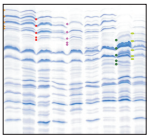
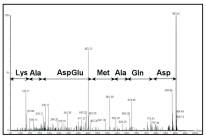
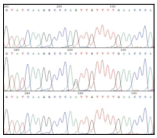
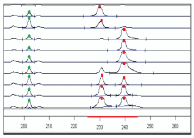
OVINE MILK PROTEINS:
DNA, mRNA, and protein analyses and their associations to milk performance traits


Isabella Jasmin Giambra

A dissertation submitted to the Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management, Justus-Liebig-University Gießen, for the doctoral degree of Dr. agr.





édition scientifique
VVB LAUFERSWEILER VERLAG

1. Auflage 2011

1st Edition 2011

Printed in Germany

édition scientifique
VVB LAUFERSWEILER VERLAG

STAUFENBERGRING 15, D-35396 GIESSEN
Tel: 0641-5599888 Fax: 0641-5599890
email: redaktion@doktorverlag.de

www.doktorverlag.de

Department of Animal Breeding and Genetics, Justus-Liebig-University Gießen

Supervisor: Prof. Dr. G. Erhardt

Ovine milk proteins: DNA, mRNA, and protein analyses and their associations to milk performance traits

A dissertation submitted to the Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management, Justus-Liebig-University Gießen,
for the doctoral degree of Dr. agr.

Presented by Isabella Jasmin Giambra (M. Sc.)
born in Siegen

Gießen, 2011

With the consent of the Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management, Justus-Liebig-University Gießen

Dean:
Prof. Dr. Ingrid-Ute Leonhäuser

1. Referee: Prof. Dr. G. Erhardt
2. Referee: Prof. Dr. Dr. habil. E. Usleber
3. Referee: Prof. Dr. V. Dzapo

Date of disputation:
14.04.2011

Meiner Familie und in Gedenken an W. und C.

Table of contents

1. General introduction

1.1. Importance of dairy sheep production and breeding

1.1.1. Worldwide

Sheep producing milk for humans started with the beginning of domestication (Haenlein, 2007). Worldwide dairy sheep population exhibited in 2008 28.72% of all dairy animals, thus being numerically the second important dairy species all over the world. In Africa, America, Asia, and Europe sheep milk represented 4.92%, 0.02%, 1.70%, and 1.44% of total milk produced over all species (in 2008), respectively (FAOstat, 2010). In the same year 34.10% of the entire sheep milk all over the world (9,129,480 tonnes) was produced in Europe, in detail 22.43% in southern Europe (FAOstat, 2010). Caused by low numbers of dairy sheep in northern Europe, no data are available for sheep milk efficiency, demonstrating the low importance of sheep milk production in this region.

In regions of the world where the hard environmental conditions require special adaptation and for which sheep are better suited than cattle (Arranz *et al.*, 2001; Haenlein, 2001a; Park and Haenlein, 2006; Degen, 2007), they are extensively held as "the cows of the small holder" (Haenlein and Wendorff, 2006), to provide home supply and self-sufficiency for families, to avoid starving and malnutrition especially in high-quality protein (Morand-Fehr and Boyazoglu, 1999; Haenlein, 2001a; Degen, 2007). Also in least developed countries sheep were in 2008, after goat, the second important dairy species with about 55 million head producing approximately 1.6 million tonnes milk (FAOstat, 2010). In these countries numbers of dairy sheep and sheep milk production yield increased over the last centuries.

Furthermore, dairy sheep farming is a vital part of the national economy in many countries, especially in the Mediterranean and Middle East region (Pandya and Ghodke, 2007; Park *et al.*, 2007; Carta *et al.*, 2009). Dairy sheep farming varies from extensive (marked seasonal milk production, dual-purpose breeds, low feed supplementation, transhumance, hand-milking, absence of farm facilities, farm-made cheese) to intensive (seasonal or continuous milk production, improved local breeds or crosses, exploitation of forage crops, high feed supplementation, milking machine and housing facilities, industrial cheeses) systems according to the economic relevance of the production chain and the specific environment and breed (Carta *et al.*, 2009). Around the Mediterranean Sea about 60% of all ewes are milked totally or partially, and about 95% of sheep and goat milk is transformed into typical dairy products, mainly yoghurts and cheeses, because of the high fat and protein yield of sheep milk (Boyazoglu and Morand-Fehr, 2001). These products often have regional or local connotation of origin and quality (for example Roquefort, Pecorino Romano, and Feta) (Boyazoglu and

Morand-Fehr, 2001; Ugarte *et al.*, 2001; Park and Haenlein, 2006) and are regulated under a Denomination of Origin and protected by the European Union (Freitas and Malcata, 2000; Haenlein and Wendorff, 2006). Known history and origin of these natural, organic, farm-fresh, and original artisan products have been leading to a worldwide increasing demand of sheep milk products over the last 30 years, caused by searching for a change in nutrition, gourmet products, or alternatives to cow milk (Haenlein, 2001a; Park and Haenlein, 2006; Barillet, 2007). Sheep milk, cheese, and yoghurt can provide a profitable alternative to cow milk products owing to their specific taste, texture and their natural and healthy image (Raynal-Ljutovac *et al.*, 2008).

Structured breeding programmes in dairy sheep are only developed in countries like France, Italy, and Spain, where selection success is in evidence (Arranz *et al.*, 2001; Barillet *et al.*, 2001; Astruc *et al.*, 2007; Zumbach and Peters, 2008). Mostly only parts of total numbers of dairy ewes are tested for milk performance traits (Astruc *et al.*, 2004) and qualitative recording is included only for some breeds in breeding schemes (Carta *et al.*, 2009). Due to the great importance of cheese and yoghurt production in the dairy sheep sector, milk yield is not of primary interest as is fat and protein content (Haenlein and Wendorff, 2006; Barillet, 2007), which therefore should be selective objectives (Arranz *et al.*, 2001; Barillet, 2007). Unfortunately payment systems for sheep milk based on milk composition are applied only in some European countries, particularly where the dairy sheep sector is well developed (Carta *et al.*, 2009).

1.1.2. Germany

No data about dairy sheep production in Germany are available in world- or European-wide statistics (Eurostat, 2010; FAOstat, 2010). In 2003 ICAR indicated 20,000 dairy sheep in Germany (Astruc *et al.*, 2004) and Rahmann (2001) described an estimated German sheep milk production of 4,000 tonnes in 2001. From 2005 to 2009, ICAR registered the total German dairy sheep population, decreasing during this time. 8,100 East Friesian Dairy sheep and 104 Lacaune sheep were held in Germany in 2009 (http://www.waap.it/sheep_enquiry/). Participation in performance tests of East Friesian Dairy sheep was reduced by half during this time to a number of 602 tested ewes in 2009 held in 53 flocks.

In Germany test day records are starting or completely lying (in flocks which do not milk ewes) within the suckling period (Astruc *et al.*, 2004; Astruc *et al.*, 2007), whereas fat, protein, and lactose content as well as somatic cells and urea are always analysed (Astruc *et al.*, 2004; Carta *et al.*, 2009). Milk recording is delegated to several milk recording

organisations, which are dealing with all dairy species, meaning that in each organisation dairy sheep activity is not important, compared to cattle (Astruc *et al.*, 2002). An efficient breeding programme for East Friesian Dairy sheep is missing, although it is necessary for the commercial dairy sheep breeders and to stop the decreasing German dairy sheep population (Zumbach and Peters, 2008).

However, it is to conclude that in Germany a bulk of dairy sheep holders are hobby-breeders, partly aiming lamb and not milk production (Astruc *et al.*, 2002; Zumbach and Peters, 2008) and maintaining the breed with all the typical standard signs (Astruc *et al.*, 2002).

1.2. Milk performance and composition

1.2.1. Milk yield

Milk yield of dairy sheep is often recorded as "after suckling" lambs for four to eight weeks. Contrary to the practice for dairy goats or cattle, ewes are milked at least for another five months. Only in intensive management systems lambs are not suckled at all but fed on a lamb nursing bar with reconstituted commercial milk replacer and ewes are milked for ten months (Haenlein and Wendorff, 2006).

By comparing 4%-fat-corrected milk yields of different sheep breeds all over the world, and also with goat breeds, the East Friesian Dairy sheep is by far the leading dairy sheep in yield and physiologically comparable to the best dairy goat breeds (Haenlein, 2007). Table 1 demonstrates milk yield for different sheep breeds used for milk production in several countries.

Table 1. Milk yield and milking period of several (dairy) sheep breeds after suckling period (Haenlein and Wendorff, 2006)

Breed	Country	Milk yield (kg/year)	Days of milking period
Assaf	Israel	150-350	150-210
Awassi	Israel	440-550	240-300
Awassi	Turkey	130-205	120
Chios	Greece	135-300	170-250
Churra	Spain	150	150
Comisana	Italy	90-175	-
East Friesian	Germany	500-900	300-365
Lacaune	France	270	160-170
Lacha	Spain	210	180
Manchega	Spain	80-520	150-270
Massese	Italy	110-140	-
Sarda	Italy	120-195	-
Valle de Belice	Italy	270	200

1.2.2. Milk composition

Sheep milk contains higher levels of total solids and major nutrient than goat and cow milk (Table 2), especially of average protein and fat (Park *et al.*, 2007). This results in higher cheese yield (approximately 15% for sheep milk compared with 10% for cow milk), because cheese curd contains primarily fat and casein (Storry *et al.*, 1983; Haenlein and Wendorff, 2006; Recio *et al.*, 2009).

Compared to cow milk, lactose content in sheep milk is at about the same level (Park *et al.*, 2007), whereas mineral contents of sheep milk are comparable with goat and mostly higher than in cow milk (Park *et al.*, 2007).

Table 2. Composition of sheep milk (in %) in comparison to other species (Jandal, 1996; Haenlein and Wendorff, 2006; Degen, 2007; Pandya and Ghodke, 2007; Park *et al.*, 2007)

Species	Dry mass	Whole protein	Casein	Whey proteins	Fat	Lactose	Mineral nutrients
Sheep	16.0-20.0	4.5-6.6	4.2-5.2	0.8-0.9	5.3-9.3	3.9-4.9	0.8-0.9
Cattle	12.3-14.5	3.2-4.0	2.6-2.8	0.6	3.4-4.5	4.6-4.9	0.7
Goat	11.5-13.5	2.8-3.7	2.4-2.6	0.4-0.6	3.4-4.5	3.9-4.8	0.8
Buffalo	17.2	3.8	3.2	0.6	7.6	4.8	0.8
Yak	17.0-18.0	5.5	n.a.[a]	n.a.[a]	6.5-9.0	5.0-6.0	n.a.[a]
Camel	13.5-16.0	3.5-4.5	n.a.[a]	n.a.[a]	5.0-5.5	5.0-6.0	n.a.[a]
Horse	10.0-12.0	1.8-2.5	1.3	1.2	1.0-2.0	6.0-7.0	0.5
Human	12.4-12.9	0.9-1.6	0.4	0.5-0.7	3.7-4.7	6.9-7.1	0.2-0.3

[a] = Not applicable.

In addition, ovine milk protein fraction shows a higher content of essential amino acids (55.1% of all amino acids) in comparison to cow (52.1%) and goat (55.3%) (http://www.ernaehrung.de; Hinrichs, 2004).

Milk lipids are present in the form of globules, which are characteristically only 3.3 μm in ewe milk compared to cow milk, where fat globule diameter is about 4.6 μm (Park *et al.*, 2007). Furthermore, sheep milk shows a higher ratio of linoleic acid than cow milk, comparable with that of goat milk, as well as the highest concentration of linolenic acid and polyunsaturated fatty acids in total (http://www.ernaehrung.de; Park *et al.*, 2007).

Vitamin content of sheep milk is considerably higher than in goat and cattle milk, especially that of vitamin C for example, but with exception of β-carotene (http://www.ernaehrung.de; Park *et al.*, 2007; Raynal-Ljutovac *et al.*, 2008).

1.2.3. Genetic and environmental influences on milk yield and composition

The great genetic variation existing in milk composition, lactation length, lactation yield, and seasonality of milk production (Haenlein, 2001b), makes it possible to improve milk production through selection. Estimated heritabilities (h^2) of milk production traits for

different sheep breeds are in a moderate to high range (Table 3). Generally, heritabilities for milk, fat, and protein yield are smaller than for fat and protein content, comparable with cattle (Barillet, 2007).

Table 3. Estimated heritabilities for milk performance traits in dairy sheep (Barillet and Boichard, 1987; Sanna *et al.*, 1997; El-Saied *et al.*, 1998; El-Saied *et al.*, 1999; Othmane *et al.*, 2002a; Othmane *et al.*, 2002b; Hamann *et al.*, 2004)

Trait	h^2
Milk yield	0.10-0.30
Fat yield	0.15-0.27
Protein yield	0.15-0.26
Fat %	0.09-0.62
Protein %	0.20-0.55
Casein %	0.21-0.30

Moreover, correlations between milk yield and milk composition are described. In detail null or negative correlations exist between yield and content traits (0.02 to -0.51), as well as highly positive genetic correlations between yield traits (0.89 to 0.94) and between fat and protein content (0.24 to 0.69) (Sanna *et al.*, 1997; El-Saied *et al.*, 1999; Legarra and Ugarte, 2001). Protein and casein content are also strongly related (0.97 to 0.99) (Fuertes *et al.*, 1998; Othmane *et al.*, 2002a; Othmane *et al.*, 2002c).

Besides genetic factors like breed and genotype, milk yield and composition are furthermore influenced

- by physiological factors like
  - age at lambing (Carta *et al.*, 1995; Gootwine and Goot, 1996; Othmane *et al.*, 2002a)
  - number of parity (Casoli *et al.*, 1989; Gabina *et al.*, 1993; Sevi *et al.*, 2000),
  - number of suckled lambs (litter size) (Gootwine and Goot, 1996; Ruiz *et al.*, 2000),
  - lactation length (Casoli *et al.*, 1989; Gabina *et al.*, 1993; Ruiz *et al.*, 2000) or stage of lactation (Wilson *et al.*, 1971; Pellegrini *et al.*, 1997; Brozos *et al.*, 1998),
  - month of lambing or season (Gabina *et al.*, 1993; Carta *et al.*, 1995; Ruiz *et al.*, 2000), and
  - (udder) state of health (Haenlein, 2001b; Nudda *et al.*, 2003; Bianchi *et al.*, 2004),
- by management systems like
  - milking techniques and interval (Bencini, 2001; Castillo *et al.*, 2008), and
  - feeding composition and strategy (Wilson *et al.*, 1971; Haenlein and Wendorff, 2006; Park *et al.*, 2007).

1.3. Milk proteins

The principal milk proteins in sheep are about the same as in cow and goat milk (Park *et al.*, 2007) and are divided into two big classes: the caseins and the whey proteins.

1.3.1. Caseins

Within the caseins the four fractions α_{s1}- (*CSN1S1*), α_{s2}- (*CSN1S2*), β- (*CSN2*), and κ-casein (CN; *CSN3*) are distinguished according to the homology of their primary structures (Eigel *et al.*, 1984). Caseins are the major proteins in sheep milk and form 76-83% of total protein (Mercier *et al.*, 1978a; Hinrichs, 2004; Park *et al.*, 2007), whereas the four messenger ribonucleic acids (mRNAs) of the caseins are not translated with the same efficiency. Transcripts of α_{s1}- and β-CN are translated approximately 3- to 4-fold more efficiently than those of α_{s2}- and κ-CN (Bevilacqua *et al.*, 2006). Therefore, the relative percentages of the individual ovine caseins range as follows: κ-CN from 7.0% to 17.4%, α_{s2}-CN from 8.0% to 16.4%, α_{s1}-CN from 32.0% to 39.9%, and β-CN from 37.0% to 56.5% (Law, 1992; Calavia and Burgos, 1998; Bramanti *et al.*, 2003; Moatsou *et al.*, 2004).

Caseins are phosphoproteins and are organised within micelles, whereas the calcium-sensitive α_{s1}-, α_{s2}-, and β-caseins are stabilised within the micelle by κ-CN (Rijnkels, 2002). The micelles are interlinked by calcium phosphate and small amounts of magnesium, sodium, potassium, and citrate, diffusing the light and lending the milk its opaque white appearance (Park *et al.*, 2007).

Caseins play a nutritive function as a source of amino acids, calcium, and phosphor for the neonates (Holt and Sawyer, 1988) and moreover form the essential structure of the cheese matrix (Hinrichs, 2004).

1.3.1.1. Composition of casein genes

The casein genes *CSN1S1*, *CSN2*, *CSN1S2*, and *CSN3*, are tightly linked within a 250 kb cluster (Threadgill and Womack, 1990; Lévéziel *et al.*, 1991; Bevilacqua *et al.*, 2006) on ovine chromosome 6 (OAR6) (http://www.animalgenome.org/sheep/maps/; de Gortari *et al.*, 1998).

The genes encoding the family of α_{s1}-, α_{s2}- and β-caseins are evolved by both intra- and intergenic duplications of a primordial gene (Gaye *et al.*, 1977; Hobbs and Rosen, 1982; Jones *et al.*, 1985; Bonsing and Mackinlay, 1987). These lead to an extremely split architecture of the mosaic *CSN1S1* and *CSN1S2* genes (Table 4), with small exons in the coding region (Rijnkels, 2002). In *CSN1S1* for example seven out of 19 exons are only 24

base pairs (bp) in length (Groenen *et al.*, 1993; Passey *et al.*, 1996; Rijnkels, 2002). Due to the close relatedness the 5´-noncoding region, the signal peptide-coding region, the regions encoding the ubiquitous multiple phosphorylation sites and the 3´-noncoding region of the mRNA of *CSN1S1*, *CSN1S2*, and *CSN2* are conserved and show great homologies (Jones *et al.*, 1985; Mercier and Vilotte, 1993). *CSN3* exhibits a different composition and it is proposed to be evolutionary related to the γ-fibrinogen gene (Jollès *et al.*, 1974; Alexander *et al.*, 1988). Structural features of the deoxyribonucleic acid (DNA) and mRNA sequences of these four casein genes are demonstrated in Table 4.

Table 4. Known structural features of the four ovine casein genes (Boisnard and Petrissant, 1985; Mercier *et al.*, 1985; Provot *et al.*, 1989; Furet *et al.*, 1990; Groenen *et al.*, 1993; Provot *et al.*, 1995; Rijnkels, 2002)

Casein	DNA (n[a] =)		mRNA (bp)			
	Exons	Introns	Complete	5'-UTR[b]	CDS[c]	3'-UTR[b]
CSN1S1	19	18	1,100	61	618	431
CSN1S2	18	17	1,024	53	669	302
CSN2	9	8	1,088	60	669	359
CSN3	5	4	854	69	579	206

[a] = Number.
[b] = Untranslated region.
[c] = Coding sequence.

1.3.1.2. Composition of casein proteins

Structural homologies of α_{s1}-, α_{s2}-, and β-CN proteins in sheep, cattle, and goat are visible in Table 5, demonstrating lengths of their amino acid chains and signal peptides and their degrees of phosphorylation. Protein structure of κ-CN differs considerably from α_{s1}-, α_{s2}-, and β-CN, confirming that it has not arisen from the same ancestral gene as the other caseins (Holt and Sawyer, 1988).

Within the ovine caseins, α_{s1}- and α_{s2}-CN show highest phosphorylation levels (Boisnard and Petrissant, 1985; Mercier *et al.*, 1985; Ferranti *et al.*, 1995), β-casein is the most hydrophobic one (Martin *et al.*, 2003), and κ-CN is glycosylated at various levels (Jollès *et al.*, 1974; Furet *et al.*, 1990).

Table 5. Comparison of known structural features of the four caseins in ovine, bovine, and caprine milk (according to Chianese *et al.*, 1995; Martin *et al.*, 2003; Park *et al.*, 2007)

Casein	Ovine			Bovine			Caprine		
	Amino acids[a]	Amino acids[b]	P sites[c]	Amino acids[a]	Amino acids[b]	P sites[c]	Amino acids[a]	Amino acids[b]	P sites[c]
α_{s1}-CN	199	15	10/10	199	15	9/9	199	15	11/11
α_{s2}-CN	208	15	17/13	207	15	17/12[d]	208	15	16/10[e]
β-CN	207	15	7/7	209	15	6/5	207	15	6/6
κ-CN	171	21	5/3	169	21	5/3	171	21	6/3

[a] = Number of amino acid residues of the mature protein chain of the protein.
[b] = Number of amino acid residues of the signal peptide.
[c] = Number of phosphorylation sites (putative/actual).
[d] = UniProtKB/Swiss-Prot P02663.
[e] = UniProtKB/Swiss-Prot P33049.

1.3.2. Casein polymorphisms

The heterogeneity of caseins is determined either by the presence of genetic variants or by other factors like phosphorylation level (Table 5), variation in the extent of glycosylation of the κ-CN fraction, and the coexistence of proteins with different chain lengths in α_{s1}- and α_{s2}-CN due to alternative splicing events (Ferranti *et al.*, 1995; Ferranti *et al.*, 1998; Park *et al.*, 2007). These heterogeneities cause different molecular weights, electrical charges, and hydrophobicity of the proteins (Haenlein and Wendorff, 2006).

Alternative splicing has already been described in other ruminant casein genes (Leroux *et al.*, 1992; Bouniol *et al.*, 1993; Martin and Leroux, 1994; Mohr *et al.*, 1994; Grosclaude and Martin, 1997; Mahé *et al.*, 1999) and seems to be common in casein genes probably explainable by the evolutionary caused split architecture (Rijnkels, 2002).

Determination of polymorphisms within ovine milk proteins, reviewed by Amigo *et al.* (2000), started by analyses on protein level, using electrophoretic techniques. Several electrophoretic studies revealed variability in ovine milk proteins, whereas the knowledge of milk protein genetic variants is more fragmentary in ovine species in comparison to goat and cattle breeds (Amigo *et al.*, 2000; Ceriotti *et al.*, 2004; Moioli *et al.*, 2007).

1.3.2.1. α_{s1}-casein

1.3.2.1.1. Protein variability

So far, eight phenotypes (A, B, C, D, E, F, G, X) of α_{s1}-CN in ovine milk have been identified by protein electrophoresis or chromatography (Chianese *et al.*, 1996; Pirisi *et al.*, 1999; Wessels *et al.*, 2004; Chianese, 2008), whereas primary structures have been published only for α_{s1}-CN A, C, D, and E (Ferranti *et al.*, 1995; Chianese *et al.*, 2007). The α_{s1}-CN D was previously described as the Welsh variant because of its first discovery in the Welsh

Mountain sheep (King, 1966). Most common is α_{s1}-CN C with frequencies of up to 89.0% (Table 6).

Table 6. Distribution of seven published protein variants of ovine α_{s1}-CN and their corresponding allele frequencies estimated by biochemical analyses (Chianese *et al.*, 1996; Pirisi *et al.*, 1999; Amigo *et al.*, 2000; Wessels *et al.*, 2004; Barillet *et al.*, 2005; Chianese *et al.*, 2007)

Variant	Diffusion	Allele frequency
A	Common[a]	0.02-0.06
B	Common[a]	0.10-0.43
C	Common[a]	0.49-0.89
D	Common[a]	0.001-0.25
E	Leccese	0.003
F	Sarda	0.02
X	East Friesian Dairy sheep	0.12

[a] = In several breeds (different countries and purposes).

The differences of the three genetic variants *CSN1S1*A*, *C*, and *D* are single amino acid exchanges influencing the degree of phosphorylation (Ferranti *et al.*, 1995; Amigo *et al.*, 2000). α_{s1}-CN C differs from α_{s1}-CN A in the amino acid substitution of Ser13Pro, which determines the loss of the phosphate group at position 12 of the protein chain (SerP12Ser) in mature protein (Ferranti *et al.*, 1995). A further substitution, SerP68Asn, causes the disappearance of both phosphate groups in the phosphorylated residues Ser_{64} and Ser_{66} in variant D (Ferranti *et al.*, 1995). Ovine α_{s1}-CN E is characterised by the lack of the amino acid residues 70 to 77, leading to a protein with the lowest number of phosphorylated residues and a chain length shorter than the mature ovine counterparts (Chianese *et al.*, 2007). Amino acid sequence studies of α_{s1}-CN phenotypes B and G are under progress (Chianese *et al.*, 2008), whereas α_{s1}-CN F (Pirisi *et al.*, 1999) and X (Wessels *et al.*, 2004) have not been further characterised until now.

