D occurred only in the Sumava sheep population and was characterized by a very low frequency (Table 1). Moreover, in our work we did not observe the genetic variant D in Walachia sheep population. In contrast the present results, some authors (10), (26) described a higher frequency of the "Welsh allele" in the Italian Sarda breed and the Slovak Merino breed, respectively. Comparison of our results with those available in the literature showed similar results in some Italian, Spanish, Hungarian and German breeds (1), (11), (12), (27), (28). Thus indicates that variant C is characteristic in both the Sumava and Walachian sheep populations compared with variants A and D.

A single amino acid substitution at the positions Met183 and Val183 of the mature protein (part of an exon and part of intron 7) was examined at *CSN2* locus. The single nucleotide substitution between genetic variants A (EMBL X79703) and G (AY444504), amino acid exchange Met (ATG) → Val (GTG) at position 183 describe, that the A variant had a higher frequency than did variant G in both studied breeds (Table 2), thus indicating that the A variant is characteristic of sheep populations kept in the Czech Republic compared to variant G. Similar results, with variant A showing a frequency of 50-80% compared to the G variant, were described by (6), (29) in Italian sheep breeds.

At the *CSN3* locus, amino acid substitution at the positions Ser104 and Leu104 of the mature protein was observed. The results regarding exon 4 showed that variant C was predominant (Table 2) in both sheep populations, which is comparable to the results of GeneBank accession no. X51822. This finding is in agreement with the results obtained in other European sheep breeds (6). The second pattern, variant T, as described by (6), was not found in the present study. However, Italian sheep populations demonstrated a very low frequency (1%) of variant T (6).

In accordance with Hardy-Weinberg law, the Sumava and the Walachian sheep populations were consistent with followed the Hardy-Weinberg equilibrium at the *CSN1S1* and *CSN2* loci ($P < 0.05$*). The locus *CSN3* was not tested for Hardy-Weinberg equilibrium, PIC and heterozygosities (observed and expected), because was monomorphic in both sheep breeds studied. Low values were derived from heterozygosities both observed and expected. The PIC showed a small or medium value as a single locus as the mean of population in both populations (not shown). Statistical analysis showed that genetic polymorphism has low diversity (Table 3). We postulate that the low diversity of genetic polymorphism of casein loci could be related to the breed, the origin and migration in the past. Also it could be associated with the breeding program, selection pressure on production properties or genetic drift.

Molecular analysis of the genotype distribution at the casein loci (in the order to *CSN1S1-CSN2-CSN3*) showed, that the most frequent combination of aggregate genotypes was CC-AA-CC, followed by CC-AG-CC (in the order to *CSN1S1-CSN2-CSN3*) in the sheep populations, whereas combination of aggregate genotypes AC-AA-CC, AC-AG-CC, CD-AA-CC and CD-AG-CC had either very low frequency or not detected (Table 4). In default of information about combination of aggregate genotype distribution we are not able to compare our result with those available in the literature; however we suppose that there has been some difference of aggregate genotype combination across world sheep populations. The difference in distribution of the casein allele variation may have been a consequence of the geographic origin of the breeds and their reproduction management pattern as well as the association or binding sites among the loci. The similar occurrence of genetic polymorphism in the genes of milk protein may be caused by genetic relation and genetic drift in studied sheep populations.

4. CONCLUSION

This paper described the first results of genetic polymorphism at the casein loci (*CSN1S1*, *CSN2* and *CSN3*) of the greatest interest of the two Czech national sheep breeds, (Sumava and Walachian). We identified, that undesirable allele D at the *CSN1S1* locus had a very low frequency in the Sumava population and an absence in Walachian population. Observed results of molecular analysis showed that aggregate genotype CC-AA-CC, had the highest frequency of occurrence, followed by CC-AG-CC (in the order to *CSN1S1-CSN2-CSN3*), in both Czech national sheep populations (Sumava and Walachian). Such information could be used for monitor, preserve and manage their biodiversity and their allele combination could be potentially useful in future breeding schemes focusing in improving the quality of the processed milk and cheese yield of Czech sheep breeds.

5. ACKNOWLEDGMENTS

This project was supported by the Ministry of Agriculture of the Czech Republic (Institutional Programme (MZE 0002701404).