1.3.2.1.2. Non-allelic protein and mRNA forms

Within the ovine α_{s1}-CN alleles A, C, and D the appearance of at least eight protein forms is described, occurring at three different phosphorylation levels (caused by the partial phosphorylation of Ser_{41} and Ser_{115}), not related to allelic variability (Ferranti *et al.*, 1995; Ferranti *et al.*, 1998; Ferranti *et al.*, 1999; Ferranti *et al.*, 2001). They result from abbreviated mRNA forms, founded by alternative splicing events (Passey *et al.*, 1996; Ferranti *et al.*, 1999). These are probably caused by single nucleotide polymorphisms (SNPs), insertions, or deletions concerning the 3'- or 5'- splice sites or the branch point (Leroux *et al.*, 1992; Martin *et al.*, 2003). Therefore, α_{s1}-CN phenotypes in ovine milk are a mixture of these eight forms, whereas the main component corresponds to the 199 bp long form (Table 7).

Table 7. Protein forms of ovine α_{s1}-CN, missing mRNA parts and amino acids, resulting protein length, and their percentage of whole α_{s1}-CN (Ferranti *et al.*, 1998; Ferranti *et al.*, 2001)

Protein form	Missing parts		Resulting protein length (Amino acids)	% of whole α_{s1}-CN
	mRNA	Amino acids (n =)		
1	-	-	199	45.1
2	1. codon exon 11	1	198	17.1
3	Exon 16	8	191	16.9
4	Exon 13	8	191	10.0
5	1. codon exon 11 + exon 16	9	190	4.5
6	Exon 13 + exon 16	16	183	3.6
7	1. codon exon 11 + exon 13	9	190	3.0
8	1. codon exon 11 + exon 13 + exon 16	17	182	n.a.[a]

[a] = Not applicable in Ferranti *et al.* (2001).

1.3.2.1.3. Genetic variability

The differences of the three genetic variants *CSN1S1*A*, *C*, and *D* are SNPs lying within exon 3 and exon 9 (Ramunno *et al.*, 1997; Pilla *et al.*, 1998). Ceriotti *et al.* (2005) differentiated *CSN1S1*C* in *C'*, probably the ancestral form, and *C''* caused by a C>T-SNP lying within exon 17 of *CSN1S1* and leading to the amino acid replacement Ile186Thr (GenBank accession number (Acc. No.) NP_001009795). This amino acid exchange is not identifiable at protein level by isoelectric focusing (IEF), because ancestral and exchanged amino acids are both neutral (Ceriotti *et al.*, 2004; Ceriotti *et al.*, 2005). Pariset *et al.* (2006) confirmed the nucleotide substitution within exon 17. Furthermore, an additional A>G-SNP within the 5'-UTR of *CSN1S1* in eight different European sheep breeds was identified (Pariset *et al.*, 2006). Bastos (2001) detected a single strand conformation polymorphism (SSCP) in ovine *CSN1S1* exons 10 and 11.

1.3.2.2. α_{s2}-casein

1.3.2.2.1. Protein variability

Protein analyses suggested the occurrence of three not further characterised α_{s2}-CN protein variants (Chianese *et al.*, 1993). Chessa *et al.* (2003) named two phenotypes identified by IEF α_{s2}-CN A and B, which were characterised by amino acid exchanges at positions 75 and 105 (Asp75Tyr and Ile105Val) of mature protein (Picariello *et al.*, 2009). The allele frequencies of these two protein patterns in different sheep breeds are demonstrated in Table 8.

Table 8. Distribution of the two protein variants α_{s2}-CN A and B and their corresponding allele frequencies determined by IEF (Chessa *et al.*, 2003; Wessels *et al.*, 2004[a]; Caroli *et al.*, 2007a; Picariello *et al.*, 2009)

Variant	Diffusion	Allele frequency
A	Common[b]	0.60-0.86
B	Common[b]	0.14-0.40

[a] = Denomination of Wessels *et al.* (2004) was adapted according to Chessa *et al.* (2003).
[b] = In several breeds (different countries and purposes).

1.3.2.2.2. Non-allelic protein and mRNA forms

Two non allelic forms arising from casual alternative splicing were also identified in ovine α_{s2}-CN by Boisnard *et al.* (1991). They differ by an internal deletion of the nine amino acid residues 34-42 of the α_{s2}-CN chain and are produced by different mRNAs (Boisnard *et al.*, 1991).

1.3.2.2.3. Genetic variability

Boisnard *et al.* (1991) described a total of four types of *CSN1S2* complementary DNA (cDNA) caused by nucleotide substitutions and the presence or absence of an internal stretch of 44 nucleotides in the 5'-UTR. These cDNA types lead to deduced differences in the amino acid sequence at positions p.49 and p.200 (Asp49Asn and Asn200Lys; GenBank Acc. No. X03238).

Recently Chessa *et al.* (2010) confirmed the sequence difference leading to the amino acid exchange Asn200Lys, which seems to be associated with an additional T/A-SNP in intron 16, by polymerase chain reaction (PCR)-SSCP. The variant coding for Lys_{200} was called *CSN1S2*Y* in comparison to *CSN1S2*N* (Asn_{200}). In addition, an isoconformer of *CSN1S2*N* was identified and named *CSN1S2*X* (Chessa *et al.*, 2010).

1.3.2.3. β-casein

1.3.2.3.1. Protein variability and non-allelic protein forms

Within ovine β-CN heterogeneity in electrophoretic patterns was found in different breeds (Richardson and Mercier, 1979; Chianese *et al.*, 1995; Martin *et al.*, 2003). Two main bands, β_1- and β_2-CN, were identified by Richardson and Mercier (1979) whereas the different mobility is due to the number of their phosphate groups, five and six respectively; Thr_{12} is phosphorylated in β_1-CN but not in β_2-CN. Chianese *et al.* (1995) identified further β-CN patterns, expressed in five types of phenotypes. They differed in the relative intensity and the number of bands, also caused by a different number of phosphate residues in the same protein chain (Chianese *et al.*, 1995).

Further differences in electrophoretic mobility were not found (Moioli *et al.*, 2007). However, comparing the amino acid sequences published by Richardson and Mercier (1979), Provot *et al.* (1989), and Ceriotti *et al.* (2004) sequence differences concerning Thr55Ala and Ala67Pro are indicated.

1.3.2.3.2. Genetic variability

PCR-SSCP-analyses lead to the identification of molecular genetic variability within the coding sequence of ovine *CSN2*. Two conformational patterns, not further characterised, were demonstrated by Bastos (2001) for *CSN2* exon 7. An A>G-SNP lying within this exon 7 and leading to the deduced amino acid exchange Met183Val was identified as being causal for these patterns (Ceriotti *et al.*, 2004; Corral *et al.*, 2010). This SNP was recently confirmed in Spanish Merinos (Corral *et al.*, 2010) and other Italian sheep breeds (Chessa *et al.*, 2010) and the resulting PCR-SSCP patterns were named *CSN2*A* (Met_{183}) and *CSN2*G* (Val_{183}) (Chessa *et al.*, 2010). The Met183Val amino acid exchange could not be recognised at the protein level using electrophoretic methods, because both amino acids are neutral (Ceriotti *et al.*, 2004).

In addition, two new PCR-SSCP patterns, called *CSN2*X* and *CSN2*Y*, were identified by Chessa *et al.* (2010). Sequencing revealed that *CSN2*X* is due to a silent mutation (G>A) concerning Gln_{192}, and that *CSN2*Y* is caused by a C>A transversion leading to the amino acid exchange Leu196Ile (Chessa *et al.*, 2010).

1.3.2.4. κ-casein

1.3.2.4.1. Protein variability and non-allelic protein forms

Most studies could not identify any electrophoretic protein variation for ovine κ-CN (Moioli *et al.*, 1998; Chessa *et al.*, 2003), but non-genetic polymorphisms occur due to varying degrees of glycosylation and phosphorylation (Calavia and Burgos, 1998; Moreno *et al.*, 2000). However, two κ-CN protein fractions were isolated by Alais and Jollès (1967), but a corresponding genetic polymorphism was never suggested at the protein level (Ceriotti *et al.*, 2004). Comparison of protein or cDNA sequences of Jollès *et al.* (1974), Gaye *et al.* (1977), and Trujillo *et al.* (2000) revealed sequence differences concerning the amino acid residues 2 and 7 of mature protein (Gln2Glu and Glu7Gln).

1.3.2.4.2. Genetic variability

More recently, molecular genetic analyses of ovine *CSN3*, mainly of exon 4, started in different sheep breeds, and partly synonymous and non-synonymous sequence differences were identified (Barroso *et al.*, 1996; Bastos, 2001; Ceriotti *et al.*, 2004; Feligini *et al.*, 2005). Ceriotti *et al.* (2004) identified a C>T-SNP in one Italian sheep by PCR-SSCP, leading to the amino acid exchange Ser104Leu. Both amino acids are neutral and, therefore, the exchange can not be revealed by protein electrophoresis (Ceriotti *et al.*, 2004). Feligini *et al.* (2005), Pariset *et al.* (2006) and Hartmann (2010) identified in different sheep breeds a synonymous T>C-SNP at position 237 of the mRNA sequence (GenBank Acc. No. AY237637), concerning amino acid Tyr_{35}. Within studies of Bastos (2001) and Barroso *et al.* (1996) ovine *CSN3* was monomorphic and showed an identical pattern in PCR-SSCP.

Additionally, a microsatellite within intron 3 of ovine *CSN3*, previously described for bovine *CSN3*, was identified with five alleles (Maddox *et al.*, 2001; Corral *et al.*, 2010).

1.3.3. Whey proteins

Sheep milk whey proteins account for 17-22% of total proteins and compose a soluble phase (Park *et al.*, 2007). The two main whey proteins in ruminant milk, synthesised within the udder, are α-lactalbumin (α-LA; *LAA*) and β-lactoglobulin (β-LG; *LGB*) (Mercier *et al.*, 1978a), whereas β-LG is the major whey protein in sheep milk (Gaye *et al.*, 1986). Sheep milk whey protein characterisation by high performance liquid chromatography gave an α-LA content of 1.6 g/l and a β-LG content of about 11.0 g/l in Comisana milk (Vincenzetti *et al.*, 2008). Further whey proteins, like immunoglobulins, serum albumin, and lactoferrin for example, are present in smaller concentrations (Haenlein and Wendorff, 2006; Park *et al.*, 2007).

α-LA is essential for the biosynthesis of lactose in the mammary gland (Brew and Grobler, 1992; Martin *et al.*, 2002) and β-LG can form stable complexes with retinol, suggesting a possible role in the transport of vitamin A (Gaye *et al.*, 1986; Hambling *et al.*, 1992).

In cheese-making process the majority of whey proteins is separated out in the form of whey (Haenlein and Wendorff, 2006).

1.3.3.1. Composition of α-lactalbumin and β-lactoglobulin genes

LAA and *LGB* genes are both localised on ovine chromosome 3 (http://www.animalgenome.org/sheep/maps/; Imam Ghali *et al.*, 1991; Hayes and Petit, 1993;

Hayes *et al.*, 1993). Structural features of DNA and mRNA sequences of these main whey proteins are demonstrated in Table 9.

Table 9. Known structural features of the ovine main whey protein genes (Gaye *et al.*, 1986, 1987; Harris *et al.*, 1988)

Whey protein	DNA (n =)		mRNA (bp)			
	Exons	Introns	Complete	5'-UTR	CDS	3'-UTR
LAA	4[a]	3[a]	723	26	426	271
LGB	7	6	785	39	540	206

[a] = According to GenBank Acc. No. AB052168.

1.3.3.2. Composition of α-lactalbumin and β-lactoglobulin proteins

Ovine α-LA and β-LG primary structures are presented in Table 10. Both amino acid chains form non-phosphorylated proteins, whereas they contain two or three potential phosphorylation and two potential glycosylation sites only in α-LA (Mercier *et al.*, 1978a; Mercier *et al.*, 1978b).

Table 10. Known structural features of the two main whey proteins in sheep milk (Mercier *et al.*, 1978a; Mercier *et al.*, 1978b; Gaye *et al.*, 1986, 1987)

Whey protein	Amino acids[a]	Amino acids[b]	P sites[c]
α-LA	123	19	2/0
β-LG	162	18	3/0

[a] = Number of amino acid residues of the mature protein chain of the protein.
[b] = Number of amino acid residues of the signal peptide.
[c] = Number of phosphorylation sites (putative/actual).

Also in whey proteins the signal peptides consist of highly conserved amino acids when comparing different species (Hambling *et al.*, 1992). Both α-LA and chick-type lysozyme seem to be diverged from a common ancestor, as they show similar intron-exon organisation and resulting amino acid sequences (McKenzie, 1996). Primary structure of β-LG shows some homology to human serum retinol binding protein, leading to the conclusion that they are evolutionary related (Hambling *et al.*, 1992).

1.3.4. Whey protein polymorphisms

1.3.4.1. α-lactalbumin

1.3.4.1.1. Protein variability

Two α-LA protein patterns A and B were made evident by gel electrophoresis of sheep milk (Schmidt and Ebner, 1972), with α-LA A being the most common allele, while variant B has rarely been detected and seems to be confined to specific breeds (Erhardt, 1989a; Amigo *et*

al., 2000). Protein sequences of these α-LA variants are still unknown and no associations to the published coding sequences of *LAA* (GenBank Acc. No. NM_001009797 and AB052168), which differ in several nucleotides partly leading to differences in the amino acid sequence, were made.

1.3.4.1.2. Genetic variability

Within ovine *LAA* molecular genetic analyses are rare. PCR-SSCP-analysis of exon 1 lead to the identification of three conformational patterns, whereas DNA or mRNA analyses or biochemical characterisations of these patterns were not provided (Bastos, 2001).

1.3.4.2. β-lactoglobulin

1.3.4.2.1. Protein variability

Within ovine β-LG three electrophoretic patterns and their amino acid sequence differences have been described, while β-LG variants A and B, differing in the Tyr20His amino acid exchange (Bell and McKenzie, 1967; Kolde and Braunitzer, 1983) are present in almost all breeds. β-LG C, a subtype of β-LG A, caused by an Arg148Gln amino acid exchange (Erhardt *et al.*, 1989) is rather rare and confined to specific breeds (Table 11) (Erhardt, 1989b; Recio *et al.*, 1997; Amigo *et al.*, 2000).

Table 11. Distribution of the three known protein variants of ovine β-LG and their corresponding allele frequencies estimated by biochemical and molecular genetic analyses (Erhardt, 1989b; Recio *et al.*, 1997; Anton *et al.*, 1999; Giaccone *et al.*, 2000; Elmaci *et al.*, 2006; Jaayid, 2006; Ramos *et al.*, 2009)

Variant	Diffusion	Allele frequency
A	Common[a]	0.35-0.98
B	Common[a]	0.02-0.65
C	Merinos, Serra da Estrela, Precos	0.001-0.18

[a] = In several breeds (different countries and purposes).

1.3.4.2.2. Genetic variability

Sequence differences leading to exchanges within the amino acid sequence and causing the known *LGB* variants *A*, *B*, and *C* were confirmed on molecular genetic level (Ali *et al.*, 1990; Prinzenberg and Erhardt, 1999). The *LGB*B* specific Tyr20His amino acid exchange is due to a T>C-SNP in exon 2 (Harris *et al.*, 1988; Ali *et al.*, 1990), while the Arg148Gln amino acid exchange of *LGB*C* is caused by a G>A-transition within exon 5 and is linked to a further mutation in intron 5 (Prinzenberg and Erhardt, 1999).

1.3.5. Importance of milk protein polymorphisms on production traits

The genetic polymorphisms of milk proteins are of importance as associations to quantitative and qualitative parameters in milk are described especially with regard to milk protein composition affecting technological properties of milk in cattle and goats (Ikonen *et al.*, 1999; Martin *et al.*, 2002; Boettcher *et al.*, 2004) and are proposed to be or are already included in breeding strategies (Manfredi *et al.*, 1995; Sanchez *et al.*, 2005). Also in sheep the objective to include molecular genetic markers in selection to improve milk production and composition was mentioned (Árnyasi *et al.*, 2009). However, association studies between milk protein polymorphisms and milk performance traits revealed controversial results in sheep (Amigo *et al.*, 2000; Barillet *et al.*, 2005), partly no influence of milk protein polymorphisms on milk production traits were found (Staiger *et al.*, 2010). Mostly single gene influences were analysed and no casein haplotypes were included in association studies until now, being almost standard in cattle or goat (Sacchi *et al.*, 2005; Heck *et al.*, 2009).

Ovine milk containing *CSN1S1* genotype *CC* showed a higher protein and/or fat content than *AC*, *CD*, *DD*, or *CX* milk (Chianese *et al.*, 1997; Pirisi *et al.*, 1999; Mroczkowski *et al.*, 2004; Wessels *et al.*, 2004). Therefore, *CSN1S1 CC* milk had better renneting properties, and better cheese-making characteristics than *CD* and *DD* milk (Chianese *et al.*, 1997; Pirisi *et al.*, 1999). *CSN1S2* genotype *AB* (nomenclature after Chessa *et al.*, 2003) was significantly advantageous in comparison to *AA* in German East Friesian Dairy sheep for milk and fat yield (Wessels *et al.*, 2004). Type 1 of the five β-CN electrophoretic patterns of Chianese *et al.* (1995) was associated with higher protein content, dry matter, and non-fatty solids than type 2 (Caio *et al.*, 2007). Animals with *CSN2 GG* (Chessa *et al.*, 2010) showed an increased milk yield in comparison to those with *CSN2 AA* genotype, whereas the latter ones were characterised by an increase in fat and protein content (Corral *et al.*, 2010). Corral *et al.* (2010) identified an association between the *CSN3* microsatellite and milk yield and composition in Merino sheep breed, whereas microsattelite genotypes K_1K_4 and K_3K_3 were correlated with the highest milk yield.

Mostly a favourable effect of *LGB BB* genotype on milk yield (Bolla *et al.*, 1989; Rampilli *et al.*, 1997; Corral *et al.*, 2010) was determined, whereas *LGB AA* genotyped ewes produced the highest fat and protein percentages (Wessels *et al.*, 2004; Ramos *et al.*, 2009) and higher cheese yield (Rampilli *et al.*, 1997). Contrary fat or protein content were highest in milk with *LGB BB* in other studies (Giaccone *et al.*, 2000; Mroczkowski *et al.*, 2004) and in some cases no associations between *LGB* polymorphisms and milk production traits were identified (Recio *et al.*, 1997; Mele *et al.*, 2007).

Due to the effects of milk protein polymorphisms on milk yield and composition they influence also lamb growth (Snowder and Glimp, 1991; Amigo *et al.*, 2000; Moroni *et al.*, 2007). In beef cattle for example, *CSN3* polymorphisms showed association to the weight gain from birth to weaning (Moody *et al.*, 1996).

In addition, milk protein polymorphisms can be used for molecular tracing of typical cheeses (Chianese *et al.*, 2009), and for species identification of milk or dairy products (Addeo *et al.*, 1990; Otaviano *et al.*, 2008; Reale *et al.*, 2008). Testing for milk protein polymorphisms may aid in identifying admixtures of goat and cow milk in sheep milk products, where this is either of commercial concern or prohibited (Rodriguez *et al.*, 1990; Anguita *et al.*, 1997; Haenlein and Wendorff, 2006).

Furthermore, milk protein polymorphisms can influence functional quality of milk and of dairy products, concerning bioactive peptides (Weimann *et al.*, 2009) or negative allergenic epitopes, derived from proteolysis of milk proteins (Chatchatee *et al.*, 2001; Busse *et al.*, 2002; Park *et al.*, 2007; Recio *et al.*, 2009).

1.4. Aims of the study

In sheep milk protein variability is often described on the basis of electrophoretic patterns but without further characterisation. Thus, there is a lack of protein analyses for example within α_{s1}-CN patterns B, F, and X or for both known α-LA protein forms. Additionally, molecular genetic analysis of ovine milk protein polymorphisms is fragmentary in comparison to other ruminant species. Often only parts of the DNA are amplificated by PCR and scanned for polymorphisms, a more complete knowledge of the degree of variability in ovine milk protein genes is missing but desirable. In order to get an overview of the genetic potential within sheep and to identify possibly further variability within β- and κ-CN, it is also necessary to include more breeds in studies on milk protein polymorphisms. Nomenclature is inconsistent within several protein fractions of ovine milk, whereas its standardisation is important for further allele, genotype, and haplotype studies. Subsequently, association studies concerning milk performance traits should be expanded.

Therefore, the aims of this study were i) to describe milk protein variability in different sheep breeds with different purposes by isoelectric focusing; ii) to identify reasons for protein variability on molecular genetic level; iii) to develop DNA-based tests for several casein variants; iv) to uniform nomenclature; v) to establish casein haplotype studies and with this vi) to create conditions for association studies.

1.5. Study design

Isoelectric focusing was chosen for screening milk protein variants in milk samples from 1,078 animals of six different sheep breeds. Pedigree information were available to identify inheritance of milk protein polymorphisms.

For continuative biochemical clarification of α_{s1}-CN protein sequences of selected milk samples, cooperation with an Italian working group of the University of Naples was initiated.

Additionally, molecular genetic analyses using extracted mRNA from milk and DNA from milk and blood were planned for *CSN1S1* and *CSN1S2* to complete characterisation of variability and to standardise nomenclature within these milk proteins.

PCR-RFLP-technique and fragment length analysis were choosen as basis for the establishement of DNA-based tests to enable typing of rams and non-lactating females.

Milk recording data of 76 East Friesian Dairy sheep were available to start association studies regarding milk protein genotypes and casein haplotypes.

1.6. Structure of the thesis

The general introduction (Chapter 1) demonstrates the importance of dairy sheep production and breeding and the milk performance traits of sheep. Ovine milk protein fractions and their variability are presented as well as the aims of the study and the study design.

Chapter 2 describes IEF of milk samples and the resulting allele, genotype, and haplotype frequencies of the milk proteins α_{s1}-, α_{s2}-, and κ-CN, α-LA and β-LG. IEF lead to the identification of, until now, undescribed protein patterns within ovine caseins, causing new genotypes and haplotypes. The following chapters concentrate on the deeper analysis of these new casein variants.

Chapter 3 specifies further protein analyses and molecular genetic characterisation of α_{s1}-CN allele I and the detection of skipping of exon 7 within this ovine casein variant. Alternative splicing is caused by a sequence difference affecting the 5'-donor splice site of intron 7.

Chapter 4 focuses mainly on the molecular genetic characterisation of α_{s1}-CN H, a further milk protein variant caused by alternative splicing. In this *CSN1S1* allele exon 8 is skipped due to both an insertion and a deletion within exon 8/intron 8 DNA sequence.

Chapter 5 demonstrates detailed molecular genetic analyses of α_{s2}-CN C and D as well as the identification of further variability within ovine *CSN1S2*. A possible evolutionary pathway of *CSN1S2* in sheep and a standardised nomenclature of *CSN1S2* polymorphisms were proposed.

Chapter 6 describes association studies between milk protein polymorphisms and production data in East Friesian Dairy sheep. Several correlations were detected between milk yield

and/or composition and α_{s1}-CN, α_{s2}-CN, and β-LG alleles, genotypes and α_{s1}-/α_{s2}-CN-haplotypes.

At last, chapter 7 completes the whole thesis with a general discussion on all themes covered and analysed, which are summarised within the chapters 8 and 9.

2. Isoelectric focusing reveals additional casein variants in German sheep breeds

I. J. Giambra, S. Jäger, and G. Erhardt

Department of Animal Breeding and Genetics, Justus-Liebig-University, 35390 Gießen, Germany

Published in:
Small Ruminant Research 90 (1-3), 11-17 (2010)

Submitted: 23.07.2009
Accepted: 07.12.2009

ABSTRACT

Isoelectric focusing (IEF) was applied for screening milk protein variants in milk samples from altogether 1078 sheep of different breeds, in detail Black Faced Mutton sheep (SKF; $n = 57$), East Friesian Milk sheep (OMS; $n = 254$), Gray Horned Heath (GGH; $n = 190$), Merinoland sheep (MLS; $n = 363$), Merino Mutton sheep (MMS; $n = 88$), and Rhön sheep (RHO; $n = 126$). Besides the known genetic variants of α_{s1}-casein (*CSN1S1*) (*A*, *C*, *D*), α_{s2}-casein (*CSN1S2*) (*A*, *B*), and β-lactoglobulin (*LGB*) (*A*, *B*, *C*) additional variants could be demonstrated in *CSN1S1* (*H*, *I*) and *CSN1S2* (*C*, *D*) and their genetic control confirmed by segregation analyses. *CSN1S1*H* corresponds to a previously mentioned phenotype "X" occurring in OMS, whereas *CSN1S1*I* was identified for the first time in GGH. *CSN1S2*C* appeared in OMS, GGH, MLS, and RHO in low frequencies and *CSN1S2*D* in MLS. Within *LGB* all three alleles occurred in Merino breeds while α-lactalbumin (*LAA*) and κ-CN (*CSN3*) were monomorph at protein level. The haplotype *CSN1S1*C–CSN1S2*A* was predominant in five out of six breeds with frequencies between 0.325 and 0.919.

Keywords: Isoelectric focusing; Milk protein; Sheep; Haplotype

1. Introduction

Caseins (CN) in sheep milk form up to 85% of the protein content (Mercier et al., 1978) distributed in the four fractions α_{s1}-CN, α_{s2}-CN, β-CN, and κ-CN, and encoded by *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*. The two whey proteins, α-lactalbumin and β-lactoglobulin, are encoded by *LAA* and *LGB*, and account for the main part of the remaining 15% of milk protein content in sheep milk (Mercier et al., 1978). Among small ruminants, goats have been thoroughly investigated for milk protein genes and noticeable genetic variation has been identified, whereas the knowledge of milk protein genetic variants is more fragmentary in ovine species (Amigo et al., 2000; Moioli et al., 2007).

So far, seven phenotypes (A, B, C, D, E, F, "X") of α_{s1}-CN in ovine milk have been identified by protein electrophoresis (Chianese et al., 1996; Pirisi et al., 1999; Wessels et al., 2004), whereas primary structures have been determined only for CSN1S1 A, C, D (Ferranti et al., 1995), and E (Chianese et al., 2007). While α_{s1}-CN C occurs in all sheep breeds with frequencies from 0.485 in Sarda up to 0.890 in Segureña (Chianese et al., 1996; López-Gálvez et al., 1999; Amigo et al., 2000), CSN1S1 E and F were identified so far at low frequencies in Italian breeds only (Chianese et al., 1996; Pirisi et al., 1999).

Boisnard et al. (1991) described three types of *CSN1S2* cDNA without further characterisation at protein level, while Chianese et al. (1993) identified three phenotypes, differing in electrophoretic behaviour. Chessa et al. (2003) called two phenotypes identified by IEF CSN1S2 A and B, which were recently characterised by amino acid exchanges at positions 75 and 105 of mature protein (Picariello et al., 2009).

Within ovine CSN3 at protein level no variation could be identified (Moioli et al., 1998; Chessa et al., 2003). This is in contrast to the high degree of polymorphisms described at *CSN3* in goat (Jann et al., 2004; Prinzenberg et al., 2005) and cattle (Prinzenberg et al., 2008).

For ovine *LGB* three genetic variants and their sequence differences have been described, whereas *LGB* variants *A* and *B* (Bell and McKenzie, 1967) show widespread distribution and *LGB*C* (Erhardt, 1989b; Erhardt et al., 1989; Recio et al., 1997) seems to be specific for Merino breeds. The whey protein LAA is distinguished in two phenotypes A and B (Schmidt and Ebner, 1972), however LAA variant B seems to be rare and confined to specific breeds (Erhardt, 1989a; Amigo et al., 2000). Protein and/or DNA sequence differences of these LAA variants are still unknown.

The genetic polymorphisms of milk proteins are of importance as associations to quantitative and qualitative parameters in milk are described especially with regard to milk protein composition affecting technological properties of milk in cattle and goats (Rampilli et al., 1997; Martin et al., 2002; Mroczkowski et al., 2004; Caroli et al., 2009; Heck et al., 2009), while in sheep the results are controversial (Amigo et al., 2000; Barillet et al., 2005; Barillet, 2007). In goat selection for advantageous alleles is already used in breeding strategies (Sanchez et al., 2005). In addition milk protein polymorphisms are valuable markers for population and evolutionary studies, too (Mahé et al., 1999; Ibeagha-Awemu et al., 2007; Chessa et al., 2008; Küpper et al., 2010).

Casein genes are organized as a tightly linked cluster on ovine chromosome 6 in a 250 kb DNA segment (Threadgill and Womack, 1990; Lévéziel et al., 1991; Bevilacqua et al., 2006) and therefore the estimation of the relationship between casein variants and milk production traits can be improved by considering the entire casein haplotype instead of single gene typing (Sacchi et al., 2005; Heck et al., 2009). Therefore, it is important to know the variability of caseins (Ordás, 2001) and it is to consider that the detection of further variants leads to further haplotypes in the casein cluster with potentially different effects.

In this context it could be clearly demonstrated in the past that IEF is an effective method for studying genetic variants in milk proteins. It resulted in the detection of further alleles in

different species (Erhardt, 1989a, 1996; Baranyi et al., 1993; Erhardt et al., 2002) and was therefore the basis for their further characterisation at protein and DNA level.

The importance of going deeper into the knowledge on milk protein polymorphisms in sheep is evident. Consequently it was the aim of this study to analyse milk protein polymorphisms in sheep breeds of different purposes by IEF to present current knowledge on milk protein genetic variability in sheep.

2. Materials and methods

2.1. Milk samples

A total of 1078 individual milk samples were collected from the following six different sheep breeds kept in Germany: Black Faced Mutton sheep (SKF; n = 57), East Friesian Milk sheep (OMS; n = 254), Gray Horned Heath (GGH; n = 190), Merinoland sheep (MLS; n = 363), Merino Mutton sheep (MMS; n = 88), and Rhön sheep (RHO; n = 126). In addition for segregation analyses informative families were available at Research Station "Oberer Hardthof" and in private flocks.

2.2. Isoelectric focusing

Separation and identification of milk proteins were done by IEF of skimmed milk samples according to Erhardt (1989c) in 0.3 mm thin polyacrylamide gels using carrier ampholytes. The modified gel (T = 4.98; C = 3.75; 7.05 M urea) contained 0.7 mL of the following mixture of carrier ampholytes: 0.86% (w/v) Servalyte pH 3.0–5.0 (Serva Electrophoresis, Heidelberg); 0.65% (w/v) Pharmalyte pH 4.2–4.9 (GE Healthcare Europe GmbH, Freiburg); 0.86% (w/v) Pharmalyte pH 4.5–5.4 (GE Healthcare Europe GmbH, Freiburg) and 0.57% (w/v) Servalyte pH 4.0–6.0 (Serva Electrophoresis, Heidelberg). After fixation and staining phenotypes were manually scored using sheep milk samples with known CSN1S1, CSN1S2, CSN3, LAA, and LGB phenotypes as reference samples. Nomenclature for CSN1S2 was done according to Chessa et al. (2003).

2.3. Statistical analysis

Allele and genotype frequencies were calculated with program PopGene V 1.31 (Yeh et al., 1997). A chi-square (χ^2) test was performed to test the goodness of fit to Hardy–Weinberg equilibrium expectations for the distribution of genotypes.

Haplotype frequencies for *CSN1S1* and *CSN1S2* and the occurrence of linkage disequilibrium were estimated with EH software (Xie and Ott, 1993), considering only the alleles with frequencies > 0.05 in minimum in one breed.

Pedigree data were analysed with Pedigree-Viewer Version 5.3 (Kinghorn and Kinghorn, 2005) and CFC Release 1.0 (Sargolzaei et al., 2006).

3. Results

3.1. Isoelectric focusing

Separation of CSN1S1, CSN1S2, CSN3, LAA, and LGB phenotypes in skimmed milk samples by IEF is shown in Fig. 1. In the samples analysed we could demonstrate phenotypes with CSN1S1 A, C, and D, but could neither differentiate between CSN1S1 B and C nor identify phenotypes with E and/or F. On the other side phenotypes with CSN1S1 "X" were identified and designated now as CSN1S1 H. The electrophoretic pattern of α_{s1}-CN H is characterised by a more alkaline isoelectric point (pI) in comparison to CSN1S1 A and C, but a more acidic one compared to D. Depending on the protein content of the sample, it is possible that the minor bands of CSN1S1 H are not visible and phenotyping has to be done on the basis of the major bands. Additional phenotypes with pattern belonging to the CSN1S1 fraction, were named CSN1S1 I and occurred in homozygous and heterozygous form (Fig. 1). CSN1S1 I showed the most alkaline pI in comparison to CSN1S1 A, C, D, and H.

As demonstrated in Fig. 1 the CSN1S2 fraction is located with the variants *CSN1S2*A* and *B* in the more alkaline region of the gel below CSN3 and LAA and extends into the area of LGB. Additionally two new patterns were separated and named CSN1S2 C and D. Phenotypes with CSN1S2 A, B, C, and D were distinguished by their different pI in the order C, D, A, and B with the most alkaline one. Although the main bands of CSN1S2 D are located very close to the ones of C, different phenotypes with these alleles are clearly identified by IEF and confirmed by molecular analyses (not shown).

Within CSN2 we could not differentiate alleles because we did not find differences in pI, main difference within CSN2 is caused by relative intensity of the β-CN bands.

While CSN3, as well as LAA showed identical pattern in the milk samples analysed, LGB occurred in different phenotypes in the analysed skimmed milk samples (Fig. 1).

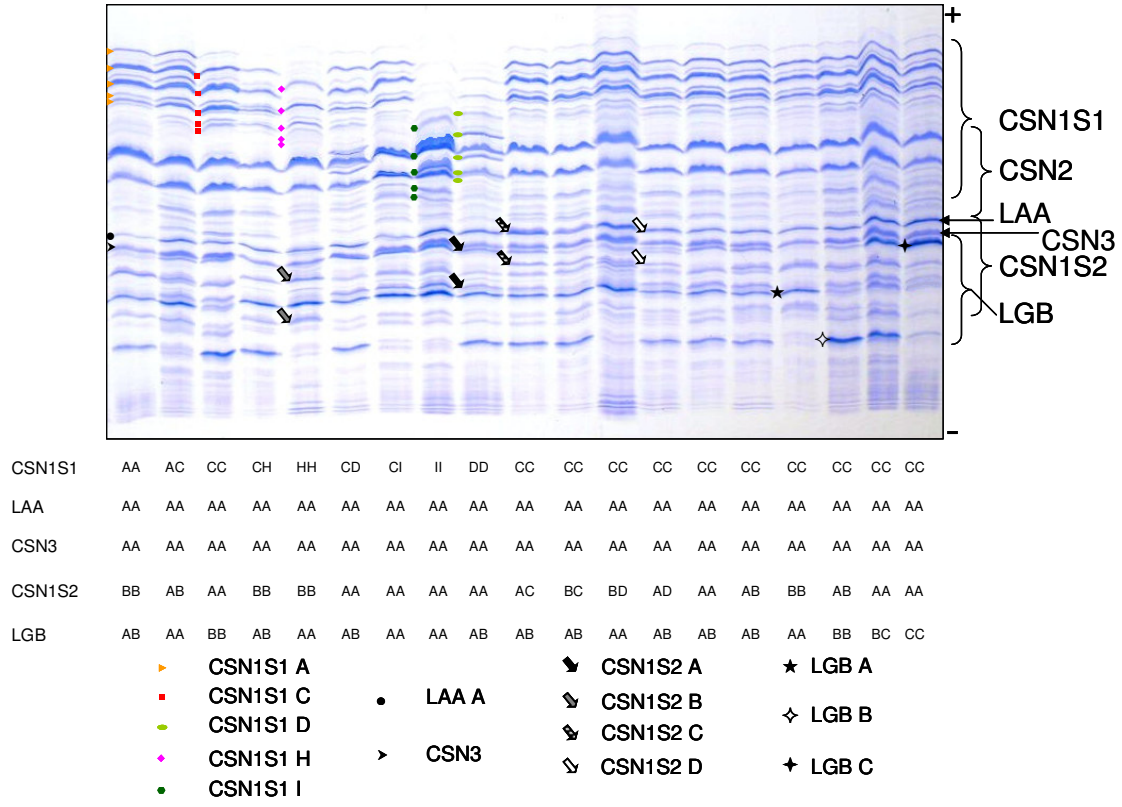


Fig. 1. Pattern of ovine milk samples after IEF representing different genotypes within *CSN1S1*, *CSN1S2*, and *LGB*.

3.2. Family studies

Transmission of *CSN1S1*H* could be reconstructed within informative families with 20 daughters, descending from four rams most probably *CSN1S1 CH*, where 14 carried *CSN1S1*H* (χ^2 = 2.11; degrees of freedom (df) = 4). In addition *CSN1S1*H* was transferred within three dam-daughter pairs.

Transmission of *CSN1S1*I* could be demonstrated within two informative families.

Within two informative half-sib families the *CSN1S2*C* allele was transferred to 7 out of 14 daughters assuming that the rams were heterozygous *CSN1S2 AC* (χ^2 = 1.97; df = 2).

In a large half-sib family where the dams were non-carrier of *CSN1S2*D* 8 of 14 daughters got *CSN1S2*D* from the ram (χ^2 = 0.4; df = 3; assuming that the rams genotype is *CSN1S2 AD*).

In all cases the observed ratio does not deviate significantly from the expected 1:1 ratio and confirms codominant autosomal inheritance of *CSN1S1*H* and *I* as well as from *CSN1S2*C* and *D*.

3.3. Allele and genotype frequencies

The estimated allele frequencies are demonstrated in Table 1. At *CSN1S1* RHO was monomorphic while in the other breeds up to three alleles occurred. *CSN1S1*C* showed the highest frequency in all breeds analysed, while *CSN1S1*A* and *CSN1S1*D* are rare, and could only be demonstrated in two (OMS and MLS) respectively in four of the six breeds analysed. On the other side *CSN1S1*H* seems to be a private allele for OMS as well as *CSN1S1*I* for GGH, which was demonstrated for the first time and occurred in a frequency of 0.029.

Within *CSN1S2* only two alleles could be demonstrated in MMS and SKF, while MLS was characterised by the occurrence of four alleles. *CSN1S2*A* showed highest frequency in GGH, MLS, as well as in MMS and RHO, while in SKF and OMS variant *B* was predominant. From the two variants *CSN1S2*C* and *D*, demonstrated for the first time, *CSN1S2*C* occurred in OMS, GGH, MLS, and RHO with low frequencies as well as *CSN1S2*D* in MLS.

Besides *CSN3* also the whey protein *LAA* showed no variation. *LGB*A* and *B* occurred in all breeds and *LGB*C* in addition in both Merino breeds (MLS and MMS) at low frequency.

Within genotype frequencies *CSN1S1 CC*, *CSN1S2 AA*, *LGB AA* and *AB* showed highest frequencies in most breeds, with exception within SKF and OMS, where *CSN1S2 AB* was predominant. There was a good agreement between the observed genotype frequencies and those expected on the basis of Hardy–Weinberg equilibrium in all milk proteins with exception of *CSN1S1* in GGH and *LGB* in RHO.

3.4. Casein haplotype frequencies

Only *CSN1S1–CSN1S2* haplotypes were evaluated, because *CSN3* was monomorphic in all breeds. *CSN1S1–CSN1S2*-haplotype frequencies are shown in Table 2, also indicating the expected frequencies under the hypothesis of independence.

On the basis of the expected nine haplotypes most of the breeds showed three of them with a frequency >0.01 with exception of OMS, where four different haplotypes occurred. Haplotype *CA* (in the order: *CSN1S1*C–CSN1S2*A*) was dominant in most breeds while haplotype *CB* prevailed in SKF. The distribution and frequency of the other haplotypes varied between the analysed breeds, whereby haplotype *HB* occurred only in OMS.

The χ^2 test showed a highly significant association between *CSN1S1* and *CSN1S2* locus in OMS ($P<0.001$) and over all breeds ($P<0.001$).

Table 1. Distribution of *CSN1S1*, *CSN1S2*, *LAA* and *LGB* allele frequencies, identified by IEF, in different sheep breeds.

Breed	*n*	*CSN1S1*					*CSN1S2*				*LAA*		*LGB*		
		A	*C*[a]	*D*	*H*	*I*	*A*	*B*	*C*	*D*	*A*	*B*	*A*	*B*	*C*
SKF	57	-	0.903	0.097	-	-	0.421	0.579	-	-	1.000	-	0.684	0.316	-
OMS	254	0.004	0.913	-	0.083	-	0.474	0.514	0.012	-	1.000	-	0.758	0.242	-
GGH	190	-	0.963	0.008	-	0.029	0.926	0.026	0.048	-	1.000	-	0.668	0.332	-
MLS	363	0.035	0.924	0.041	-	-	0.786	0.179	0.007	0.028	1.000	-	0.536	0.317	0.147
MMS	88	-	0.966	0.034	-	-	0.909	0.091	-	-	1.000	-	0.801	0.193	0.006
RHO	126	-	1.000	-	-	-	0.730	0.175	0.095	-	1.000	-	0.567	0.433	-

[a] According to Chessa et al. (2003) *CSN1S1*B* and *C* were not differentiated by IEF-typing.

Table 2. *CSN1S1–CSN1S2*-haplotype frequencies for different sheep breeds.

Haplotype[a]	SKF (*n* = 57)		OMS* (*n* = 252)		GGH (*n* = 180)		MLS (*n* = 318)		MMS (*n* = 88)		RHO (*n* =126)		Over all breeds* (*n* = 1021)	
	Exp.[b]	Obs.[c]	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
CA	0.380	0.325	0.435	0.474	0.920	0.919	0.801	0.794	0.878	0.875	0.730	0.730	0.701	0.710
CB	0.523	0.579	0.471	0.431	0.028	0.028	0.149	0.156	0.088	0.091	0.175	0.175	0.232	0.221
CC	-[d]	-	0.011	0.012	0.050	0.050	0.008	0.008	-	-	0.095	0.095	0.025	0.026
DA	0.041	0.096	-	-	0.003	0.003	0.036	0.042	0.031	0.034	-	-	0.016	0.022
DB	0.056	<0.000	-	-	<0.000	Not obs.[e]	0.007	<0.000	0.003	Not obs.	-	-	0.005	<0.000
DC	-	-	-	-	<0.000	Not obs.	<0.000	Not obs.	-	-	-	-	0.001	Not obs.
HA	-	-	0.040	<0.000	-	-	-	-	-	-	-	-	0.015	<0.000
HB	-	-	0.043	0.083	-	-	-	-	-	-	-	-	0.005	0.021
HC	-	-	0.001	Not obs.	-	-	-	-	-	-	-	-	0.001	Not obs.

* Significant (P<0.001) association between *CSN1S1* and *CSN1S2*.

[a] For haplotype evaluation, alleles with frequencies < 0.05 in all breeds were not considered.

[b] Expected haplotype frequency under independence hypothesis.

[c] Observed haplotype frequency.

[d] Minimum one of the concerned alleles did not occur.

[e] Not observed.

4. Discussion

Analysis of ovine milk proteins in six different sheep breeds using IEF separated the known *CSN1S1* variants *A*, *C*, and *D*, whereas *CSN1S1*C* was predominant in all breeds. This is in agreement with studies in Italian, German, and Polish breeds (Chianese et al., 1996; Pirisi et al., 1999; Mroczkowski et al., 2004; Wessels et al., 2004). The high frequency of *CSN1S1*C* is most probably the result of indirect selection in the past, as *CSN1S1*C* showed correlation with higher total protein and casein content (Pirisi et al., 1999; Amigo et al., 2000). On the other side *CSN1S1*D*, which showed negative effects on milk composition and cheese yield (Pirisi et al., 1999) could not be demonstrated in OMS, a breed highly selected for milk production. *CSN1S1*H* (formerly α_{s1}-CN "X"; Wessels et al., 2004) is clearly demonstrated by IEF and seems to be specific for OMS as in Italian milk sheep breeds also characterised by the same technique in a comparable pH range it was not identified (Chianese et al., 1996; Pirisi et al., 1999). Samples of animals with *CSN1S1*H* are also associated with a reduced protein expression level of this protein (Giambra et al., in press-a) and this results in a reduced intensity of the bands, especially the minor bands. The demonstration of *CSN1S1*I* in GGH confirms the possibility to identify new variants by IEF. Partly *CSN1S1*I* allele may be difficult to identify as some of the bands are close related to the position of *CSN2*, whereas in *CSN1S1*I* heterozygous samples the intensity of the bands of the other *CSN1S1* allele gave a hint that the sample is heterozygous. Furthermore, the close related positions of *CSN1S1*I* and *CSN2* are leading to a stronger focusing of the first two main bands of *CSN2* (Fig. 1).

It could be excluded that *CSN1S1*H* and *I* are identical to phenotype E or F because latter one showed the most acidic pI of all known variants (Pirisi et al., 1999) and CSN1S1 E has a higher pI compared with H and I (Lina Chianese, personal communication). In addition molecular genetic analyses confirmed the identity of the new CSN1S1 variants (Giambra et al., in press-b). A differentiation between α_{s1}-CN B and C by IEF was not possible, which is in agreement with Chessa et al. (2003). Additional biochemical and molecular genetic data are necessary to confirm that CSN1S1 B and C are different variants.

Within *CSN1S2* variants *A* and *B* occurred in all breeds and the high frequency of *CSN1S1*A* in most breeds is in accordance with observations in the Italian breeds Gentile di Puglia, Sarda, and Comisana (Chessa et al., 2003; Picariello et al., 2009). In addition two further variants could be demonstrated for the first time and their genetic control confirmed. The nomination of the variant with the most acidic pI was *CSN1S2*C* followed by *CSN1S2*D* with a pI more alkaline than *C*, but more acidic in comparison to *A* and *B*. Further biochemical and molecular characterisation of these variants are in progress to identify the

mutations responsible for the different pI and to compare them with the cDNA-sequences described by Boisnard et al. (1991).

The high resolution power of the IEF could be confirmed as it was possible to demonstrate besides the known alleles additional variants during routinely screening of skimmed milk without isolation of the caseins as it was necessary by Chessa et al. (2003) and Picariello et al. (2009). Therefore, this method can be effectively used in screening further sheep breeds. Within CSN2 we identified only differences in the intensity of the bands. This is comparable to studies of Chianese et al. (1995), where the different intensities of the β-casein bands were mainly related to the degree of phosphorylation.

Codominant autosomal inheritance of the new variants *CSN1S1*H* and *I* as well as *CSN1S2*C* and *D* was confirmed by segregation studies.

CSN3 was monomorphic after IEF in all samples which is in accordance to Chessa et al. (2003). Only at molecular level sequence differences are demonstrated resulting in synonymous and non-synonymous mutations (Ceriotti et al., 2004; Feligini et al., 2005).

While *LAA* was monomorphic *LGB* phenotypes, with the three known alleles, could be clearly differentiated within the skimmed milk samples analysed. *LGB*A* and *B* showed a widespread distribution in all breeds, while *LGB*C* was specific for MLS and MMS. This confirms the results of Erhardt (1989a) and Recio et al. (1997), that *LGB*C* is specific for Merinos and related breeds.

Hardy–Weinberg equilibrium was demonstrated with exception of *CSN1S1* in GGH and *LGB* in RHO. As both breeds were endangered in the past genetic drift cannot be excluded.

The linkage equilibrium between *CSN1S1* and *CSN1S2* in most analysed breeds is comparable with studies in goat (Finocchiaro et al., 2008). The *CSN1S1–CSN1S2* haplotype linkage disequilibrium within OMS is mainly caused by linkage between *CSN1S1*H* and *CSN1S2*B*.

Based on the pattern after IEF it can be postulated that the new variants in *CSN1S1* and *CSN1S2* are the results of protein electric charge, and could affect milk properties both from the cheese making and nutritional point of view. Further investigations are needed to evaluate the influence of the casein as well as the haplotype variability on quality and properties of ovine milk and related products (Barillet, 2007). Therefore, investigations of ovine milk protein variability not only at protein but also at the DNA level are recommended to get a more complete picture about the genetic diversity. This could be a valuable basis for an effective approach to identify association to economic traits or for breed specific dairy products followed by integration in breeding like in goat (Sanchez et al., 2005) and cattle (Chessa et al., 2007; Nilsen et al., 2009).

5. Conclusion

Knowledge of variability in sheep casein is still incomplete, confirmed by the finding of several new casein variants by screening milk samples of six breeds by IEF. Therefore, this technique could be widely exploited for typing lactating ewes at milk protein polymorphisms in a first step followed by molecular based methods. This offers the possibility to get a more complete picture about the milk protein genes in sheep and to consider milk protein variants/haplotypes in specific breeding programs with regard to preserve biodiversity and/or to improve dairy sheep breeds for specific milk protein production.