REFERENCE

1. AMIGO, L.; RECIO, I.; RAMO, M. Genetic polymorphism of ovine milk proteins: its influence on technological properties of milk – a review. *International Dairy Journal.* 2000, vol. 10, 3, 135-149. ISSN 0958-6946.
2. KUCHTÍK, J.; ŠUSTOVÁ, K.; URBAN, T.; ZAPLETAL, D. Effect of the stage lactation on milk composition, its properties and the quality of rennet curdling in East Friesian ewes. *Czech Journal of Animal Science*. 2008, vol. 53, 2, p. 55-63. ISSN 1212-1819.
3. MATEJÍČEK, A.; MATEJÍČKOVÁ, J.; ŠTÍPKOVÁ, M.; HANUŠ, O.; GENČUROVÁ, V.; KYSEĽOVÁ , J.; NĚMCOVÁ, E.; KOTT, T.; ŠEFROVÁ, J.; KREJČOVÁ , M.; MELČOVÁ, S.;, HÖLZELOVÁ, I.; BOUŠKA, J.; FRELICH, J. Join effects of *CSN3* and *LGB* genes on milk quality and coagulation properties in Czech Fleckvieh. *Czech Journal of Animal Science.* 2008, vol. 53, 6, p. 246-252. ISSN 1212-1819.