Acknowledgements

The authors thank the staff of Research Station "Oberer Hardthof" of the Department of Animal Breeding and Genetics, especially A. Kaspar, S. Mandler, and R. Zartner, and of Research Station LVA Iden and all private sheep holders, in detail B. and R. Althoff, F. Atema, R. and R. Bergmann, M. Dors, H.-U. Hartmann, Heidschnuckenhof Jeversen, B. Hucke, M. Ise, O. Junker-Matthes, W. Rompf, G. Rüpke, S. Prokasky, E. Prunzel-Ulrich, D. Schnehage, and Verein Schäferhof Neuenkirchen for their support in the collection of milk samples and for allocation of sheep data. Furthermore, we thank the professors A. Caroli and L. Chianese for providing reference samples and for fruitful discussion about nomenclature.

References

Amigo, L., Recio, I., Ramos, M., 2000. Genetic polymorphism of ovine milk proteins: its influence on technological properties of milk–a review. Int. Dairy J. 10, 135–149.

Baranyi, M., Bosze, Z.S., Buchberger, J., Krause, I., 1993. Genetic polymorphism of milk proteins in Hungarian spotted and Hungarian grey cattle: a possible new genetic variant of β-lactoglobulin. J. Dairy Sci. 76, 630–636.

Barillet, F., 2007. Genetic improvement for dairy production in sheep and goats. Small Ruminant Res. 70, 60–75.

Barillet, F., Arranz, J.J., Carta, A., 2005. Mapping quantitative trait loci for milk production and genetic polymorphisms of milk proteins in dairy sheep. Genet. Sel. Evol. 37, S109–S123.

Bell, K., McKenzie, H.A., 1967. The whey proteins of ovine milk: β-lactoglobulins A and B. Biochim. Biophys. Acta 147, 123–134.

Bevilacqua, C., Helbling, J.C., Miranda, G., Martin, P., 2006. Translational efficiency of casein transcripts in the mammary tissue of lactating ruminants. Reprod. Nutr. Dev. 46, 567–578.

Boisnard, M., Hue, D., Bouniol, C., Mercier, J.C., Gaye, P., 1991. Multiple mRNA species code for two non-allelic forms of ovine α_{s2}-casein. Eur. J. Biochem. 201, 633–641.

Caroli, A.M., Chessa, S., Erhardt, G.J., 2009. Invited review: milk protein polymorphisms in cattle: effect on animal breeding and human nutrition. J. Dairy Sci. 92, 5335–5352.

Ceriotti, G., Chessa, S., Bolla, P., Budelli, E., Bianchi, L., Duranti, E., Caroli, A., 2004. Single nucleotide polymorphisms in the ovine casein genes detected by polymerase chain reaction-single strand conformation polymorphism. J. Dairy Sci. 87, 2606–2613.

Chessa, S., Bolla, P., Dario, C., Pieragostini, E., Caroli, A., 2003. Polimorfismi genetici lattoproteici nella razza ovina gentile di puglia: Monitoraggio mediante focalizzazione isoelectrica. Sci. Tecn. Latt.-Cas. 54, 1–8.

Chessa, S., Chiatti, F., Ceriotti, G., Caroli, A., Consolandi, C., Pagnacco, G., Castiglioni, B., 2007. Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms. J. Dairy Sci. 90, 451–464.

Chessa, S., Rignanese, D., Küpper, J., Pagnacco, G., Erhardt, G., Caroli, A., 2008. Short communication: the beta-casein (CSN2) silent allele C1 is highly spread in goat breeds. J. Dairy Sci. 91, 4433–4436.

Chianese, L., Garro, G., Ferranti, P., Malorni, A., Addeo, F., Rabasco, A., Pons, P.M., 1995. Discrete phosphorylation generates the electrophoretic heterogeneity of ovine β-casein. J. Dairy Res. 62, 89–100.

Chianese, L., Caira, S., Garro, G., Addeo, F., 2007. Primary structure of ovine deleted variant α_{s1}-CN E. In: Proceedings of the 5th International Symposium on the Challenge to Sheep and Goats Milk Sectors, Alghero/Sardinia, Italy, p. 70.

Chianese, L., Garro, G., Addeo, F., Lopez-Galvez, G., Ramos, M., 1993. Discovery of an ovine α_{s2}-casein variant. J. Dairy Res. 60, 485–493.

Chianese, L., Garro, G., Mauriello, R., Laezza, P., Ferranti, P., Addeo, F., 1996. Occurrence of five α_{s1}-casein variants in ovine milk. J. Dairy Res. 63, 49–59.

Erhardt, G., 1989a. Evidence for a third allele at the β-lactoglobulin (β-Lg) locus of sheep milk and its occurrence in different breeds. Anim. Genet. 20, 197–204.

Erhardt, G., 1989b. Genetic polymorphisms of α-lactalbumin and β-lactoglobulin in sheep milk. Anim. Genet. 20, 76–77.

Erhardt, G., 1989c. K-Kaseine in Rindermilch–Nachweis eines weiteren Allels (k-Cn^{E}) in verschiedenen Rassen. J. Anim. Breed. Genet. 106, 225–231.

Erhardt, G., 1996. Detection of a new kappa-casein variant in milk of Pinzgauer cattle. Anim. Genet. 27, 105–107.

Erhardt, G., Godovac-Zimmermann, J., Conti, A., 1989. Isolation and complete primary sequence of a new ovine wild-type β-lactoglobulin C. Biol. Chem. Hoppe-Seyler 370, 757–762.

Erhardt, G., Jäger, S., Budelli, E., Caroli, A., 2002. Genetic polymorphism of goat α_{s2}-casein (*CSN1S2*) and evidence for a further allele. Milchwissenschaft 57, 137–140.

Feligini, M., Vlaco, S., Curik, V.C., Parma, P., Greppi, G., Enne, G., 2005. A single nucleotide polymorphism in the sheep κ-casein coding region. J. Dairy Res. 72, 317–321.

Ferranti, P., Malorni, A., Nitti, G., Laezza, P., Pizzano, R., Chianese, L., Addeo, F., 1995. Primary structure of ovine α_{s1}-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. J. Dairy Res. 62, 281–296.

Finocchiaro, R., Hayes, B.J., Siwek, M., Spelman, R.J., van Kaam, J.B.C.H.M., Adnoy, T., Portolano, B., 2008. Comparison of casein haplotypes between two geographically distant European dairy goat breeds. J. Anim. Breed. Genet. 125, 68–72.

Giambra, I.J., Chianese, L., Ferranti, P., Erhardt, G., in press-a. Short communication: molecular genetic characterization of ovine α_{s1}-casein allele H caused by alternative splicing. J. Dairy Sci.

Giambra, I.J., Chianese, L., Ferranti, P., Erhardt, G., in press-b. Genomics and proteomics of deleted ovine *CSN1S1*I*. Int. Dairy J.

Heck, J.M., Schennink, A., van Valenberg, H. J., Bovenhuis, H., Visker, M.H., van Arendonk, J.A., van Hooijdonk, A.C., 2009. Effects of milk protein variants on the protein composition of bovine milk. J. Dairy Sci. 92, 1192–1202.

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

Jann, O.C., Prinzenberg, E.M., Luikart, G., Caroli, A., Erhardt, G., 2004. High polymorphism in the κ-casein (*CSN3*) gene from wild and domestic caprine species revealed by DNA sequencing. J. Dairy Res. 71, 188–195.

Kinghorn, B., Kinghorn, S., 2005. Pedigree-Viewer. 5.3., http://www.personal.-une.edu.au/~bkinghor/pedigree.htm.

Küpper, J., Chessa, S., Rignanese, D., Caroli, A., Erhardt, G., 2010. Divergence at the casein haplotypes in dairy and meat goat breeds. J. Dairy Res. 77, 56–62.

Lévéziel, H., Metenier, L., Guerin, G., Cullen, P., Provot, C., Bertaud, M., Mercier, J.C., 1991. Restriction fragment length polymorphism of ovine casein genes: close linkage between the α_{s1}-, α_{s2}-, β- and κ-casein loci. Anim. Genet. 22, 1–10.

López-Gálvez, G., Amigo, L., Ramos, M., 1999. Polymorphism of α_s-caseins in the milk of two Spanish ovine breeds. Milchwissenschaft 54, 17–19.

Mahé, M.F., Miranda, G., Queval, R., Bado, A., Zafindrajaona, P.S., Grosclaude, F., 1999. Genetic polymorphism of milk proteins in African Bos taurus and Bos indicus populations. Characterization of variants α_{s1}-Cn H and κ-Cn J. Genet. Sel. Evol. 31, 239–253.

Martin, P., Szymanowska, M., Zwierzchowski, L., Leroux, C., 2002. The impact of genetic polymorphisms on the protein composition of ruminant milks. Reprod. Nutr. Dev. 42, 433–459.

Mercier, J.C., Haze, G., Gaye, P., Hue, D., 1978. Amino terminal sequence of the precursor of ovine β-lactoglobulin. Biochem. Biophys. Res. Commun. 82, 1236–1245.

Moioli, B., D' Andrea, M., Pilla, F., 2007. Candidate genes affecting sheep and goat milk quality. Small Ruminant Res. 68, 179–192.

Moioli, B., Pilla, F., Tripaldi, C., 1998. Detection of milk protein genetic polymorphisms in order to improve dairy traits in sheep and goats: a review. Small Ruminant Res. 27, 185–195.

Mroczkowski, S., Korman, K., Erhardt, G., Piwczynski, D., Borys, B., 2004. Sheep milk protein polymorphism and its effect on milk performance of Polish Merino. Arch. Tierz. 47, SI, 114–121.

Nilsen, H., Olsen, H.-G., Hayes, B., Sehested, E., Svendsen, M., Nome, T., Meuwissen, T., Lien, S., 2009. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. Genet. Sel. Evol. doi:10.1186/1297-9686-41-24.

Ordás, J.G., 2001. Technical Note: a short note on DNA lactotypes in the Churra and Manchega sheep breeds. Small Ruminant Res. 41, 71–75.

Picariello, G., Rignanese, D., Chessa, S., Ceriotti, G., Trani, A., Caroli, A., Di Luccia, A., 2009. Characterization and genetic study of the ovine α_{s2}-casein (*CSN1S2*) allele *B*. Protein J. 28, 333–340.

Pirisi, A., Piredda, G., Papoff, C.M., Di Salvo, R., Pintus, S., Garro, G., Ferranti, P., Chianese, L., 1999. Effects of sheep α_{s1}-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. J. Dairy Res. 66, 409–419.

Prinzenberg, E.M., Gutscher, K., Chessa, S., Caroli, A., Erhardt, G., 2005. Caprine κ-casein (*CSN3*) polymorphism: new developments in molecular knowledge. J. Dairy Sci. 88, 1490–1498.

Prinzenberg, E.M., Jianlin, H., Erhardt, G., 2008. Genetic variation in the κ-casein gene (*CSN3*) of Chinese yak (Bos grunniens) and phylogenetic analysis of *CSN3* sequences in the genus Bos. J. Dairy Sci. 91, 1198–1203.

Rampilli, M., Cecchi, F., Giuliotti, L., Cattaneo, T.M.P., 1997. The influence of β-lactoglobulin genetic polymorphism on protein distribution and coagulation properties in milk of massesse breed ewes. In: Proceedings of the 2nd IDF Seminar "Milk Protein Polymorphism", International Dairy Federation, Palmerston North, New-Zealand, pp. 311–315.

Recio, I., Fernández-Fournier, A., Martín-Álvarez, P.J., Ramos, M., 1997. β-lactoglobulin polymorphism in ovine breeds: influence on cheesemaking properties and milk composition. Lait 77, 259–265.

Sacchi, P., Chessa, S., Budelli, E., Bolla, P., Ceriotti, G., Soglia, D., Rasero, R., Cauvin, E., Caroli, A., 2005. Casein haplotype structure in five Italian goat breeds. J. Dairy Sci. 88, 1561–1568.

Sanchez, A., Ilahi, H., Manfredi, E., Serradilla, M., 2005. Potential benefit from using the α_{s1}-casein genotype information in a selection scheme for dairy goats. J. Anim. Breed. Genet. 122, 21–29.

Sargolzaei, M., Iwaisaki, H., Colleau, J.-J., 2006. CFC, Contribution, Inbreeding (F), Coancestry. Release 1.0 (April 2006)., http://www.agr.niigata-u.ac.jp/~iwsk/cfc.html.

Schmidt, D.V., Ebner, K.E., 1972. Multiple forms of pig, sheep and goat α-lactalbumin. Biochim. Biophys. Acta 263, 714–720.

Threadgill, D.W., Womack, J.E., 1990. Genomic analysis of the major bovine milk protein genes. Nucleic Acids Res. 18, 6935–6942.

Wessels, G., Hamann, H., Erhardt, G., Distl, O., 2004. Genotypeffekte von Milchprotein-Polymorphismen auf die Milchleistung beim Ostfriesischen Milchschaf. Berl. Münch. Tierärztl. Wochenschr. 117, 414–419.

Xie, X., Ott, J., 1993. Testing linkage disequilibrium between a disease gene and marker loci. Am. J. Hum. Genet. 53 (Suppl.), 1107.

Yeh, C.Y., Yang, R., Boyle, T., Ye, Z.-H., Moa, J., 1997. POPGENE–The User-friendly Freeware for Population Genetic Analysis., http://www.ualberta.ca/~fyeh/.

3. Genomics and proteomics of deleted ovine *CSN1S1*I*

I. J. Giambra*, L. Chianese[#], P. Ferranti[#], and G. Erhardt*

*Department of Animal Breeding and Genetics, Justus-Liebig-University, 35390 Gießen, Germany

[#]Dipartimento di Scienza degli Alimenti, University of Naples "Federico II", 80055 Portici, Naples, Italy

Published in:

International Dairy Journal 20 (3), 195-202 (2010)

Submitted: 25.05.2009

Accepted: 16.09.2009

ABSTRACT

Ovine α_{s1}-casein (*CSN1S1*) allele *I* (*CSN1S1*I*) was characterized at the molecular genetic and protein level. Sequencing of *CSN1S1* cDNA and mature protein showed the absence of exon 7 from *CSN1S1*I* in comparison with the *C''* genetic variant of the C phenotype. This allelic aberration is correlated with a sequence difference in 5´-splice donor sequence of intron 7 (g.656T > A), leading to upstream skipping of exon 7. Consequently mRNA sequence of ovine *CSN1S1*I* is 24 bp shorter than complete coding sequence leading to an abbreviation of eight amino acids in the mature protein, resulting in a lower degree of phosphorylation in comparison with *CSN1S1*C''*. However, *CSN1S1*I* was expressed at a quantitative level similar to that for the *C''* reference variant. Using amplified created restriction site polymerase chain reaction, a DNA-based test for identification of *CSN1S1*I* was developed. Altogether six nucleotide substitutions were identified within intron 6 and intron 7 of *CSN1S1* variants, forming three different haplotypes.

1. Introduction

Ovine caseins are distinguished into four fractions, namely α_{s1}-casein (α_{s1}-CN), α_{s2}-CN, β-CN and κ-CN, coded by the genes *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*, respectively. Within α_{s1}-CN (CSN1S1), seven phenotypes (A, B, C, D, E, F, and H) have been identified by electrophoretic techniques (Chianese et al., 1996; Pirisi et al., 1999) and one, CSN1S1 phenotype G, by high performance liquid chromatography (HPLC) analysis (Chianese et al., unpublished results). Primary structures have been described for CSN1S1 A, C, D, and E (Chianese, Caira, Garro, & Addeo, 2007; Ferranti et al., 1995), B and G (Chianese et al., unpublished results) and H (Giambra, Chianese, Ferranti, & Erhardt, unpublished results). On the basis of molecular analysis Ceriotti, Chiatti, Bolla, Martini, and Caroli (2005) distinguished within *CSN1S1*C* two variants, *C'* and *C''*, that were not identifiable at the protein level by isoelectric focusing (IEF). Ovine *CSN1S1*E* (Chianese et al., 2007) and *H* are characterized by the missing exons 10 and 8, respectively, as a consequence of alternative splicing demonstrated in *CSN1S1*H* (Giambra et al., unpublished results). This genetic mechanism has been described in other ruminants' milk such as bovine *CSN1S1*A* (exon 4) (Mohr, Koczan, Linder, Hobom, & Erhardt, 1994) and caprine *CSN1S1*F* (exons 9, 10 and 11) (Leroux, Mazure, & Martin, 1992) and *G* (exon 4) (Grosclaude & Martin, 1997; Martin & Leroux, 1994) and was presumed for bovine *CSN1S1*H* (exon 8) (Mahé et al., 1999). Part of the reason for exon deletion was identified as single nucleotide polymorphisms (SNPs)

concerning splice site sequences, leading to alternative splicing (Grosclaude & Martin, 1997; Martin & Leroux, 1994; Mohr et al., 1994).

In each of the cases described above, milk containing the deleted *CSN1S1* variants exhibited the worst aptitude for milk clotting, due either to the CSN1S1 and lower overall whole casein content or to the decrease in the number of phosphates anchoring calcium ions in curd. In this regard, Pirisi et al. (1999) showed that ovine *CSN1S1*D*, a non-deleted variant containing six phosphate groups (6 P), gave the worst aptitude for milk clotting compared with the A and C variants (10 P and 9 P, respectively). Therefore, milk proteins and their genetic polymorphisms are of increasing interest in ruminant breeding, because of their effects on production traits (Boettcher et al., 2004; Martin, Szymanowska, Zwierzchowski, & Leroux, 2002) and on cheese making properties (Hallén, Allmere, Naslund, Andren, & Lunden, 2007; Ikonen, Ahlfors, Kempe, Ojala, & Ruottinen, 1999; Pirisi et al., 1999), which in turn influence economic traits. In addition to their potential for population genetic markers (Ibeagha-Awemu et al., 2007; Mahé et al., 1999), milk protein variants and their allelic and/or phenotypic frequency can be used: (i) for the molecular tracing of typical cheeses (Chianese et al., 2009); (ii) to improve functional quality of milk and of derived products for the release of bioactive peptides (casomorphins, angiotensin I converting enzyme inhibitory peptides, phosphopeptides) through enzyme action in human digestion or during cheese ripening (Meisel, 1998; Park, Juarez, Ramos, & Haenlein, 2007); and (iii) identification of negative allergenic epitopes (Cocco, Jarvinen, Sampson, & Beyer, 2003; Spuergin, Mueller, Walter, Schiltz, & Forster, 1996).

The aims of this study were (i) to characterize the primary structure of ovine CSN1S1 I, recently identified by IEF in ovine milk of Gray Horned Heath ewes, on both genetic and biochemical basis; (ii) to develop a DNA-based test for its rapid identification, independently of the availability of milk samples; (iii) to trace the evolutionary pathway of ovine *CSN1S1*; and (iv) to hypothesize its role in micellar building and/or nutritional properties of milk.

2. Materials and methods

2.1. Materials

Milk and blood samples of five Gray Horned Heath sheep, identified as carriers of *CSN1S1*I* by IEF of milk samples (*II*: n = 1; *C''I*: n = 2; *DI*: n = 2) were used for molecular genetic analysis. As control eight animals from the carrier families and additionally 16 unrelated

animals were included in the DNA studies. For comparative analysis, seven milk and blood samples of reference sheep with known *CSN1S1*-genotypes (*AC''*: n = 1; *C''C''*: n = 2; *C''D*: n = 2; *DD*: n = 1; *HH*: n = 1) typed by IEF followed by DNA sequencing were used.
Protein sequencing was accomplished with milk samples of further animals with CSN1S1 AA, C''C'', C''D, C''E in comparison with CSN1S1 C''I.

2.2. Electrophoretic and immunoblotting analyses

Electrophoretic and immunoblotting analyses were carried out as reported by Chianese et al. (2009). Briefly, casein samples containing CSN1S1 C''I were dissolved in 9 M urea (0.2 g L^{-1}), containing 2-mercaptoethanol (1 mL L^{-1}). Urea-polyacrylamide gel electrophoresis (urea-PAGE) at pH 8.6 was carried out with a vertical electrophoretic apparatus (Protean II, Bio-Rad, Richmond, CA, USA) at 200 V and 6 °C for 7 h.
Ultra-thin layer isoelectric focusing (UTLIEF) on polyacrylamide gels (0.25 mm) was carried out in a pH gradient 2.5-6.5 by mixing ampholine buffers (GE Healthcare Amersham Bioscience, Buckinghamshire, UK) of pH 2.5-5.0, pH 4.5-5.4 and pH 4.0-6.5 (1.6:1.4:1, by vol). The gel was stained with Coomassie Brilliant Blue G-250.
For immunoblotting analysis, the casein fractions separated either by PAGE or UTLIEF analysis were transferred by capillary diffusion from the gel onto a nitrocellulose membrane (0.45 μm, Trans-Blot, Bio-Rad, Richmond, CA, USA) and immunostained using home-made polyclonal antibodies against bovine α_{s1}-CN, α_{s2}-CN, and κ-CN as primary antibodies.

2.3. HPLC purification of casein samples

Native or dephosphorylated whole casein samples (*C''C''* and *C''I*) were fractionated by HPLC on a Hewlett-Packard 1100 System (Agilent, Palo Alto, CA, USA) using a reverse phase Jupiter C4 column (250 × 2.0 mm, 5 μm, 300 Å; Phenomenex, St. Torrance, CA, USA) at 0.2 mL min^{-1} flow rate. Solvent A was 0.1% (v/v) trifluoroacetic acid (TFA) in water; solvent B was 0.1% (v/v) TFA in acetonitrile. Sample elution was achieved using a gradient from 5 to 70% of solvent B over 60 min. A volume of 100 μL was injected for each chromatographic run. The UV detector was set at 280 nm.

2.4. Enzymatic hydrolysis

The HPLC-purified casein samples containing the variant casein were freeze-dried before tryptic hydrolysis, which was carried out at 37 °C for 6 h using sequencing grade trypsin (7500 U mg^{-1} protein; Sigma-Aldrich, Munich, Germany) in 0.4% ammonium bicarbonate (Fluka, Milan, Italy) pH 8.5, with a 1:100 E:S ratio. Alkaline phosphatase hydrolysis of casein samples was carried out at 37 °C for 18 h in 0.4% (w/v) ammonium bicarbonate (Fluka, Milan, Italy) pH 8.5, using alkaline phosphatase (10 U mg^{-1} protein) from Roche (Switzerland) with a 1:100 E:S ratio. The enzymatic reactions were stopped by addition of 1 μL 10% formic acid (v/v) solution.

2.5. Mass spectrometry analysis of casein and casein peptides

Matrix-assisted-laser-desorption-time-of-flight-ionization mass spectra (MALDI-TOF-MS) were acquired using a Voyager DE-PRO (Perseptive Biosystem, Framingham, MS, USA) reflectron time-of-flight instrument operating in post source decay mode and equipped with a N_2 laser (337 nm, 3 ns pulse width). The accelerating voltage was 20 kV, the grid voltage was 85%, and the guide wire voltage was 0.02%.

Electrospray mass spectrometric (ESI-MS) analysis of casein samples was performed with a Micromass Q-TOF Ultima hybrid orthogonal mass spectrometer (Waters, Manchester, UK). For nano-electrospray mass spectrometry (nano-ESI/MS) analysis, HPLC-purified samples (25-50 pmol) were resuspended in 50 μL of 50:50:0.1 water/acetonitrile/TFA (by vol) and then were injected into the ion source at a flow rate of 1 μL min^{-1} operating in positive ion mode to run nano-ESI/MS and MS/MS experiments; the spectra were scanned from *m/z* 1600 to 600 at 10 s per scan.

2.6. Isolation of mRNA and DNA, cDNA synthesis, PCR amplifications, cloning and sequencing

About 10 mL of fresh milk of *CSN1S1*I* carriers and the reference sheep were centrifuged (10 min, 2200 x g) and the resulted pellet was washed twice with phosphate buffered saline/0.5 M ethylenediaminetetraacetic acid according to Boutinaud, Rulquin, Keisler, Djiane, and Jammes (2002). Isolation of total RNA was effected with the Invisorb Spin Plant RNA Mini Kit (Invitek GmbH, Berlin, Germany). This total RNA was immediately assembled for cDNA

synthesis in four separate reactions for each sample with the Verso cDNA Kit using an Oligo-dT-Primer (Thermo Fisher Scientific,Waltham, MA, USA).

The resulting cDNAs were used as PCR template for each animal performed in duplicate using two of the four cDNAs generated per mRNA sample. The forward primer (5´-cttcttcccagtcttgggttc-3´) used is located in the first exon of *CSN1S1* mRNA, whereas the reverse primer (5´-aattttcaatcccatcaaagacc-3´) lies partly in exon 18 and in exon 19. PCR amplification was carried out in the following way: after denaturation of 15 min at 95 °C the temperature cycling was as follows: 35 cycles 30 s at 95 °C, 30 s at 60 °C, 1 min at 72 °C and a final extension of 7 min at 72 °C.

Amplified PCR fragments were cloned with pGEM-T and pGEM-T Easy Vectors (Promega, Mannheim, Germany) according to Shin et al. (2005). Fifty recombinant clones per animal were sequenced bi-directionally using an ABI Prim 377 DNA analyzer, Big Dye Terminator sequencing kit v1.1, and vector primers M13-Forward and M13-Reverse (Applied Biosystems, Foster City, CA, USA).

Isolation of genomic DNA was carried out from milk somatic cells according to Lühken, Caroli, Ibeagha-Awemu, and Erhardt (2009) or from whole blood or buffy coat using the high-salt method of Montgomery and Sise (1990). DNA sequences from exon 6 to exon 8 including the introns were determined by four individual PCRs (Table 1; PCR 1-4) with overlapping fragments. Bi-directional direct sequencing of the four PCR fragments was done as described above, using the same primers used for PCR.

Table 1

Primer sequences, reaction conditions and resulted product sizes for characterization of exon 7 DNA sequence and flanking intron sequences of ovine *CSN1S1*.

PCR-number	PCR-name[a]	Primer forward (5´- 3´) (GenBank acc. no. and location)	Primer reverse (5´- 3´) (GenBank acc. no. and location)	Mg concentration (mM)	AT[b] (°C)	Cycles (n)	Product size (bp)
1	Ex7F-In8R	gggagtgaatcaattgaggtaag (NM_001009795; r.239-256 in combination with X59856; g.16685-16689)	caattgcagtggatagctttctt (X59856; g.17487-17465)	1.0	60.0	45	821
2	In7F-In8bR	ttctctattgccacccatttcta (FJ440847; g.603-625)	actgactgaatggacgtgagttt (X59856; g.17904-17882)	3.0	59.0	35	643
3	Ex6F-In7R	aaggagaacatcaatgaactgag (NM_001009795; r.209-228) in combination with X59856; g.16031-16034)	tttggccttagagaattgttcat (FJ440847; g.148-126)	3.0	64.0	35	776
4	Ex5F-In6R	tgtagccttttccagaagtgttt (NM_001009795; r.185-202) in combination with X59856; g.15365-15369)	ggttgcaaattatgattttgagg (FJ695514; g.86-64)	2.0	62.0	35	767

[a] Ex, Exon; In, Intron.

[b] AT, annealing temperature.

2.7. Restriction fragment length polymorphism analysis

To screen for an identified nucleotide substitution in intron 7 of ovine *CSN1S1*I* we analyzed DNA with a PCR based on amplification created restriction site (ACRS-PCR) (Haliassos et al., 1989; Lien, Alestrom, Klungland, & Rogne, 1992). A modified forward primer (5´-tattgggagtgaatcaattgaag-3´) and the reverse primer 5´-agccaattttggtcaacatacat-3´ were used. Amplification was carried out with the following PCR conditions: after a denaturation step of 90 s at 95 °C the temperature cycling was as follows: 35 cycles 15 s at 95 °C, 30 s at 50 °C, 30 s at 72 °C and a final extension of 5 min at 72 °C. ACRS-PCR created a DNA fragment of 211 bp with an artificial restriction site for *Mbo*II. PCR products were digested 2 h at 37 °C. Digested fragments were separated in a 3.5% agarose gel and detected after ethidium bromide staining.

2.8. Bioinformatic analyses

ChromasPro 1.32 (Technelysium Pty Ltd, Queensland, Australia) and DNASTAR (DNASTAR, Inc., Madison, WI, USA) were used for cDNA/mRNA and DNA sequence analysis.

The influence of a nucleotide sequence variation on attachment of splice factors to splicing signals in ovine *CSN1S1* gene areas was analyzed in silico with NNSPLICE 0.9 (Reese, Eeckman, Kulp, & Haussler, 1997).

3. Results

3.1. Electrophoretic and immunoblotting analyses

The Coomassie Blue immunostained PAGE profiles at alkaline pH of CSN1S1 AA, C''C'', C''D, C''E in comparison with CSN1S1 C''I sample and after immunoblotting with anti-α_{s1}-CN specific antibodies are shown in Fig. 1a and b, respectively.

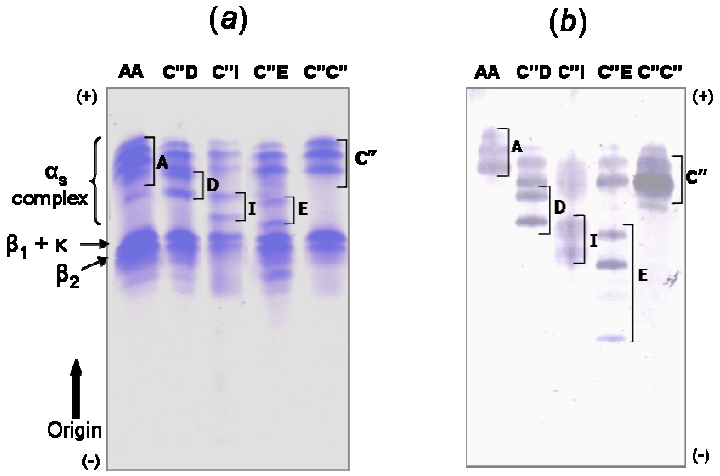


Fig. 1. Disc-PAGE at pH 8.6 of ovine CSN1S1 variants. (a) Coomassie brilliant blue staining, (b) Immunoblotting with polyclonal antibodies against CSN1S1.

The results indicate that the CSN1S1 I variant had an electrophoretic mobility towards the anode similar to or slightly higher than that of E variant. The UTLIEF analysis of the same samples is reported in Fig. 2 (immunostaining with specific polyclonal anti-α_{s1}-CN antibodies). This procedure was necessary because of the partial overlapping between CSN1S1 I and CSN2 caused by their very similar pI values. Comparison of the electrophoretic profiles of the CSN1S1 variants by both electrophoretic techniques showed that their relative mobility was in the order: A > C'' > D > I > E.

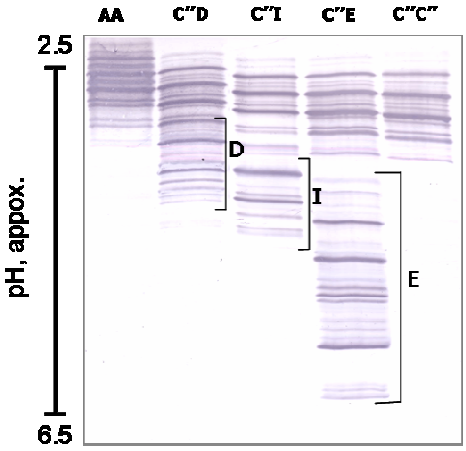


Fig. 2. Immunoblotting with polyclonal antibodies against α_{s1}-casein of ovine CSN1S1 variants.

3.2. Determination of primary structure by mass spectrometry analysis

The reversed phase HPLC separation of casein fractions and subsequent MALDI-TOF-MS analysis of the eluted CSN1S1 peaks gave rise two clusters of components (one corresponding to the C'' variant and the other corresponding to the I variant), whose average molecular mass values differed by 831 Da. Furthermore, on the basis of the mass shift between native and dephosphorylated proteins (480 and 560 Da respectively; Table 2) it was found that the I variant had a lower degree of phosphorylation (6/7 phosphate groups per molecule instead of 8/9 phosphate groups per molecule occurring in variant C'').

Table 2

Average molecular mass of native and dephosphorylated CSN1S1 variants C'' and I and phosphate groups per molecule.

CSN1S1 variant	Average molecular mass (Da)			Phosphate groups per molecule
	Native	Dephosphorylated	Native - dephosphorylated	
C''	23,401/23,481	22,761	640/720	8/9
I	22,410/22,490	21,930	480/560	6/7

The identity of the variants was ascertained by mass spectrometry analysis of the tryptic digest of the dephosphorylated casein, showing the presence of a variant peptide at 906 *m/z*. The other mass signals were in agreement with those given by the sequence of the common CSN1S1 C'' phenotype (data not shown). Tandem mass spectrometry analysis of the variant peptide carried out by nano-ESI/MS/MS analysis (Fig. 3) provided the sequence Asp-Gln-Ala-Met-Glu-Asp-Ala-Lys, unmatched by any of the tryptic peptides expected on the basis of full-length CSN1S1.

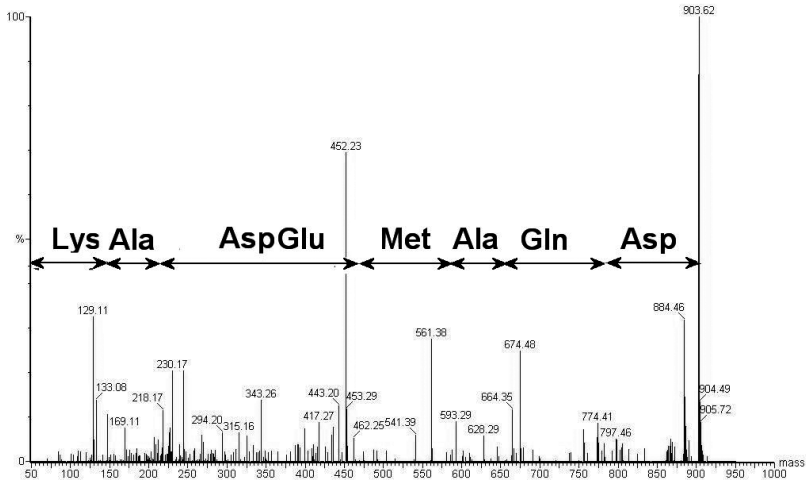


Fig. 3. Deconvoluted *nano*-ESI-Q-TOF-MS/MS spectrum of variant tryptic peptide at 906.5 Da, for which the amino acid sequence Asp-Ile-Gly-Ser-Glu-Ser-Ile-Glu was determined.

Sequence analysis showed instead that it corresponded to the sequence 51-58. This peptide is normally included in the tryptic peptide 43-58 in full-length CSN1S1 variants, and this finding indicated that the complementary sequence 43-50 (Asp-Ile-Gly-Ser-Glu-Ser-Ile-Glu) was missing in the deleted variant. The CSN1S1 I is therefore a deleted variant of sheep CSN1S1 C'' lacking the sequence 43-50. It is important to underline that the missing sequence contained two phosphorylated Ser residues (namely p.SerP46 and p.SerP48).
In the dephosphorylated spectrum, the main component was also accompanied by several others with lower mass (Fig. 4).

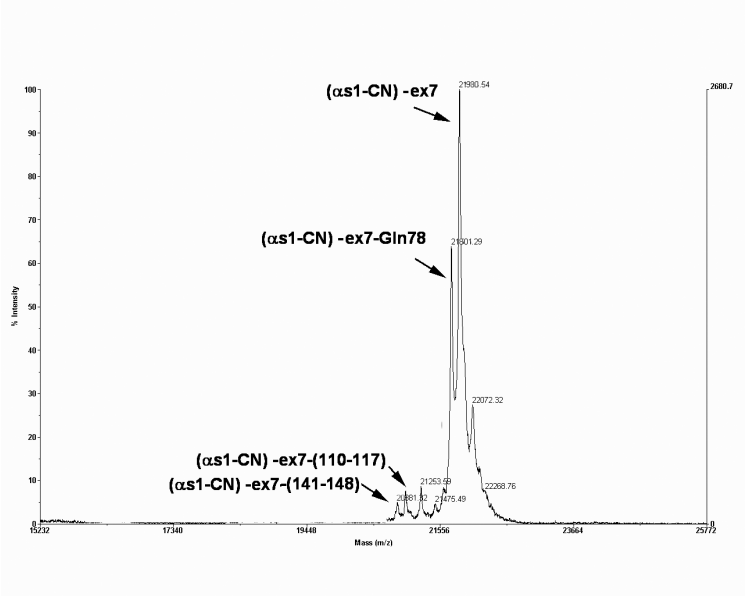


Fig. 4. MALDI-TOF-MS spectra of dephosphorylated CSN1S1 I. A major component was observed at 21,930 Da. Arrows indicate non-allelic deleted species previously shown to occur, to a variable extent, in normal sheep CSN1S1 phenotypes (Chianese et al., 1996; Ferranti et al., 1995).

It needs to be underlined that the integration of multicharged ion peaks of dephosphorylated CSN1S1 C''I sample allowed us to determine that the quantitative level of the variant I was the same as that for the C'' reference variant (data not shown).

3.3. Ovine CSN1S1 mRNA sequences

Sequence analysis of cloned PCR fragments of ovine cDNA spanning the coding region of exon 1 to exon 18 showed that exon 7 was missing in *CSN1S1*I* cDNA/mRNA (GenBank Acc. No. FJ695513) in comparison with *CSN1S1*C''* (GenBank Acc. No. FJ440845), leading to the shortening of the mature protein by eight amino acids (Fig. 5) as described above.

Except for the loss of exon 7, we could not identify any further sequence difference within the amplified cDNA/mRNA sequences of *CSN1S1*A*, *C''*, *D*, and *I*.

```
                          35                                  45                                  54
amino acid sequence        E  N  I  N  E  L  S  K  D  I  G  S  E  S  I  E  D  Q  A  M
CSN1S1*C''(FJ440845)      GAGAACATCAATGAACTGAGTAAGGATATTGGGAGTGAATCAATTGAGGATCAAGCCATG
CSN1S1*I  (FJ695513)      **********************----------------------**********
                          184                                                       243
```

Fig. 5. Detail of the amino acid sequence constituting peptides 35-54 of mature protein (single letter code) and the corresponding cDNA/mRNA (nucleotides 184 to 243) sequence of *CSN1S1*C''* in alignment with the corresponding coding sequence of *CSN1S1*I*. The remaining sequences were identical. Within the sequence of *CSN1S1*I* (GenBank Acc. No. FJ695513), stars denote analogue cDNA/mRNA sequences and dashes indicate the missing exon 7 accentuated by bold letters in *CSN1S1*C''* (GenBank Acc. No. FJ440845) sequence.

3.4. Ovine CSN1S1 DNA sequences

Analysis of the genomic ovine *CSN1S1* sequence from exon 6 to exon 8 showed that the sequence corresponding to exon 7 is present in DNA sequence of *CSN1S1*I* carriers. However, sequencing of PCR fragments of *CSN1S1 C''I*, *II*, *AC''*, *C''C''* and *DD* revealed two sequence differences (g.656T > A and g.668T > C) in intron 7 in *CSN1S1*I* (GenBank Acc No. FJ695515) whereas g.656T > A affects 5´-splice donor sequence of intron 7 and is correlated with g.668T > C (Fig. 6) without direct influence on splicing process (NNSPLICE 0.9; Reese et al, 1997).

```
                          630                                                                 690
CSN1S1*C''(FJ695514)      GGATATTGGGAGTGAATCAATTGAGGTAAGATTCTTTATTTTAAAACTATTAAATATAAT
CSN1S1*I  (FJ694415)      GGATATTGGGAGTGAATCAATTGAGGAAAGATTCTTTACTTTAAAACTATTAAATATAAT
                          ************************* *********** *********************
```

Fig. 6. Detail (g.630-690) of DNA sequence of *CSN1S1*C''* and *CSN1S1*I*. Circles designate sequence differences between *CSN1S1*C''* (GenBank Acc. No. FJ695514) and *CSN1S1*I* (GenBank Acc. No. FJ695515); the first circle marks g.656T > A, responsible for exon skipping of exon 7 in pre-mRNA processing, while the second circle designates correlated sequence alteration g.668T > C. The underlined sequence in *CSN1S1*C''* corresponds to the 5´-donor splice site.

The g.656T > A substitution leads to upstream exon skipping during the serial splice reactions of the *I* allele *CSN1S1* pre-mRNA. That results in an interruption of the normal complementarity of the intron 7 splice donor signal with U1 small nuclear RNA (U1-snRNA). In the passage of exon 7 to intron 7 in *CSN1S1*C''*, the binding score for splice signals was 1.00, whereas in *CSN1S1*I* no splice site was found in silico by NNSPLICE 0.9; hence exon 7 is spliced out in pre-mRNA processing. Furthermore, the sequence differences g.656T > A

and g.668T > C were identified in heterozygous form in nine out of a total of 24 control animals, not typed by IEF previously. Therefore, it can be concluded that these animals are also carriers of *CSN1S1*I*.

Sequence analysis of DNA fragment of exon 6 to exon 8 of animals with *CSN1S1 AC''*, *C''C''*, *C''D*, *DD*, *HH*, *C''I*, *DI* and *II* as well as sequencing of the additional 24 samples without IEF-pre-typing revealed further sequence differences within intron 6 and intron 7, without influence on splicing process after in silico analysis with NNSPLICE 0.9. Different genotypes were identified for the positions 105, 295, 402, 656, 668 and 1195 (Table 3) resulting in three different haplotypes within *CSN1S1*C''* (GenBank Acc. No. FJ695514) and *CSN1S1*I* (GenBank Acc. No. FJ695515). Haplotypes H01 and H02 occurred in animals with *CSN1S1 AC''*, *C''C''*, *C''D*, *DD* and *HH*, whereas haplotype H03 could be demonstrated in an animal with *CSN1S1 II.*

Table 3

Observed *CSN1S1*-genotypes (G01-G11) with their distribution and inferred haplotypes (H01-H03).

Genotype or haplotype	n[a]	SNP position[b]					
		105	295	402	656	668	1195
G01	6	A	A	C	T	T	G
G02	3	A	A	T	T	T	G
G03	5	A	A	Y[c]	T	T	G
G04	3	M[d]	W[e]	T	T	T	R[f]
G05	5	M	W	Y	T	T	R
G06	1	M	W	T	W	Y	G
G07	3	M	W	T	W	Y	R
G08	4	M	W	Y	W	Y	R
G09	4	C	T	T	W	Y	A
G10	1	C	T	T	A	C	A
G11	1	C	W	Y	W	Y	R
H01		A	A	C	T	T	G
H02		A	A	T	T	T	G
H03		C	T	T	A	C	A

[a] n = number of observed animals.
[b] The single nucleotide polymorphism (SNP) positions are relative to the GenBank Acc. No. FJ695514 and FJ695515.
[c] Y = CT.
[d] M = AC.
[e] W = AT.
[f] R = AG.

3.5. Restriction fragment length polymorphism analysis

The PCR products (211 bp) of animals with *CSN1S1*I* were hydrolysed after *Mbo*II-digestion into two fragments of 32 and 179 bp while PCR products of animals with remaining *CSN1S1* variants were not digested. The results were demonstrated by agarose gel electrophoresis (Fig.

7) and were in full agreement with sequencing results of 24 control animals and animals with known *CSN1S1* genotype.

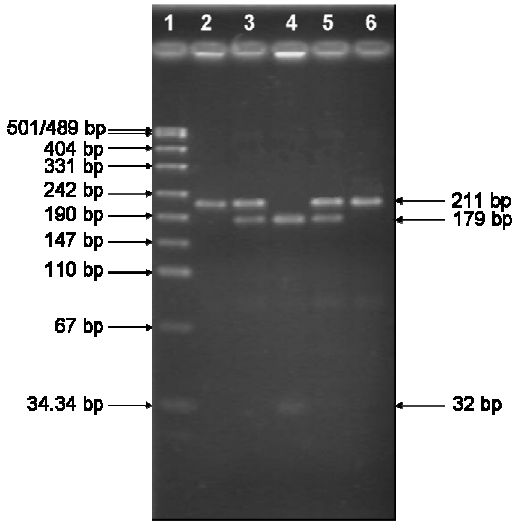


Fig. 7. Agarose gel electrophoresis of restriction fragment length polymorphism (RFLP) analysis of different ovine *CSN1S1*-genotypes. Lane 1: marker; Lanes 2 + 6: *CSN1S1 C''C''*; Lanes 3 + 5: *CSN1S1 C''I*; Lane 4: *CSN1S1 II*.

4. Discussion

In this study an integrated approach based on genomic and proteomic methodologies was applied to define the genetic and molecular aspects leading to the occurrence of a novel ovine *CSN1S1* genetic variant and determining its structural and functional characteristics. The proteomic approach, by means of immuno-electrophoretic techniques coupled to mass spectrometry analysis, showed the occurrence of a new deleted ovine CSN1S1, caused by skipping of exon 7. This deletion produces a shortened protein chain that also determines a decrease of the protein net negative charge and a highest pI value as confirmed by UTLIEF in comparison with the reference ovine A, C and D variants. Taking into account that electrophoretic behaviour for the known variants A, C, D and E were explained by a decreasing number of phosphate groups (10, 9, 6 and 5 P respectively) it is important to underline that the missing sequence in variant I contained two phosphorylated serine residues (namely p.SerP46 and p.SerP48). This explains the observed lower phosphorylation degree of CSN1S1 I in comparison with CSN1S1 C'' (Table 2) and in turn explained the highest net charge value and the lowest pI than variant E (Chianese et al., 2007).

Beside the two phosphorylated serine residues, the deleted amino acid residues include three acidic residues (1 Asp and 2 Glu) that can be involved in electrostatic bonds with calcium

ions, contributing to micellar size and curd consistency (Pirisi et al., 1999; Tziboula & Horne, 1999). As a consequence, it can be expected that parameters such as micellar size and curd firming rate of CSN1S1 I would be higher and curd firmness lower than those given by CSN1S1 A and C''. Furthermore, the new *CSN1S1*I* was expressed at same level of the full-length *C''* variant, differently from the known goat *CSN1S1* deleted variants expressed at a lower level than strong goat *CSN1S1* variants, and from the lower expression of deleted ovine *CSN1S1*E* (Chianese et al., 2007) and *H* (Giambra et al., unpublished results).

Molecular characterization revealed that *CSN1S1*I* is a further variant caused by posttranscriptional alteration of pre-mRNA, due to exon skipping, showing a widespread distribution, even in bovine *CSN1S1*A* (Mohr et al., 1994), caprine *CSN1S1*F* and *G* (Grosclaude & Martin, 1997; Leroux et al., 1992), and as presumed for bovine *CSN1S1*H* (Mahé et al., 1999). Additionally, ovine *CSN1S1*E* is characterized by loss of exon 10 in mature protein, but no further molecular genetic analyses are available (Chianese et al., 2007). In bovine *CSN1S1*A* (Mohr et al., 1994) and caprine *CSN1S1*G* (Grosclaude & Martin, 1997; Martin & Leroux, 1994), SNPs in the splice donor sequence of the following intron are responsible for alternative splicing. Also in ovine *CSN1S1*I* an SNP within the 5´-donor splice site of intron 7 is the cause of alternative splicing; base pairing of U1-snRNA is avoided. Therefore, it can be assumed that these sequence differences in 5´-splice sites of introns are the main reason for alternative splicing in ruminant casein genes, resulting in a large variability.

Evolution and composition of casein genes can probably explain this phenomenon. The characteristically extremely split architecture of the *CSN1S1* and *CSN1S2*, with small exons in the coding region evolved by both intragenic and intergenic duplication of a primordial gene (Groenen, Dijkhof, Verstege, & Van Der Poel, 1993), can support exon skipping (Rijnkels, 2002).

The additionally identified sequence differences within intron 6 and 7 of *CSN1S1*C''* and *I* and the resulting haplotypes show potential for identification of further variability within ovine caseins as it could be already demonstrated for bovine *CSN3* (Chen et al., 2008).

The results of this study show that it is possible to genotype the milk protein variants in lactating ewes by IEF, allowing simultaneous screening of the main milk protein variants at the phenotype level in a fast, cheap and high resolution manner (Seibert, Erhardt, & Senft, 1985). In addition, non-lactating females, as well as rams, could be typed by ACRS-PCR and restriction fragment length polymorphism (RFLP) analysis. It should be noted that this test, and also other molecular genetic tests already available in sheep (Ceriotti et al., 2004; Pilla,

Bevilacqua, Leroux, Fraghi, & Martin, 1998; Ramunno et al., 1997), are able to identify only a certain *CSN1S1* allele. Therefore, these molecular genetic tests are very expensive for screening large populations and are at the moment only a complementation to biochemical analysis. On the other hand, further characterization of the genetic variation within caseins and whey protein genes in sheep will offer the chance to apply microarray technology for screening SNPs, as already used in bovine milk protein genes (Chessa et al., 2007).