4. ORAVCOVÁ, M.; MARGETÍN, M.; PEŠKOVIČOVÁ, D.; DAŇO, J.; MILERSKI, M.; HETÉNYI, L.; POLÁK, P. Factors affecting ewe's milk fat and protein content and relationships between milk yield and milk components. *Czech Journal Animal Science.* 2007, vol. 52, 7, p. 189-198. ISSN 1212-1819.
5. FERRANTI, P.; MALORNI, A.; NITTI, G.; LAEZZA, P.; PIZZANO, R.; CHIANESE, L.; ADDEO, F. Primary structure of ovine α_{S1}-caseins: localization of phosphorylation sites and characterization of genetic variant A, C, and D. Journal of *Dairy Research.* 1995, vol. 62, 2, p. 281-296. ISSN 0022-0299.
6. CERIOTTI, G.; CHESSA, S.; BOLLA, P.; BUDELLI, E.; BIANCHI, L.; DURANTI, E.; CAROLI, A. Single in the ovine casein genes detected by polymerase chain reaction - single strand conformation polymorphism. *Journal of Dairy Science.* 2004, vol. 87, 8, p. 2606-2613. ISSN 0022-0302.
7. PARK, Y. W.; JUARÉZ, M.; RAMOS, M.; HAENLEIN, G. F. W. Physico-chemical characteristics of goat and sheep milk. *Small Ruminant Research.* 2007, vol. 68, 1-2, p. 88-113. ISSN 0921-4488.
8. CHIANESE, L.; GARRO, G.; MAURIELLO, R.; LAEZZA, P.; FERRANTI, P.; ADDEO, F. Occurrence of five alpha (s1)-casein variants in ovine milk. *Journal of Dairy Research.* 1996, vol. 63, 1, p. 49-59. ISSN 0022-0299.
9. RECIO, I.; RAMOS, M.; AMIGO, L. Study of the polymorphism of ovine α_{s1}- and α_{s2}-caseins by capillary electrophoresis. *Journal of Dairy Research.* 1997, vol. 64, 1, p. 525-534. ISSN 0022-0299.
10. PIRISI, A.; PIREDDA, G.; PAPOFF, C. M.; DI SALVO, R.; PINTUS, S.; GARRO, G.; FERRANTI, P., CHIANESE, L. Effect of sheep α-$_{s1}$ casein CC, CD and DD genotypes on milk composition and cheese-making properties. *Journal of Dairy Research.* 1999, vol. 66, 1, p. 409-419. ISSN 0022-0299.
11. GIAMBRA, I. J.; JÄGER, S.; ERHARDT, G. Isoelectric focusing reveals additional casein variants in German sheep breeds. Small Ruminant Research. 2010, vol. 90, 1-3, p. 11-17. ISSN 0921-4488.
12. RAMUNNO, L.; COSENZA, G.; RANDO, A.; MACCIOTTA, N. P. P.; PAPPALARDO, M.; MASINA, P. Identification of carriers of the Welsh *CACA1* variant using an allele-specific PCR method. *Animal Genetics* 1997, vol. 28, 1, p. 154-155. ISSN 0268-9146.
13. PILLA, F.; BEVILACQUEA, C.; LEROUX, C.; FRAGHI, A.; MARTIN, P. Genotyping of α-$_{s1}$ casein in sheep. *Animal Genetics*. 1998, vol. 29, 6, p. 472-473. ISSN 0268-9146.
14. MERCIER, J. C.; GAYE, P.; SOULIER, S.; HUE-DELAHAIE, D.; VILOTTE, J. L. Construction and identification of recombinant plasmids carrying cDNAs coding for α_{s1} – casein, α_{s2} – casein, β - casein, κ-casein and beta lactoglobulin. Nucleotide sequence of for α_{s1} – casein cDNA. *Biochimie*. 1985, vol. 67, 9, p. 959-971. ISSSN 0300-9084.
15. RICHARDSON, B. C.; MERCIER, J. C. The primary structure of the ovine β-caseins. *European Journal Biochemistry.* 1979, vol. 99, 2, p. 285-297. ISSN 0014-2956.
16. PROVOT, C.; PERSUY, M. A.; MERCIER, J. C. Complete sequence of ovine β-casein encoding gene and interspecies comparison inter species comparison. *Gene*. 1995, vol. 154, 2, p. 259-263. ISSN 0378-1119.
17. CHIANESE, L.; GARRO, G.; FERRANTI, P.; MALORNI, A.; ADDEAO, F.; RABASCO, A.; MOLINA, P. Discrete phosphorylation generates the electrophoretic heterogeneity of ovine beta casein. *Journal of Dairy Research.* 1995, vol. 62, 1, p. 9-100. ISSN 0022-0299.
18. SWAISGOOD, H.E. Symposium: Genetic perspectives on milk proteins: comparative studies and nomenclatures. Review and update of casein chemistry. *Journal of Dairy Science.* 1993, vol. 76, 10, p. 3054-3061. ISSN 0022-0302
19. GUTIERREZ, A., MAGA E. A., MEADE H., SHOEMAKER C. F., MEDRANO J. F., ANDESON G. B, MURRAY J. D. Alterations of the physical characteristics of milk from transgenic mice producing bovine κ-casein. *Journal of Dairy Science.* 1996, vol. 79, p. 791-799. ISSN 0022-0302.
20. KUČEROVÁ , J.; MATĚJÍČEK, A.; JANDUROVÁ, O. M.; SØRENSEN, P.; NĚMCOVÁ, E.; ŠTÍPKOVÁ, M.; KOTT, T.; BOUŠKA, J.; FRELICH, J. Milk protein genes *CSN1S1*, *CSN2*, *CSN3*, *LGB* and their relation to genetic values of milk production parameters in Czech Fleckvieh. Czech Journal of Animal Science. 2006, vol. 51, 6, p. 241-247. ISSN 1212-1819.
21. FURET, J. P.; MERCIER, J. C.; GAYE, P.; HUE-DELAHAIE D.; VILOTTE, J. L. Nucleotide sequence of ovine κ-casein cDNA. *Nucleic Acids Research.*. 1990, vol. 18, 17, p.5286. ISSN 0305-1048.
22. SZTANKOOVA, Z.; KYSELOVA, J.; RYCHTAROVA, J.; CZERNEKOVA, V. Technical note: A novel method for routine genotyping of the G allele of beta-casein (CSN2) and T allele of kappa-casein (CSN3) in a sheep population using LightCycler. *Journal of Animal Science*. 2011, vol. 89, 12, p. 3843-3845. ISSN 0021-8812.
23. FELIGINI, M.; VLACO, S.; CURIK, V. C.; PARMA, P., GREPPI, G.; ENNE, G. A single nucleotide polymorphism in the sheep kappa - casein coding region. *Journal of Dairy Research.* 2005, vol. 72, 3, p. 317-321. ISSN 0022-0299.
24. LIU, K. J. and MUSE, S. V. PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics*. 2005, vol. 21, p. 2128-2129. ISSN 1367-4803.
25. BOTSTEIN, D.; WHITE, R. L.; SKOLICK, M.; DAVIS, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics.* 1980, vol. 32, 3, p. 314-*33.* ISSN 0002-9297.
26. BANYKÓ, J. Distribution of α_{S1}-casein "Welsh" variant in some Slovak and Czech sheep breeds. *Archiv fur Tierzuch t- Archives of Animal Breeding.* 2007, vol. 50, 4, p.381-387. ISSN 0003-9438.
27. ANTON, I.; ZSOLNAI, A.; FÉSÜS, L.; KUKOVICS, S.; MOLNÁR, A. Survey of β-Lactoglobulin and α_{s1}-casein polymorphisms in Hungarian dairy sheep breeds and crosses on DNA level (short communication) *Archiv fur Tierzuch t- Archives of Animal Breeding.* 1999, vol. 42, 4, p.387-392. ISSN 0003-9438.
28. LOPEZ-GALVEZ, G.; CHIANESE, L.; ADDEO F.; AMIGO, L.; RAMOS, M. Polymorphism of alpha(s)-casein in the milk of two Spanish ovine breeds. *Milchwissenschaft – Milk Science International*. 1999, vol. 54, 1, p. 17-19. ISSN 0026-3788.
29. CHESSA, S.; RIGNANESE, D.; BERBENNI, M.; CERIOTTI, G.; MARTINI, M.; PAGNACCO G. New genetic polymorphisms within ovine β – and α_{s2} – caseins. Small Ruminant Research. 2010, vol. p. 88, 84-87. ISSN 0921-4488.