In the meantime, it is necessary to analyze the effects of variant *I* on milk yield and composition, especially protein content, as correlations are published between abbreviated CSN1S1 forms, reduced protein content, and cheese making ability in ruminants (Grosclaude & Martin, 1997; Leroux et al., 1992; Ng-Kwai-Hang, 2006). Additionally, milk composition has an influence on development and growth of the offspring (Moroni, Pisoni, Varisco, & Boettcher, 2007; Snowder & Glimp, 1991). Because of the tight association among casein genes, consideration of the entire casein haplotype instead of single gene typing is important for the improvement of milk production traits, as is done in cattle and goat (Heck et al., 2009; Nilsen et al., 2009; Sacchi et al., 2005). Similarly, studies concerning the variability within the other casein genes are necessary in sheep.

Further analyses are important to clarify if *CSN1S1*I* is specific for German Gray Horned Heath with potential for breed characterisation as well as diversity and phylogenetic studies, as in other ruminants (Caroli et al., 2006; Erhardt, 1993; Farrell et al., 2004; Jann et al., 2004).

In this context, appearance of biological active peptides or allergenic epitopes by enzymatic hydrolysis should be implicated in further studies to estimate the value of new milk protein variants for human nutrition and production of functional food. We have to analyze if the abbreviation within ovine *CSN1S1*I* and the correlated rearrangement of amino acid sequence leads to occurrence of (new) peptides with bioactive value, as could be shown within bovine kappa casein variants (Weimann, Meisel, & Erhardt, 2009).

In this respect, it can be envisaged that integration of the complementary genomic and proteomic analytical approaches will give a strong contribution to assess the structure-function relationship of milk protein variants and their role in determining the functional properties of the relative milks.

5. Conclusions

The biochemical characterization of ovine *CSN1S1*I* confirmed electrophoretic mobility in isoelectric focusing and identified the amino acid sequence differences in comparison to known ovine *CSN1S1* variants. Genomic analysis revealed cDNA/mRNA sequence and its abbreviation caused by alternative splicing and underlying sequence difference affecting 5´-splice site of intron 7. Given the fact that *CSN1S1*I* was expressed at a quantitative level similar to the *C''* reference variant, further studies have to include effects on economical and nutritional traits.

Acknowledgement

We thank Gesine Lühken for valuable comments.

References

Boettcher, P. J., Caroli, A., Stella, A., Chessa, S., Budelli, E., Canavesi, F., et al. (2004). Effects of casein haplotypes on milk production traits in Italian Holstein and Brown Swiss cattle. *Journal of Dairy Science, 87*, 4311-4317.

Boutinaud, M., Rulquin, H., Keisler, D. H., Djiane, J., & Jammes, H. (2002). Use of somatic cells from goat milk for dynamic studies of gene expression in the mammary gland. *Journal of Animal Science, 80*, 1258-1269.

Caroli, A., Chiatti, F., Chessa, S., Rignanese, D., Bolla, P., & Pagnacco, G. (2006). Focusing on the goat casein complex. *Journal of Dairy Science, 89*, 3178-3187.

Ceriotti, G., Chessa, S., Bolla, P., Budelli, E., Bianchi, L., Duranti, E., et al. (2004). Single nucleotide polymorphisms in the ovine casein genes detected by polymerase chain reaction-single strand conformation polymorphism. *Journal of Dairy Science, 87,* 2606-2613.

Ceriotti, G., Chiatti, F., Bolla, R., Martini, M., & Caroli, A. (2005). Genetic variability of the ovine α_{s1}-casein. *Italian Journal of Animal Science, 4*, 64-66.

Chen, S. Y., Costa, V., Azevedo, M., Baig, M., Malmakov, N., Luikart, G., et al. (2008). Short communication: new alleles of the bovine κ-casein gene revealed by resequencing and haplotype inference analysis. *Journal of Dairy Science, 91*, 3682-3686.

Chessa, S., Chiatti, F., Ceriotti, G., Caroli, A., Consolandi, C., Pagnacco, G., et al. (2007). Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms. *Journal of Dairy Science, 90,* 451-464.

Chianese, L., Caira, S., Garro, G., & Addeo, F. (2007). Primary structure of ovine deleted variant α_{s1}-CN E. In: Proceedings of the 5th international symposium on the challenge to sheep and goats milk sectors (p. 70). Alghero/Sardinia. Italy.

Chianese, L., Garro, G., Mauriello, R., Laezza, P., Ferranti, P., & Addeo, F. (1996). Occurrence of five α_{s1}-casein variants in ovine milk. *Journal of Dairy Research, 63*, 49-59.

Chianese, L., Quarto, M., Pizzolongo, F., Calabrese, M. G., Caira, S., Mauriello, R., et al. (2009). Occurrence of genetic polymorphism at the α_{s1}-casein locus in Mediterranean water buffalo milk. *International Dairy Journal, 19*, 181-189.

Cocco, R. R., Jarvinen, K. M., Sampson, H. A., & Beyer, K. (2003). Mutational analysis of major, sequential IgE-binding epitopes in α_{s1}-casein, a major cow's milk allergen. *Journal of Allergy and Clinical Immunology, 112*, 433-437.

Erhardt, G. (1993). Allele frequencies of milk proteins in German cattle breeds and demonstration of αs2-casein variants by isoelectric focusing. *Archiv für Tierzucht.* [*Archives of Animal Breeding*]*, 36*, 145-152.

Farrell, H. M., Jr., Jimenez-Flores, R., Bleck, G. T., Brown, E. M., Butler, J. E., Creamer, L. K., et al. (2004). Nomenclature of the proteins of cows' milk-sixth revision. *Journal of Dairy Science, 87*, 1641-1674.

Ferranti, P., Malorni, A., Nitti, G., Laezza, P., Pizzano, R., Chianese, L., et al. (1995). Primary structure of ovine α_{s1}-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. *Journal of Dairy Research, 62*, 281-296.

Groenen, M. A., Dijkhof, R. J., Verstege, A. J., & Van Der Poel, J. J. (1993). The complete sequence of the gene encoding bovine α_{s2}-casein. *Gene, 123,* 187-193.

Grosclaude, F., & Martin, P. (1997). Casein polymorphisms in the goat. In: Proceedings of the 2nd IDF seminar "Milk Protein Polymorphism" (pp. 241-253). Palmerston North, New Zealand: International Dairy Federation.

Haliassos, A., Chomel, J. C., Grandjouan, S., Kruh, J., Kaplan, J. C., & Kitzis, A. (1989). Detection of minority point mutations by modified PCR technique: a new approach for

a sensitive diagnosis of tumor-progression markers. *Nucleic Acids Research, 17,* 8093-8099.

Hallén, E., Allmere, T., Naslund, J., Andren, A., & Lunden, A. (2007). Effect of genetic polymorphism of milk proteins on rheology of chymosin-induced milk gels. *International Dairy Journal, 17,* 791-799.

Heck, J. M., Schennink, A., Van Valenberg, H. J., Bovenhuis, H., Visker, M. H., Van Arendonk, J. A., et al. (2009). Effects of milk protein variants on the protein composition of bovine milk. *Journal of Dairy Science, 92,* 1192-1202.

Ibeagha-Awemu, E. M., Prinzenberg, E. M., Jann, O. C., Lühken, G., Ibeagha, A. E., Zhao, X., et al. (2007). Molecular characterization of bovine *CSN1S2*B* and extensive distribution of zebu-specific milk protein alleles in European cattle. *Journal of Dairy Science, 90,* 3522-3529.

Ikonen, T., Ahlfors, K., Kempe, R., Ojala, M., & Ruottinen, O. (1999). Genetic parameters for the milk coagulation properties and prevalence of noncoagulating milk in Finnish dairy cows. *Journal of Dairy Science, 82*, 205-214.

Jann, O. C., Ibeagha-Awemu, E. M., Ozbeyaz, C., Zaragoza, P., Williams, J. L., Ajmone-Marsan, P., et al. (2004). Geographic distribution of haplotype diversity at the bovine casein locus. *Genetics, Selection, Evolution, 36,* 243-257.

Leroux, C., Mazure, N., & Martin, P. (1992). Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat α_{s1}-casein transcripts. Structural organization of the relevant gene. *Journal of Biological Chemistry, 267*, 6147-6157.

Lien, S., Alestrom, P., Klungland, H., & Rogne, S. (1992). Detection of multiple beta-casein (CASB) alleles by amplification created restriction sites (ACRS). *Animal Genetics, 23*, 333-338.

Lühken, G., Caroli, A., Ibeagha-Awemu, E. M., & Erhardt, G. (2009). Characterization and genetic analysis of bovine α_{s1}-casein *I* variant. *Animal Genetics, 40*, 479-485

Mahé, M. F., Miranda, G., Queval, R., Bado, A., Zafindrajaona, P. S., & Grosclaude, F. (1999). Genetic polymorphism of milk proteins in African Bos taurus and Bos indicus populations. Characterization of variants α_{s1}-Cn H and κ-Cn J. *Genetics, Selection, Evolution, 31,* 239-253.

Martin, P., & Leroux, C. (1994). Characterization of a further α_{s1}-casein variant generated by exon skipping. In: Proceedings of the 24th international society of animal genetics conference (p. 86). Prague.

Martin, P., Szymanowska, M., Zwierzchowski, L., & Leroux, C. (2002). The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reproduction Nutrition Development, 42,* 433-459.

Meisel, H. (1998). Overview on milk protein-derived peptides. *International Dairy Journal, 8*, 363-373.

Mohr, U., Koczan, D., Linder, D., Hobom, G., & Erhardt, G. (1994). A single point mutation results in A allele-specific exon skipping in the bovine α_{s1}-casein mRNA. *Gene, 143*, 187-192.

Montgomery, G. W., & Sise, J. A. (1990). Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research, 33*, 437-441.

Moroni, P., Pisoni, G., Varisco, G., & Boettcher, P. (2007). Effect of intramammary infection in Bergamasca meat sheep on milk parameters and lamb growth. *Journal of Dairy Research, 74*, 340-344.

Ng-Kwai-Hang, K. F. (2006). Genetic variants of milk proteins and their effects on the yield and quality of cheese. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, *56*, 1-11.

Nilsen, H., Olsen, H.-G., Hayes, B., Sehested, E., Svendsen, M., Nome, T., et al. (2009). Casein haplotypes and their association with milk production traits in Norwegian Red cattle. *Genetics, Selection, Evolution, 41*, 24.

Park, Y. W., Juarez, M., Ramos, M., & Haenlein, G. F. W. (2007). Physico-chemical characteristics of goat and sheep milk. *Small Ruminant Research, 68*, 88-113.

Pilla, F., Bevilacqua, C., Leroux, C., Fraghi, A., & Martin, P. (1998). Genotyping of α-s1 casein in sheep. *Animal Genetics, 29*, 472-473.

Pirisi, A., Piredda, G., Papoff, C. M., Di Salvo, R., Pintus, S., Garro, G., et al. (1999). Effects of sheep α_{s1}-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. *Journal of Dairy Research, 66,* 409-419.

Ramunno, L., Cosenza, G., Rando, A., Macciotta, N. P. P., Pappalardo, M., & Masina, P. (1997). Identification of carriers of the Welsh CASA1 variant using an allele-specific PCR method. *Animal Genetics, 28,* 154-155.

Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved splice site detection in Genie. *Journal of Computational Biology, 4*, 311-323.

Rijnkels, M. (2002). Multispecies comparison of the casein gene loci and evolution of casein gene family. *Journal of Mammary Gland Biology and Neoplasia, 7,* 327-345.

Sacchi, P., Chessa, S., Budelli, E., Bolla, P., Ceriotti, G., Soglia, D., et al. (2005). Casein haplotype structure in five Italian goat breeds. *Journal of Dairy Science, 88*, 1561-1568.

Seibert, B., Erhardt, G., & Senft, B. (1985). Procedure for simultaneous phenotyping of genetic variants in cow's milk by isoelectric focusing. *Animal Blood Groups and Biochemical Genetics, 16*, 183-191.

Shin, H. J., Park, H. Y., Jeong, S. J., Park, H. W., Kim, Y. K., Cho, S. H., et al. (2005). STAT4 expression in human T cells is regulated by DNA methylation but not by promoter polymorphism. *Journal of Immunology, 175*, 7143-7150.

Snowder, G. D., & Glimp, H. A. (1991). Influence of breed, number of suckling lambs, and stage of lactation on ewe milk-production and lamb growth under range conditions. *Journal of Animal Science, 69,* 923-930.

Spuergin, P., Mueller, H., Walter, M., Schiltz, E., & Forster, J. (1996). Allergenic epitopes of bovine α_{s1}-casein recognized by human IgE and IgG. *Allergy, 51*, 306-312.

Tziboula, A., & Horne, D. S. (1999). The role of α_{s1}-casein in the structure of caprine casein micelles. *International Dairy Journal, 9,* 173-178.

Weimann, C., Meisel, H., & Erhardt, G. (2009). Bovine κ-casein variants result in different angiotensin I converting enzyme (ACE) inhibitory peptides. *Journal of Dairy Science, 92,* 1885-1888.

4. *Short Communication*: Molecular genetic characterization of ovine α_{s1}-casein allele H caused by alternative splicing

I. J. Giambra*, L. Chianese[#], P. Ferranti[#], and G. Erhardt*

*Department of Animal Breeding and Genetics, Justus-Liebig-University, 35390 Gießen, Germany

[#]Dipartimento di Scienza degli Alimenti, University of Naples "Federico II", 80055 Portici, Naples, Italy

Published in:

Journal of Dairy Science 93 (2), 792-795 (2010)

Submitted: 05.08.2009

Accepted: 30.10.2009

ABSTRACT

Sequencing of ovine *CSN1S1*H* cDNA showed an absence of exon 8 in comparison with GenBank sequences; the absence was confirmed by protein sequencing. We demonstrated that this allelic aberration is the result of a deletion of 4 nucleotides, the last 3 of exon 8 and the first 1 of intron 8, which are replaced by an insertion of 13 nucleotides in the DNA sequence. The insertion is a precise duplication of a part of the adjacent intronic sequence of *CSN1S1*C''*. These sequence differences result in an inactivation of the splice donor sequence distal to exon 8, leading to upstream exon skipping during the serial splice reactions of the ovine *CSN1S1*H* pre-mRNA, and may affect the specific casein expression as well as protein characteristics.

Key words: ovine, *CSN1S1*, messenger ribonucleic acid, exon skipping

Among the ovine CN, the 4 fractions α_{s1}-, α_{s2}-, β-, and κ-CN, encoded by the genes *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*, respectively, are distinguished. Milk proteins and their genetic polymorphisms are of interest in ruminant breeding; several studies have revealed their effects on production traits (Martin et al., 2002; Boettcher et al., 2004) and on cheese-making properties (Ikonen et al., 1999; Pirisi et al., 1999; Hallén et al., 2007). Furthermore, milk protein variants and their allelic or phenotypic frequencies can be used for the molecular tracing of typical cheeses (Chianese et al., 2009) and as population genetic markers (Mahé et al., 1999; Ibeagha-Awemu et al., 2007).

So far, 7 phenotypes (A, B, C, D, E, F, and X) of α_{s1}-CN in ovine milk have been identified by protein electrophoresis (Chianese et al., 1996; Pirisi et al., 1999; Wessels et al., 2004), whereas primary structures have been determined only for α_{s1}-CN A, C, D, and E (Ferranti et al., 1995; Chianese et al., 2007). Within *CSN1S1*C*, Ceriotti et al. (2005) distinguished between 2 haplotypes, *C'* and *C''*, at the DNA level. Recently the α_{s1}-CN phenotype X was recovered by isoelectric focusing (**IEF**) in East Friesian dairy sheep and named *CSN1S1*H* (Giambra et al., 2009).

Therefore, the aims of this study were to characterize *CSN1S1*H* at the protein, mRNA, and DNA levels to better understand *CSN1S1* variants in sheep, and to develop a DNA-based test for their identification.

Milk samples of 4 East Friesian dairy sheep with *CSN1S1* genotypes *HH* (n = 1) and *CH* (n = 3), previously typed by IEF, were used for molecular genetic analysis. For comparative analysis, 6 milk and blood samples of reference sheep with known *CSN1S1* genotypes

(*AC''*: n = 1; *C''C''*: n = 2; *C''D*: n = 2; *DD*: n = 1), typed by IEF and by DNA sequencing, were used.

From each of the 10 samples, about 10 mL of fresh milk were collected and centrifuged (10 min, 2,200 x *g*) and the resulting pellet was washed according to Boutinaud et al. (2002). Total RNA was extracted with the Invisorb Spin Plant RNA mini kit (Invitek, Berlin, Germany). The RNA was assembled immediately for cDNA synthesis in 4 separate reactions for each sample with the Verso cDNA kit using an Oligo-dT primer (both from Thermo Fisher Scientific, Waltham, MA).

The resulting cDNA was used as PCR template for each animal using 2 of the 4 cDNA generated per mRNA sample in 2 separate PCR. The forward primer (5'-cttcttcccagtcttgggttc-3') is located in the first exon of *CSN1S1* mRNA, whereas the reverse primer (5'-aattttcaatcccatcaaagacc-3') spans exons 18 and 19. Amplified PCR fragments were cloned with pGEM-T and pGEM-T Easy Vectors (Promega, Mannheim, Germany) according to Shin et al. (2005). Fifty recombinant clones per animal were sequenced bidirectionally using an ABI Prim 377 DNA analyzer, Big Dye Terminator sequencing kit v1.1, and vector primers M13-Forward and M13-Reverse (all from Applied Biosystems, Foster City, CA).

Genomic DNA was isolated from blood according to Montgomery and Sise (1990) and from milk somatic cells according to Lühken et al. (2009) and used for 4 PCR (Table 1; PCR 1-4) concerning DNA sequence from exon 7 to exon 9 in overlapping fragments. Bidirectional direct sequencing of the 4 PCR fragments (1 to 4) was done as described above, using the same primers used for PCR.

To establish a DNA-based test, the forward primer used in PCR 5 (Table 1) was labeled with indodicarbocyanine (Cy5) to enable fragment length analysis using an A.L.F.express sequencer and the AlleleLocator 1.03 software (both from Amersham Biosciences, Freiburg, Germany) for all carriers of *CSN1S1*H* and the reference samples. The PCR products were run on a 5.5% Long Ranger gel (0.5 mm, 6 M urea) at 800 V, 60 mA, and 49°C for 280 min using internal (202 bp and 266 bp) and external (266 bp ladder) size markers. The PCR 5 product was also used for RFLP analysis with the restriction enzyme *DdeI* for identification of *CSN1S1*H*.

To supplement mRNA and DNA sequencing results, protein isolation and mass spectrometry analysis were performed simultaneously for the *CSN1S1 C''C''*, *CH*, and *HH* samples according to Chianese et al. (2009).

Bioinformatic analysis of nucleotide and AA sequences were performed with the programs ChromasPro 1.32 (Technelysium Pty Ltd., Queensland, Australia), DNASTAR (DNAstar Inc., Madison, WI), and NNSPLICE 0.9 (Reese et al., 1997).

Nucleotide sequences of cloned PCR fragments revealed that exon 8 was missing in *CSN1S1*H* cDNA/mRNA (GenBank acc. no. FJ440846) compared with GenBank acc. no. NM_001009795 and GenBank acc. no. FJ440845, which is the full-length coding sequence of *CSN1S1*C''*, containing exon 16, established in this study for the first time. The lack of exon 8 coding for 8 AA (AA 51-58; DQAMEDAK), which does not influence the reading frame, resulted in a truncated protein. Mass spectrometry analysis of HPLC-purified dephosphorylated H variant provided a molecular mass of 21,872 Da, which indicated the presence of a deleted variant of α_{s1}-CN that lacked the sequence 51 to 58; this was further confirmed by tandem mass spectrometry analysis of tryptic digest of the variant protein (data not shown).

Analysis of the genomic ovine *CSN1S1* sequence from exon 7 to exon 9 showed that the sequence corresponding to exon 8 is present in the DNA sequence. The alignment of *CSN1S1*H* with *CSN1S1*A*, *C''*, and *D* DNA sequences showed a deletion of 4 nucleotides (the last triplet of exon 8 and the first nucleotide of intron 8) and an insertion of 13 bp (g.739_742delAAGGinsTTATTTTAATAAA) in *CSN1S1*H* (Figure 1). This insertion is a precise duplication of the adjacent intronic sequence (g.751-763) of *CSN1S1*C''* (GenBank acc. no. FJ440847). These changes delete the 5'-donor splice site of intron 8 with its special AG/GT-configuration (g.740-747; GenBank acc. no. FJ440847) in *CSN1S1*H* as shown in Figure 1.

In the junction of exon 8 and intron 8 in *CSN1S1*H* no splice site could be identified in silico by NNSPLICE 0.9 (Reese et al., 1997). The sequence differences in ovine *CSN1S1*H* at the 5'-splice site alter the U1 small nuclear RNA recognition sequence and inhibit accurate splice site selection and the splicing process.

Table 1. Primer sequences, reaction conditions, and resulting product sizes used for characterization of ovine *CSN1S1* DNA sequence from exon 7 to exon 9

PCR number	PCR name[1]	Primer forward, 5'- 3' (GenBank acc. no. and location)	Primer reverse, 5'- 3' (GenBank acc. no. and location)	Mg concentration (mM)	AT[2] (°C)	Cycles (n)	Product size (bp)
1	Ex7F-In8R	gggagtgaatcaattgaggtaag (NM_001009795; r.239-256 in combination with X59856; g.16685-16689)	caattgcagtggatagctttctt (X59856; g. 17487-17465)	1.0	60	45	821
2	In7F-In8bR	ttctctattgccacccatttcta (FJ440847; g.603-625)	actgactgaatggacgtgagttt (X59856; g.17904-17882)	3.0	59	35	643
3	In8F-In8cR	gctatgctgttgttcagtcactc (FJ440847; g.1126-1148)	gaacaatcaatttcagaatggaa (X59856; g.18478-18456)	1.0	56	35	699
4	In8bF-In9R	cccaagagatgactattttgctg (FJ440847; g.1639-1661)	tgttttcttacattttggggttg (X59856; g.19081-19103)	1.0	53	35	805
5	In7FCy5-In8dR	atttcctttggcatccattttat (FJ440847; g.625-646)	tagcactgcttgaggagttcaat (FJ440847; g.855-832)	2.0	60	35	231

[1] Ex = exon; In = intron.
[2] AT = annealing temperature.

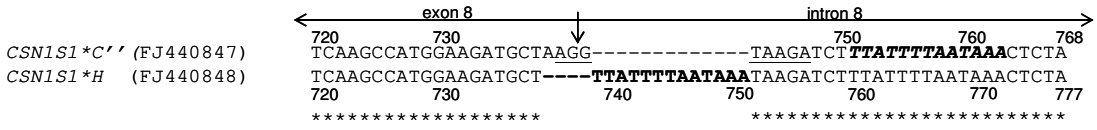


Figure 1. Genomic DNA sequence alignment of ovine *CSN1S1* variants *C''* (nucleotides 720-768; GenBank acc. no. FJ440847) and *H* (nucleotides 720-777; GenBank acc. no. FJ440848). Stars mark identical nucleotides. Blanks within the sequence of *CSN1S1*C''* mark a region where an insertion in *CSN1S1*H* occurred. Blanks in *CSN1S1*H* mark a deletion of the last 3 nucleotides of exon 8 and the first nucleotide of intron 8. Bold letters mark an insertion in *CSN1S1*H* corresponding to intron 8 nucleotides 751-763 in *CSN1S1*C''*. Underlined nucleotides display the 5'-donor splice site of intron 8 in *CSN1S1*C''*. The arrow indicates the junction of exon 8 and intron 8.

The characterization of *CSN1S1*H* with a deletion in the exon and an insertion in the intron DNA sequence describes an ovine CN variant caused by exon skipping, which is different in comparison with *CSN1S1* in cattle (Mohr et al., 1994) and goats (Leroux et al., 1992; Grosclaude and Martin, 1997), where SNP in the splice donor sequence of the following intron are responsible for alternative splicing.

Fragment length analysis of PCR 5 revealed sequence length differences of 9 nucleotides between *CSN1S1*H* and *C''* (nonallele H: 231 bp; allele H: 240 bp; Figure 2).

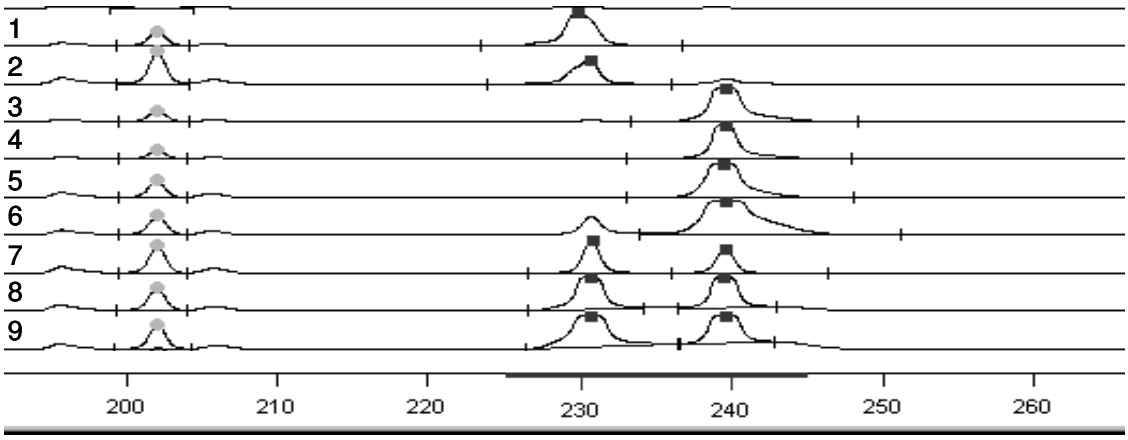


Figure 2. Fragment length analysis of ovine *CSN1S1* using an A.L.F.express sequencer and AlleleLocator 1.03 software (both Amersham Biosciences, Freiburg, Germany). Peak at position of 202 bp indicates external marker. Peak at position of 231 bp marks non-*CSN1S1*H*, whereas peak at position of 240 bp marks *CSN1S1*H*. Lanes 1 and 2: non-*CSN1S1*H*; lanes 3, 4, 5, and 6: *CSN1S1 HH*; lanes 7, 8, and 9: *CSN1S1*H*-carrier.

These results were confirmed by PCR-RFLP analyses, whereas restriction of PCR 5 product with *Dde*I resulted in a control fragment of 89 bp in all cases, 2 additional fragments in samples without *CSN1S1* allele *H* (24 bp and 118 bp), and only 1 additional fragment in carriers of *CSN1S1*H* (151 bp; Figure 3).

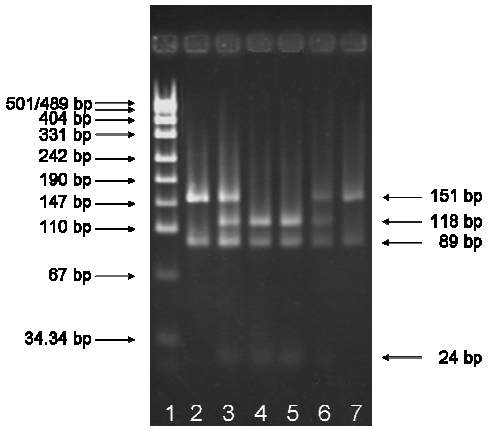


Figure 3. Separation of *Dde*I-digested PCR products in a 3.5% agarose gel stained with ethidium bromide. Lane 1: marker; lanes 2, 7: *CSN1S1 HH*; lanes 3, 6: *CSN1S1 C''H*; lanes 4, 5: *CSN1S1 C''C''*.

Fragment length and PCR-RFLP analyses can be used as DNA tests for the identification of *CSN1S1*H* in animals independent from age, sex, and lactation.

The widespread missing of short exon sequences in different CN and in different species could be supported by the extremely split architecture of *CSN1S1* and *CSN1S2* in various species with small exons in the coding region (Rijnkels, 2002), caused by evolutionary inter- and intragenic duplications (Groenen et al., 1993).

The possible effects of *CSN1S1*H* on milk composition, milk yield, and cheese making abilities are the subject of ongoing studies. This is interesting because of the correlation between abbreviated caprine *CSN1S1*F* and *G* and reduced protein content (Leroux et al., 1992; Grosclaude and Martin, 1997) and the correlation between bovine *CSN1S1*A* and lower cheese-making ability and cheese quality (Ng-Kwai-Hang, 2006). Quantitative analysis of the expressed protein by integration of mass ion peaks of dephosphorylated CN resulted in a 74% lower expression level for α_{s1}-CN H compared with the reference *C″* variant (data not shown). This is in agreement with the results of Wessels et al. (2004), who found significantly lower CN content in milk from sheep with *CSN1S1 CX*, which is identical to the *CSN1S1* genotype *CH*, compared with *CSN1S1 CC* sheep.

ACKNOWLEDGEMENTS

We thank Gesine Lühken (Department of Animal Breeding and Genetics, Justus-Liebig-University Gießen, Germany) and the anonymous reviewers for valuable comments.

REFERENCES

Boettcher, P. J., A. Caroli, A. Stella, S. Chessa, E. Budelli, F. Canavesi, S. Ghiroldi, and G. Pagnacco. 2004. Effects of casein haplotypes on milk production traits in Italian Holstein and Brown Swiss cattle. J. Dairy Sci. 87:4311-4317.

Boutinaud, M., H. Rulquin, D. H. Keisler, J. Djiane, and H. Jammes. 2002. Use of somatic cells from goat milk for dynamic studies of gene expression in the mammary gland. J. Anim. Sci. 80:1258-1269.

Ceriotti, G., F. Chiatti, R. Bolla, M. Martini, and A. Caroli. 2005. Genetic variability of the ovine α_{s1}-casein. Ital. J. Anim. Sci. 4:64-66.

Chianese, L., S. Caira, G. Garro, and F. Addeo. 2007. Primary structure of ovine deleted variant α_{s1}-CN E. Page 70 in Proc. 5th International Symposium on the Challenge to Sheep and Goats Milk Sectors, Alghero-Sardinia. Italy. International Dairy Federation, Brussels, Belgium.

Chianese, L., G. Garro, R. Mauriello, P. Laezza, P. Ferranti, and F. Addeo. 1996. Occurrence of five α_{s1}-casein variants in ovine milk. J. Dairy Res. 63:49-59.

Chianese, L., M. Quarto, F. Pizzolongo, M. G. Calabrese, S. Caira, R. Mauriello, S. De Pascale, and F. Addeo. 2009. Occurrence of genetic polymorphism at the α_{s1}-casein locus in Mediterranean water buffalo milk. Int. Dairy J. 19:181-189.

Ferranti, P., A. Malorni, G. Nitti, P. Laezza, R. Pizzano, L. Chianese, and F. Addeo. 1995. Primary structure of ovine α_{s1}-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. J. Dairy Res. 62:281-296.

Giambra, I. J., S. Jäger, and G. Erhardt. 2009. Isoelectric focusing reveals additional casein variants in German sheep breeds. Small Rumin. Res. (accepted)

Groenen, M. A., R. J. Dijkhof, A. J. Verstege, and J. J. Van Der Poel. 1993. The complete sequence of the gene encoding bovine α_{s2}-casein. Gene 123:187-193.

Grosclaude, F., and P. Martin. 1997. Casein polymorphisms in the goat. Pages 241-253 in Milk Protein Polymorphism. International Dairy Federation, Brussels, Belgium.

Hallén, E., T. Allmere, J. Naslund, A. Andren, and A. Lunden. 2007. Effect of genetic polymorphism of milk proteins on rheology of chymosin-induced milk gels. Int. Dairy J. 17:791-799.

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

Ikonen, T., K. Ahlfors, R. Kempe, M. Ojala, and O. Ruottinen. 1999. Genetic parameters for the milk coagulation properties and prevalence of noncoagulating milk in Finnish dairy cows. J. Dairy Sci. 82:205-214.

Leroux, C., N. Mazure, and P. Martin. 1992. Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat α_{s1}-casein transcripts. Structural organization of the relevant gene. J. Biol. Chem. 267:6147-6157.

Lühken, G., A. Caroli, E. M. Ibeagha-Awemu, and G. Erhardt. 2009. Characterization and genetic analysis of bovine α_{s1}-casein *I* variant. Anim. Genet. 40:479-485.

Mahé, M. F., G. Miranda, R. Queval, A. Bado, P. S. Zafindrajaona, and F. Grosclaude. 1999. Genetic polymorphism of milk proteins in African Bos taurus and Bos indicus populations. Characterization of variants α_{s1}-Cn H and κ-Cn J. Genet. Sel. Evol. 31:239-253.

Martin, P., M. Szymanowska, L. Zwierzchowski, and C. Leroux. 2002. The impact of genetic polymorphisms on the protein composition of ruminant milks. Reprod. Nutr. Dev. 42:433-459.

Mohr, U., D. Koczan, D. Linder, G. Hobom, and G. Erhardt. 1994. A single point mutation results in A allele-specific exon skipping in the bovine α_{s1}-casein mRNA. Gene 143:187-192.

Montgomery, G. W., and J. A. Sise. 1990. Extraction of DNA from sheep white blood cells. N. Z. J. Agric. Res. 33:437-441.

Ng-Kwai-Hang, K. F. 2006. Genetic variants of milk proteins and their effects on the yield and quality of cheese. CAB Reviews 56:1-11. doi:10.1079/PAVSNNR20061056

Pirisi, A., G. Piredda, C. M. Papoff, R. Di Salvo, S. Pintus, G. Garro, P. Ferranti, and L. Chianese. 1999. Effects of sheep α_{s1}-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. J. Dairy Res. 66:409-419.

Reese, M. G., F. H. Eeckman, D. Kulp, and D. Haussler. 1997. Improved splice site detection in Genie. J. Comput. Biol. 4:311-323.

Rijnkels, M. 2002. Multispecies comparison of the casein gene loci and evolution of casein gene family. J. Mammary Gland Biol. Neoplasia 7:327-345.

Shin, H. J., H. Y. Park, S. J. Jeong, H. W. Park, Y. K. Kim, S. H. Cho, Y. Y. Kim, M. L. Cho, H. Y. Kim, K. U. Min, and C. W. Lee. 2005. STAT4 expression in human T cells is regulated by DNA methylation but not by promoter polymorphism. J. Immunol. 175:7143-7150.

Wessels, G., H. Hamann, G. Erhardt, and O. Distl. 2004. Genotypeffekte von Milchprotein-Polymorphismen auf die Milchleistung beim Ostfriesischen Milchschaf. Berl. Munch. Tierarztl. Wochenschr. 117:414-419.

5. Molecular genetic characterisation of ovine *CSN1S2* variants *C* and *D* revealed further variability within *CSN1S2*

I. J. Giambra and G. Erhardt

Department of Animal Breeding and Genetics, Justus-Liebig-University, 35390 Gießen, Germany

Submitted to:
Animal Genetics

Submitted: 28.09.2010

Summary

Molecular knowledge about genetic variation of ovine casein genes, especially α_{s2}-casein (*CSN1S2*), and therefore resulting integration in animal breeding are very limited. Within this study recently identified ovine *CSN1S2* variants *C* and *D* were characterised on molecular genetic level. Sequencing of the mRNA/cDNA and of parts of DNA identified several sequence differences within *CSN1S2*C* and *D* in comparison to *CSN1S2*A* and *B*. *CSN1S2*C* is characterised by two non-synonymous SNPs lying within exon 7 leading to the amino acid substitutions p.Val45Ile and p.Ala48Ser. *CSN1S2*D* is caused by a SNP also located in exon 7 resulting in an amino acid replacement at position 46 (p.Arg46Ser) in mature protein and additionally by an insertion of one single nucleotide in intron 6. Furthermore, compared to *CSN1S2*A* and *B*, in both variants *CSN1S2*C* and *D* a SNP within 5´-untranslated region (exon 2) was identified, which is associated with five further SNPs in intron 1 and intron 2 and an insertion of 100 bp in intron 2.

Molecular genetic analyses of ovine *CSN1S2* lead to the identification of an additional very common G>A-SNP within exon 15 resulting in an amino acid substitution (p.Arg161His), causing new *CSN1S2*G*, not detectable by isoelectric focusing and misidentified as *CSN1S2*A* before. Additionally a sub-form of *CSN1S2*A*, *A'*, was detected and implemented in a possible phylogenetic pathway of ovine *CSN1S2*.

On the basis of the identified sequence differences DNA-based tests for *CSN1S2*B*, *C*, *D*, and *G* were established and a new nomenclature was proposed.

Kewords ovine, casein, *CSN1S2*, phylogenetic pathway, nomenclature.

Introduction

In the recent past additional ovine casein (CN) variants, caused by several polymorphisms not known before, were identified (Picariello *et al.* 2009; Chessa *et al.* 2010; Giambra *et al.* 2010a; 2010b). This is caused by the increased interest in milk protein polymorphisms in ruminants as they show influence on production traits (Pirisi *et al.* 1999; Martin *et al.* 2002; Moioli *et al.* 2007; Caroli *et al.* 2009) and on nutrition-physiological aspects of milk products (Meisel 1998; Cocco *et al.* 2003; Park *et al.* 2007). Furthermore, they can be used as population genetic markers (Mahé *et al.* 1999; Ibeagha-Awemu *et al.* 2007) and are therefore of interest for animal breeding.

The α_{s2}-CN (CSN1S2) is one of four caseins (α_{s1}-, α_{s2}-, β-, and κ-CN), coded by *CSN1S2* gene, lying within the tightly linked cluster on ovine chromosome 6 (Threadgill & Womack

1990; Lévéziel *et al.* 1991; Bevilacqua *et al.* 2006). Ovine α_{s2}-CN consists of 208 amino acids in mature protein and is the most highly phosphorylated of the calcium-sensitive caseins (Martin *et al.* 2003). It is coded by 1024 nucleotides of mRNA excluding the poly-A-tail (Boisnard & Petrissant 1985). Furthermore, ovine α_{s2}-CN is characterised by simultaneous occurrence of two isoforms, differing by an internal deletion of nine amino acid residues (Boisnard *et al.* 1991). However, knowledge about α_{s2}-CN variation on protein and DNA level in sheep is fragmentary and partly not consistent. Sequencing of *CSN1S2* mRNA/cDNA revealed four mRNAs, differing in the presence or absence of a stretch of 44 bp in the 5'-untranslated region (UTR) and in nucleotide substitutions at positions 243 and 698 (GenBank accession number X03238) concerning the amino acid residues p.49 and p.200 (Asn49Asp and Lys200Asn) in mature protein (Boisnard *et al.* 1991). However, Swiss-Prot accession number P04654 only mentions sequence variation Asn49Asp, as Lys200Asn occurred only in combination with the internal stretch of 44 bp (Boisnard *et al.* 1991). Independently of the 5'-UTR 44 bp stretch, protein forms with the combinations Asp_{49}/Asn_{200}, Asn_{49}/Asn_{200}, and Asn_{49}/Lys_{200} are described. Amigo *et al.* (2000) named two of them CSN1S2 A (Asn_{49}/Lys_{200}) and B (Asp_{49}/Asn_{200}). Swiss-Prot accession number P04654 shows an amino acid sequence conflict between sequences of Boisnard and Petrissant (1985) and Boisnard *et al.* (1991) concerning the amino acid residue 79 (Thr79Tyr) of mature protein, whereas this is caused by a mistake in the amino acid translation of mRNA sequence in Boisnard *et al.* (1991). Both published sequences (GenBank accession numbers X03238 and NM_001009363) are characterised by Tyr79, which was confirmed by Picariello *et al.* (2009) recently.

An additional electrophoretic pattern of α_{s2}-CN showed a lower isoelectric point (pI), than the most common variant, not named (Chianese *et al.* 1993). This pattern was later on designated by Amigo *et al.* (2000) as variant "Fast" and was characterised in a further study by a lower molecular weight, whereas the authors presumed a deletion in the amino acid sequence (Recio *et al.* 1997).

Isoelectric focusing (IEF) of ovine milk protein complex in Italian sheep breeds lead within α_{s2}-CN to the differentiation of two variants (Chessa *et al.* 2003) called CSN1S2 A and B as well. These variants were further analysed recently and protein sequence studies revealed two amino acid substitutions concerning the residues 75 and 105 (p.Asp75Tyr and p.Ile105Val within CSN1S2 B) of mature protein leading to different pIs of CSN1S2 A and B, with CSN1S2 A as the more acidic one (Picariello *et al.* 2009).

Polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP) methodology was accomplished for *CSN1S2* exon 16 and three PCR-SSCP patterns were found and named *CSN1S2*N*, *CSN1S2*X*, and *CSN1S2*Y* (Chessa *et al.* 2010). Sequence analysis revealed that *CSN1S2*N*, the most common allele, corresponds to the published *CSN1S2* sequence (GenBank accession number X03238) of Boisnard *et al.* (1991) with p.Asn200. *CSN1S2*X* is most probably an isoconformer of *CSN1S2*N* and *CSN1S2*Y* is characterised by a C>G transversion concerning the already described amino acid change Asn_{200}>Lys_{200} and by a T>A-single nucleotide polymorphism (SNP) within intron 16 (Chessa *et al.* 2010).

Considering the nomenclature of Chessa *et al.* (2003) recently two further α_{s2}-CN variants C and D were identified in German sheep breeds by IEF and their genetic control was confirmed by segregation studies (Giambra *et al.* 2010c).

Therefore, the aim of this work was to i) characterise the two new α_{s2}-CN protein patterns C and D on mRNA and DNA level, ii) describe the mRNA and DNA sequence differences of *CSN1S2*A* and *B* (after Chessa *et al.* (2003)), iii) develop DNA-based tests for *CSN1S2* genotyping also in non-lactating females or males, iv) propose uniformed nomenclature, and v) establish a possible phylogenetic pathway of ovine *CSN1S2*.

Materials and methods

Samples

Milk samples of 28 animals with different CSN1S2 genotypes (AA: n = 5; AB: n = 6; AC: n = 6; AD: n = 6; BB: n = 3; BC: n = 1; BD: n = 1) typed by IEF were used for isolation of mRNA and for mRNA sequence studies. For additional DNA analyses blood samples of further 26 animals with the following CSN1S2 genotypes were available: AA: n = 3; AB: n = 3; AC: n = 5; AD: n = 4; BB: n = 4; BC: n = 3; BD: n = 4. Nomenclature was made according to Chessa *et al.* (2003) and Giambra *et al.* (2010c). Allelic frequency of a nucleotide substitution identified within this study (*CSN1S2*G*) and described below was estimated by typing 30 animals of five different breeds (Black Faced Mutton sheep (SKF), OMS, GGH, MLS, Merino Mutton sheep (MMS)), and of 43 RHO, whereas 20 of these RHO were half-sibs and were additionally used for segregation analyses.

mRNA isolation, reverse transcription, and sequence analyses

From each of the 28 milk samples mRNA was isolated and reverse transcribed into cDNA according to Giambra *et al.* (2010a). Whole coding sequence was amplified by PCR No. 1

(Table 1) and subsequently sequenced bidirectionally with an ABI Prism 377 DNA and a 3130 Genetic Analyzer and the Big Dye Terminator sequencing kit v1.1 (all Applied Biosystems) using the same primers as used for PCR. Partly PCR products were sent for sequencing to Microsynth AG (Balgach).

Table 1 Primer sequences, reaction conditions and resulted product sizes for characterisation of mRNA and DNA sequences and development of DNA-based tests.

PCR-Number	PCR-name	Primer forward (5´- 3´) (GenBank accession number and location)	Primer reverse (5´- 3´) (GenBank accession number and location)	Mg-concentration (mM)	AT[1] (°C)	Cycles (n)	Product size (bp)
1	Ex[2]1F-Ex18R	cattgcctggactacttgtcttc (X03238; r.4–26)	caaacatgctggttgtatgaagt (X03238; r.827–805)	1.0	51	35	824
2	In[3]1For-In2Rev	catttccttctccaggttttaca (BOVCASAS2X; g.4121–4143)	tccttgttatggttccatgtttc (BOVCASAS2X; g.4739–4717)	2.0	61	35	619
3	In6For-In7Rev	ctgcggatacaacaataagcac (BOVCASAS2X; g.8170–8191)	acagggaaaaactgtcattttga (BOVCASAS2X; g.8569–8547)	3.0	60	35	400
4	In6bFor-In7bRev	acaatgaaagcgaatccagtcta (BOVCASAS2X; g.8258–8280)	aacatgttcccatatgccattac (BOVCASAS2X; g.8501–8479)	2.0	53	35	244
5	Ex7F-In7R	cttctctctttaaggaagttgt**c**ag (BOVCASAS2X; g.8372–8385 in combination with X03238; r.225–235)	tcattttgaagcaacttataaatcat (BOVCASAS2X; g.8555–8530)	3.0	53	35	184
6	Ex9F-In10R	ctgaagttgccccagaggta (X03238; r.286–302 in combination with BOVCASAS2X; g.9946–9948)	catttggagaagaagcagtgg (own unpublished sequencing results of ovine intron 10 in combination with BOVCASAS2X; g.10150–10133)	1.0	59	35	225
7	In13F-In15R	tgtcatatggacacagagaatgc (BOVCASAS2X; g.13919–13941)	ggtacctgtcttcgggataattt (BOVCASAS2X; g.14573–14551)	3.0	56	35	655

[1]Annealing temperature.
[2]Exon.
[3]Intron.
Bold and underlined: miss paired nucleotide for the establishment of an artificial restriction site.

DNA-isolation and sequence analyses

Isolation of genomic DNA was carried out from whole blood or buffy coat using the high-salt method of Montgomery and Sise (1990). DNA sequences surrounding exon 2 and exon 7 were amplified by PCR No. 2 and 3 (Table 1). First sequencing results of PCR No. 2 showed necessity of cloning before sequencing. In this case amplified PCR fragments were cloned with pGEM-T and pGEM-T Easy Vectors (Promega) according to Shin *et al.* (2005) and five recombinant clones of four animals each (CSN1S2 genotypes AA, BB, BC, and BD) were sequenced as described above.

Development of DNA-based tests

To develop DNA-based tests for *CSN1S2*C* and *D* the new PCR No. 4 and 5 (Table 1) were amplified for restriction fragment length polymorphism (RFLP) analyses. In the case of *CSN1S2*D* PCR No. 5 was based on amplification created restriction site (ACRS) (Haliassos *et al.* 1989; Lien *et al.* 1992) with a modified forward primer establishing an artificial restriction site for the restriction enzyme *Hpy188*III and was digested for 3 h at 37 °C. PCR No. 4, developed for DNA-based confirmation of *CSN1S2*C*, was also digested for 3 h at 37 °C with *Psi*I.

In order to screen for sequence differences causing *CSN1S2*B* and/or new nucleotide substitutions, identified within mRNA/cDNA sequence studies, PCRs No. 6 and 7 were established to perform a *Rsa*I/*Bme1390*I-double-digest and a *Hph*I-RFLP-analysis, respectively. Both PCR-RFLP analyses were carried out at 37 °C for 16 h. All digested fragments were separated on 3.5% agarose gels - with exception of the *Hph*I-RFLP, where a 2.0% agarose gel was sufficient - and detected after ethidium bromide staining.

Bioinformatic and statistical analyses

ChromasPro 1.32 (Technelysium Pty Ltd) and DNASTAR (DNASTAR) were used for mRNA/cDNA and DNA sequence analysis. Allele and genotype frequencies were calculated with program PopGene V 1.31 (Yeh *et al.* 1997).

Results

mRNA/cDNA sequence analyses

Sequencing of mRNA/cDNA revealed several sequence differences in the coding region and within the 5'- and 3'-UTR of ovine *CSN1S2* (Table 2). SNPs concerning the nucleotides 299 and 389 (GenBank accession numbers GU169085 and GU169086) leading to the known amino acid exchanges at positions 75 and 105 of CSN1S2 A and B variants were identified. In animals with *CSN1S2*C* a $G_{209}>A_{209}$ and a $G_{218}>T_{218}$ substitution (GenBank accession number GU169087), leading to differences in the amino acid sequence at positions 45 (p.Val45Ile) and 48 (p.Ala48Ser) respectively, were detected.

Table 2 Single nucleotide polymorphisms and their influence on deduced amino acid sequence in ovine *CSN1S2* mRNA subdivided according to *CSN1S2* genotype.