Table 1. Allele frequency at the *CSN1S1* locus (A, C and D alleles) in the Czech national sheep breeds (Sumava, Walachian).

CSN1S1	Allele	Sumava, n=133	Walachian, n=132
	A	0,008±0.005	0,008±0.005
	C	0,981±0.008	0.992±0.005
	D	0,011±0.006	0,000±0.000

CSN1S1- alpha S1 casein

Table 2. The single nucleotide polymorphism (SNP), the amino acid (AA) exchanges in the mature protein sequence, and frequencies variants at the *CSN2* and *CSN3* loci in the Sumava and Walachian populations.

Locus [1]	Acc. No.	SNP	AA	Sumava, n=133	Walachian, n=132
CSN2	X79703	A12029	Met183	0.778±0.027	0.829±0.024
	AY444504	G226	Val183	0.222±0.027	0.171±0.024
CSN3	X51822	C443	Ser104	1.000±0.000	1.000±0.000
	AY444505	T164	Leu104	0.000±0.000	0.000±0.000

[1] CSN2 = Beta casein, CSN3 = kappa casein.

Table 3. Heterozygosity (observed H(obs) and expected H(exp)), PIC, p-value and $\chi 2$ at the *CSN1S1-CSN2-CSN3* loci in the Sumava and Walachian sheep breeds.

Locus	Breed	$H_{(exp)}$	$H_{(obs)}$	PIC	p-value	$\chi 2$
CSN1S1	Sumava	0.037	0.038	0.037	0.040	0.049*
CSN2	(n=133)	0.345	0.308	0.286	0.120	1.523*
CSN3		0	0	0	0	0
CSN1S1	Walachian	0.015	0.015	0.015	1.000	0.007*
CSN2	(n=132)	0.283	0.265	0.243	0.359	0.514*
CSN3		0	0	0	0	0

(P*<0.05).

Table 4. Amino acid differences among the calcium-sensitive caseins loci, (in the order to *CSN1S1-CSN2-CSN3*) in the Sumava and Walachian sheep breeds.

CSN1S1-CSN2-CSN3	13	68	183	104	Sumava, n=133	Walachian, n=132
AC-AA-CC	Pro/Ser	Asn	Met	Ser	2	1
AC-AG-CC	Pro/Ser	Asn	Met/Val	Ser	0	1
CC-AA-CC	Ser	Asn	Met		79	91
CC-AG-CC	Ser	Asn	Met/Val	Ser	40	34
CC-GG-CC	Ser	Asn	Val	Ser	9	5
CD-AA-CC		Asn/SerP°	Met	Ser	2	0
CD-AG-CC		Asn/SerP°	Met/Val	Ser	1	0

SerP° -disappearance of two phosphate groups in the phosphorylated residues Ser64 and Ser66 in variant D