IEF genotype	mRNA genotype	*n*	SNP[1] (*Amino acid*[2])											
			22 ([3])	209 (*45*)	214 (*46*)	218 (*48*)	221 (*49*)	299 (*75*)	322 (*82*)	389 (*105*)	558 (*161*)	667 (*200*)	708 ([4])	759 ([5])
AA	AA	1	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	GG (*RR*)	CC (*NN*)	TT	GG
AA	AG	1	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	AG (*HR*)	CC (*NN*)	TT	CC
AA	AG	2	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	AG (*HR*)	CC (*NN*)	TT	CG
AA	GG	1	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	AA (*HH*)	CC (*NN*)	CC	CC
AB	BG	1	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GT (*DY*)	AC (*AA*)	AG (*IV*)	AG (*HR*)	CC (*NN*)	TT	CC
AB	BG	2	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GT (*DY*)	AC (*AA*)	AG (*IV*)	AG (*HR*)	CC (*NN*)	CT	CC
AB	AB	3	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	GT (*DY*)	AC (*AA*)	AG (*IV*)	GG (*RR*)	CC (*NN*)	TT	CG
AC	AC	2	GT	AG (*IV*)	GG (*RR*)	GT (*AS*)	AA (*NN*)	GG (*DD*)	AC (*AA*)	AA (*II*)	GG (*RR*)	CC (*NN*)	CT	CC
AC	CG	2	GT	AG (*IV*)	GG (*RR*)	GT (*AS*)	AA (*NN*)	GG (*DD*)	AC (*AA*)	AA (*II*)	AG (*HR*)	CC (*NN*)	TT	CC
AC	AC	2	GT	AG (*IV*)	GG (*RR*)	GT (*AS*)	AA (*NN*)	GG (*DD*)	AC (*AA*)	AA (*II*)	GG (*RR*)	CC (*NN*)	TT	CG
AD	AD	2	GT	GG (*VV*)	CG (*SR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	GG (*RR*)	CC (*NN*)	TT	CG
AD	DG	2	GT	GG (*VV*)	CG (*SR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	AG (*HR*)	CC (*NN*)	TT	CC
AD	AD	1	GT	GG (*VV*)	CG (*SR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	GG (*RR*)	CC (*NN*)	CT	CC
AD	DG	1	GT	GG (*VV*)	CG (*SR*)	GG (*AA*)	AA (*NN*)	GG (*DD*)	AA (*AA*)	AA (*II*)	AG (*HR*)	CC (*NN*)	CT	CC
BB	BB	1	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	TT (*YY*)	CC (*AA*)	GG (*VV*)	GG (*RR*)	CC (*NN*)	CT	CC
BB	BB	2	TT	GG (*VV*)	GG (*RR*)	GG (*AA*)	AA (*NN*)	TT (*YY*)	CC (*AA*)	GG (*VV*)	GG (*RR*)	CC (*NN*)	TT	CC
BC	BC	1	GT	AG (*IV*)	GG (*RR*)	GT (*AS*)	AA (*NN*)	GT (*DY*)	CC (*AA*)	AG (*IV*)	GG (*RR*)	CC (*NN*)	TT	CC
BD	BD	1	GT	GG (*VV*)	CG (*SR*)	GG (*AA*)	AA (*NN*)	GT (*DY*)	AC (*AA*)	AG (*IV*)	GG (*RR*)	CC (*NN*)	TT	CC

[1]SNP position in GenBank accession numbers GU169085–GU169089.
[2]Amino acid (1-letter-code) position in mature protein.
[3]In 5'-UTR.
[4]In 3'-UTR.

*CSN1S2*D* carriers showed a G>C nucleotide substitution at position 214 in mRNA sequence (GenBank accession number GU169088) causing the amino acid exchange p.Arg46Ser in mature protein product. In comparison to *CSN1S2*A* and *B CSN1S2*C* and *D* carriers both are additionally characterised by a T_{22}>G_{22}-SNP (GenBank accession numbers GU169087 and GU169088) lying within the 5'-UTR of *CSN1S2* mRNA (exon 2).

In addition one synonymous A_{322}>C_{322}-SNP was identified in mRNA/cDNA (p.Ala82; GenBank accession numbers GU169086 and GU169087) in animals carrying *CSN1S2*B* and *C*.

Furthermore, mRNA/cDNA sequencing lead to the detection of an additional G>A-SNP at mRNA position 558 (GenBank accession number GU169089) positioned in exon 15 in 12 animals with CSN1S2 IEF genotypes AA, AB, AC and AD, whereas animals with CSN1S2 BB, BC, and BD did not show this sequence variation. This transversion lead to the amino acid exchange p.Arg161His, causing a further variant, not identifiable by IEF, which we called *CSN1S2*G*.

All sequenced samples with the CSN1S2 genotypes AA, AB, AC, AD, BB, BC, and BD carried A_{221} in mRNA sequence (GenBank accession numbers GU169085–GU169089) instead of G_{243} in reference sequence (GenBank accession number X03238) and with this p.Asn49 instead of p.Asp49.

Additionally two nucleotide substitutions lying in the 3'-UTR (T_{708}>C_{708} and C_{759}>G_{759}; GenBank accession numbers GU169086 and GU169089) located in exon 17 and 18 respectively, were identified. These SNPs are not linked to any *CSN1S2* variant.

DNA sequence analyses

Sequence analyses of PCR No. 2 and 3 (Table 1) confirmed the occurrence of the SNPs within exon 2 and exon 7 identified by mRNA/cDNA sequencing and revealed several further sequence differences also concerning *CSN1S2*C* and *D*. Cloning and sequencing of PCR No. 2 identified an insertion of 100 bp (Fig. S1) in intron 2 in both *CSN1S2* variants *C* and *D* (g.550_551ins100; GenBank accession number GU169092 or g.551-650; GenBank accession number GU169094) with a high content of adenine and thymine (90%). This insertion is furthermore associated with five SNPs within intron 1 and intron 2 (g.22C>T, g.31A>G, g.449T>A, g.469G>A, and g.500C>T; GenBank accession number GU169094).

In addition all *CSN1S2*D* samples showed an insertion of one adenine residue in intron 6 (Genbank accession number GU169091) in comparison to reference samples (Genbank accession number GU169090; g.186_187insA).

Sequencing of reference samples revealed further variability within ovine *CSN1S2* demonstrated in GenBank accession numbers GU169092 and GU169093.

DNA-based tests

PCR-RFLP analyses of PCR No. 4 (244 bp) with *Psi*I was based on the detection of the G>A-SNP on position 209 (GenBank accession number GU169087) characterising *CSN1S2*C*. PCR products of samples with *CSN1S2*C* were hydrolysed after *Psi*I-digestion into two fragments of 109 bp and 135 bp while PCR products of animals with the remaining *CSN1S2* variants were not digested (Fig. S2). ACRS-PCR No. 5 (184 bp) followed by *Hpy188*III resulted in fragments of 23 bp and 161 bp in all *CSN1S2* genotypes with exception of *CSN1S2*D*, where the PCR product kept its original size caused by the $G_{214}>C_{214}$-SNP (Fig. S3). The results are in full agreement with the mRNA sequencing results.

*Rsa*I/*Bme1390*I-double-digest was used as DNA-based test for the identification of *CSN1S2*B* and the $A_{322}>C_{322}$-SNP concerning p.Ala82. *Rsa*I-RFLP analysis lead to the occurrence of a 22 bp control fragment in all cases and two further fragments of PCR No. 6 product (71 bp and 132 bp) in samples with *CSN1S2*B*. Samples with *CSN1S2*non-B* alleles were not further digested. *Bme1390*I-RFLP identified the $A_{322}>C_{322}$-SNP and lead to fragments of 69 bp and 156 bp in C_{322}-carriers. Double digestion resulted in fragments of 22 bp and 203 bp in *CSN1S2* alleles *A*, *D*, and *G*, in fragments of 22 bp, 24 bp, 47 bp, and 132 bp in allele *B* and in fragments of 22 bp, 47 bp, and 156 bp in allele *C.* In samples with *CSN1S2* genotypes *AB*, *BD*, and *BG* the 156 bp pattern still occurs without diagnostic relevance (Fig. S4).

Segregation and allelic frequencies of *CSN1S2*G* determined by DNA-based test

*Hph*I-RFLP-analysis of PCR product No. 7 (Table 1) was used as DNA-based test for the identification of the $G_{558}>A_{558}$-SNP, leading to *CSN1S2*G*. Fragments of 352 bp and 303 bp length were generated in the case of *CSN1S2*G* (Fig. S5).

This *Hph*I-RFLP-digest of PCR No. 7 was used to prove segregation of *CSN1S2*G* by typing the informative family of 20 half-sibs. *CSN1S2* genotype distribution within these 20 animals was as follows: *AB*: $n = 1$; *AG*: $n = 5$; *BG*: $n = 6$; *GG*: $n = 8$. Therefore, it could be concluded that the common father-ram should carry *CSN1S2* genotype *BG* and that a codominant autosomal inheritance is existing ($\chi^2 = 6.28$; degree of freedom = 6; $P \leq 0.5$).

Additionally, this DNA-based test was also used for estimating allele frequency of *CSN1S2*G* in different breeds and revealed frequencies between 0.017 and 0.383 (Table 3).

Table 3 Allelic frequency of *CSN1S2*G* within different sheep breeds typed by PCR-RFLP. *CSN1S2*G* can not be distinguished from *CSN1S2*A* by IEF and therefore the frequency of *CSN1S2*A* is overestimated.

Breed	*CSN1S2*G*	*CSN1S2*A*[1]
Black Faced Mutton	0.300	0.421
East Friesian Dairy sheep	0.017	0.474
Gray Horned Heath	0.117	0.926
Merinoland sheep	0.317	0.786
Merino Mutton sheep	0.383	0.909
Rhön sheep	0.350	0.730

[1]According to Giambra *et al.* (2010c).

Discussion

Molecular analyses showed that ovine *CSN1S2* is more polymorphic as demonstrated until now. The reasonable sequence differences of *CSN1S2*C* and *D* in comparison to the published mRNA sequences (GenBank accession numbers X03238 and NM_001009363) were revealed. These SNPs, causing amino acid sequence differences of CSN1S2 C and D are all located in exon 7 of the *CSN1S2* gene. Ovine CSN1S2 C is caused by two amino acid substitutions (p.Val45Ile and p.Ala48Ser) in comparison to CSN1S2 A (Picariello *et al.* 2009). Latter amino acid exchange leads to a lower pI of CSN1S2 C, which is in accordance with the IEF results (Giambra *et al.* 2010c). CSN1S2 D (Giambra *et al.* 2010c) is caused by the p.Arg46Ser substitution. In this context phosphorylation degree of CSN1S2 C and D should be clarified, as it has an effect on cheese-making aptitude of milk by influencing solubility of casein in the presence of calcium (Chianese *et al.* 1996; Chianese *et al.* 2007).

Combination of sequence analyses on the level of mRNA/cDNA and DNA and IEF results revealed full agreement between molecular genetic and biochemical methods of analysing ovine CSN1S2 A, B, C, and D. The amino acid exchange p.Arg161His in CSN1S2 G using IEF is not associated with a difference in charge, as both amino acid residues are basic. Therefore, using IEF both CSN1S2 G and A are not separated due to identical isoelectric points. CSN1S2 A is overestimated until now as allelic frequency of *CSN1S2*G* is high in several breeds (Table 3). Lowest frequency of *CSN1S2*G* occurred in OMS, a breed highly selected for milk performance traits.

Additionally, we could not confirm the sequence variation concerning the amino acids 49 and 200 described by Boisnard *et al.* (1991) and Chessa *et al.* (2010). All sequenced cDNA ($n = 28$) and DNA ($n = 26$) samples carried Asn49 and all cDNA sequences coded for

Asn200. Picariello *et al.* (2009) sequenced peptides containing amino acid residues p.75 and p.105 in CSN1S2 A and B. They published complete amino acid sequences of these protein variants, both carrying Asp49 and Asn200, by using GenBank accession number X03238 to complete peptide sequences of CSN1S2 A and B. These forms of complete amino acid sequences were not confirmed in the present study in CSN1S2 A and B, characterised by Asn49, Asp75Tyr, Ile105Val, and Asn200.

```
                          1                                                 45
Sheep          MKFFIFTCLLAVALAKHKMEHVSSSEEPINISQEIYKQEKNMAIHPRKEKLCTTSCEEVV
Goat           ****************************F***************************V
Cattle         ***************NT*********S*\****T*********N*S**N**S*F*K**V

               46                            75                          105
Sheep          RNANEEEYSIRSSSEESAEVAPEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPI
Goat           R*AN*******************I****D************************P****I
Cattle         R*AN******G*********T*******D****************************I

               106                                                 161 165
Sheep          VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNHLNFL
Goat           ************************************************R****
Cattle         ***********V*I***L***************V*****************R****

               166                             200     208
Sheep          KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL
Goat           *********************************N********
Cattle         *****R*****L*******Y*********I****KV*******
```

Figure 1 Alignment of primary structure of ovine α_{s2}-CN variant G and caprine and bovine α_{s2}-CN reference variants.

Sheep = GenBank accession number GU169089, goat = Swiss-Prot accession number P33049, and cattle = Swiss-Prot accession number P02663.
Numbering according to GenBank accession number GU169089. Stars mark identical amino acids. The diagonal slash marks a missing amino acid residue within bovine sequence. Underlined Serin-residues are phosphorylated.

Ancestral CSN1S2 form seems to carry Asn49, Asp75, Ile105 (Picariello *et al.* 2009), and Asn200 in alignment with goat and cattle primary sequence of α_{s2}-CN (Fig. 1), corresponding to CSN1S2 A after revised nomenclature we proposed in Table 4.

Table 4 Ovine *CSN1S2* variants according to the proposed nomenclature.

CSN1S2 variant	Past nomenclature	GenBank accession number	Amino acid position									Reference
			45	46	48	49	75	82	105	161	200	
A	A (Chessa *et al.* 2003)	GU169085	Val	Arg	Ala	Asn	Asp	Ala	Ile	Arg	Asn	This paper and Picariello *et al.* (Picariello *et al.* 2009)
B	B (Chessa *et al.* 2003)	GU169086	Val	Arg	Ala	Asn	Tyr	Ala	Val	Arg	Asn	This paper and Picariello *et al.* (Picariello *et al.* 2009)
C	C (Giambra *et al.* 2010c)	GU169087	Ile	Arg	Ser	Asn	Asp	Ala	Ile	Arg	Asn	This paper
D	D (Giambra *et al.* 2010c)	GU169088	Val	Ser	Ala	Asn	Asp	Ala	Ile	Arg	Asn	This paper
E	B (Amigo *et al.* 2000)	X03238	Val	Arg	Ala	Asp	Asp	Ala	Ile	Arg	Asn	(Boisnard & Petrissant 1985; Boisnard *et al.* 1991)
F	A (Amigo *et al.* 2000)	-	Val	Arg	Ala	Asn	Asp	Ala	Ile	Arg	Lys	(Boisnard & Petrissant 1985; Boisnard *et al.* 1991)
G	-	GU169089	Val	Arg	Ala	Asn	Asp	Ala	Ile	His	Asn	This paper
A'	-	-	Val	Arg	Ala	Asn	Asp	Ala[1]	Ile	Arg	Asn	This paper

[1]Synonymous SNP within coding triplet for amino acid 82.

Necessity of standardising nomenclature within ovine *CSN1S2* was also mentioned by Picariello *et al.* (2009). Uniformed nomenclature can help to clarify evolutionary pathway. The G_{558}>A_{558}-SNP (GenBank accession number GU169089) of *CSN1S2*G* gave rise to the conclusion that *CSN1S2*G* can be a subtype of *CSN1S2*A*. This is supported by the amino acid sequence comparison of α_{s2}-CN protein between sheep, goat, and cattle, shows sequence similarity at amino acid positions 45, 46, 48, 49, 75, 105, and 200, whereas goat and cattle carried Arg_{161} and not His_{161} as in ovine *CSN1S2*G* (Fig. 1).

Restriction patterns after *Rsa*I/*Bme1390*I-double digestion confirmed the SNP at mRNA position 322 (GenBank accession numbers GU169086 and GU169087) being identical to nucleotide 344 of GenBank accession number X03238, occurring in animals with *CSN1S2*B* and *C*. Therefore, this mutation should be happened before the differentiation of *CSN1S2*A* in *B* or *C* and we concluded that this C_{322}-SNP characterises a sub-form of *CSN1S2*A* and nominated it as *CSN1S2*A'* (Fig. 2).

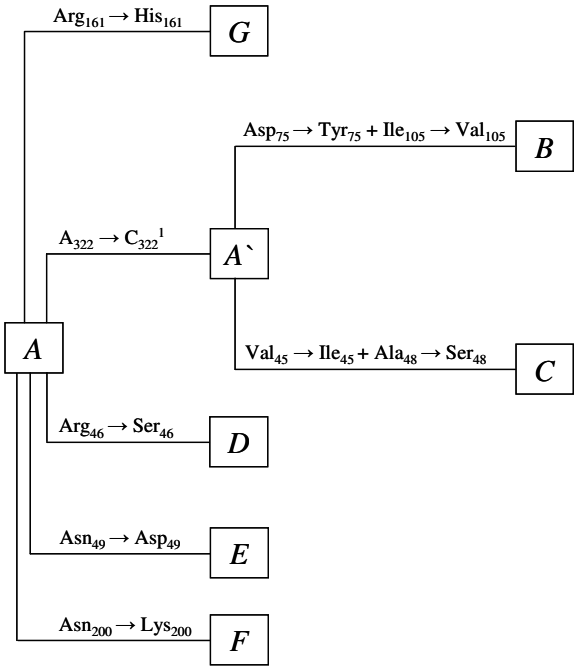


Figure 2 Possible evolutionary pathway of ovine *CSN1S2*.
[1]GenBank accession numbers GU169086 and GU169087.

Therefore, typing for *CSN1S2*B* with the *Rsa*I single RFLP-analysis will be sufficient for a correct typing of *CSN1S2*, using control animals to be sure that enzyme has worked correctly. However, evolutionary context between *CSN1S2*A'* and *CSN1S2*B* and *C* is still unclear with two mutations discriminating these alleles. This is comparable with the situation in bovine *CSN3*A* and *B*, where a precursor variant could be identified in *Bos grunniens* only

recently (Prinzenberg *et al.* 2008). This leads to the conclusion that ovine *CSN1S2* evolutionary pathway is still not complete and that more breeds, also the wild forms, have to be included in further studies.

Moreover, analysis are necessary to clarify if the protein form "Fast" (Chianese *et al.* 1993; Amigo *et al.* 2000) can be allocated to *CSN1S2*C* or *D*. In comparison to our results we can presume, that *CSN1S2*C* and CSN1S2 "Fast" might be identical, because they both show a lower pI in comparison to the most common variant, here defined as *CSN1S2*A*.

The established DNA-based tests on the basis of PCR-RFLP-analyses showed 100% agreement to mRNA/cDNA and DNA sequencing and IEF results leading to the conclusion, that they are suitable for typing sheep for *CSN1S2* independently of lactation. For simultaneous typing of the *CSN1S2* variants including the other milk protein polymorphisms described until now, the development of a microarray technology, as already used in bovine milk protein genes (Kaminski *et al.* 2005; Chessa *et al.* 2007) can be desirable depending on the SNPs in casein genes included in OvineSNP50 BeadChip (Illumina). This would also open the possibility to identify associations between *CSN1S2* variants and casein haplotypes, and economical and nutritional aspects of sheep milk production.

Acknowledgement

We thank Anna Caroli for providing reference samples and for fruitful discussion about nomenclature.

References

Amigo L., Recio I. & Ramos M. (2000) Genetic polymorphism of ovine milk proteins: its influence on technological properties of milk - a review. *International Dairy Journal* **10,** 135-49.

Bevilacqua C., Helbling J.C., Miranda G. & Martin P. (2006) Translational efficiency of casein transcripts in the mammary tissue of lactating ruminants. *Reproduction Nutrition Development* **46,** 567-78.

Boisnard M., Hue D., Bouniol C., Mercier J.C. & Gaye P. (1991) Multiple mRNA species code for two non-allelic forms of ovine α_{s2}-casein. *European Journal of Biochemistry/FEBS* **201,** 633-41.

Boisnard M. & Petrissant G. (1985) Complete sequence of ovine α_{s2}-casein messenger RNA. *Biochimie* **67,** 1043-51.

Caroli A.M., Chessa S. & Erhardt G.J. (2009) Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *Journal of Dairy Science* **92,** 5335-52.

Chessa S., Bolla P., Dario C., Pieragostini E. & Caroli A. (2003) Polimorfismi genetici lattoproteici nella razza ovina gentile di puglia: Monitoraggio mediante focalizzazione isoelectrica. *Scienza e Tecnica Lattiero-Casearia* **54,** 1-8.

Chessa S., Chiatti F., Ceriotti G., Caroli A., Consolandi C., Pagnacco G. & Castiglioni B. (2007) Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms. *Journal of Dairy Science* **90,** 451-64.

Chessa S., Rignanese D., Berbenni M., Ceriotti G., Martini M., Pagnacco G. & Caroli A. (2010) New genetic polymorphisms within ovine β- and α_{s2}-caseins. *Small Ruminant Research* **88,** 84-8.

Chianese L., Caira S., Garro G. & Addeo F. (2007) Primary structure of ovine deleted variant α_{s1}-CN E. *Proceedings of the International Symposium on the challenge to sheep and goats milk sectors* **5,** 58-60.

Chianese L., Garro G., Addeo F., Lopez-Galvez G. & Ramos M. (1993) Discovery of an ovine α_{s2}-casein variant. *Journal of Dairy Research* **60,** 485-93.

Chianese L., Garro G., Mauriello R., Laezza P., Ferranti P. & Addeo F. (1996) Occurrence of five α_{s1}-casein variants in ovine milk. *Journal of Dairy Research* **63,** 49-59.

Cocco R.R., Jarvinen K.M., Sampson H.A. & Beyer K. (2003) Mutational analysis of major, sequential IgE-binding epitopes in α_{s1}-casein, a major cow's milk allergen. *Journal of Allergy and Clinical Immunology* **112,** 433-7.

Giambra I.J., Chianese L., Ferranti P. & Erhardt G. (2010a) Genomics and proteomics of deleted ovine *CSN1S1*I*. *International Dairy Journal* **20,** 195-202.

Giambra I.J., Chianese L., Ferranti P. & Erhardt G. (2010b) Short Communication: Molecular genetic characterization of ovine α_{s1}-casein allele H caused by alternative splicing. *Journal of Dairy Science* **93,** 792-5.

Giambra I.J., Jäger S. & Erhardt G. (2010c) Isoelectric focusing reveals additional casein variants in German sheep breeds. *Small Ruminant Research* **90,** 11-7.

Haliassos A., Chomel J.C., Grandjouan S., Kruh J., Kaplan J.C. & Kitzis A. (1989) Detection of minority point mutations by modified PCR technique: a new approach for a sensitive diagnosis of tumor-progression markers. *Nucleic Acids Research* **17,** 8093-9.

Ibeagha-Awemu E.M., Prinzenberg E.M., Jann O.C., Lühken G., Ibeagha A.E., Zhao X. & Erhardt G. (2007) Molecular characterization of bovine *CSN1S2*B* and extensive distribution of zebu-specific milk protein alleles in European cattle. *Journal of Dairy Science* **90,** 3522-9.

Kaminski S., Ahman A., Ruść A., Wójcik E. & Malewski T. (2005) MilkProtChip - a microarray of SNPs in candidate genes associated with milk proteins biosynthesis - development and validation. *Journal of Applied Genetics* **46,** 45-58.

Lévéziel H., Metenier L., Guerin G., Cullen P., Provot C., Bertaud M. & Mercier J.C. (1991) Restriction fragment length polymorphism of ovine casein genes: close linkage between the α_{s1}-, α_{s2}-, β- and κ-casein loci. *Animal Genetics* **22,** 1-10.

Lien S., Alestrom P., Klungland H. & Rogne S. (1992) Detection of multiple beta-casein (CASB) alleles by amplification created restriction sites (ACRS). *Animal Genetics* **23,** 333-8.

Mahé M.F., Miranda G., Queval R., Bado A., Zafindrajaona P.S. & Grosclaude F. (1999) Genetic polymorphism of milk proteins in African Bos taurus and Bos indicus populations. Characterization of variants α_{s1}-Cn H and κ-Cn J. *Genetics, Selection, Evolution* **31,** 239-53.

Martin P., Ferranti P., Leroux C. & Addeo F. (2003) Non-bovine caseins: Quantitative variability and molecular diversity. In: *Advanced Dairy Chemistry Volume 1: Proteins* (ed. by P.F. Fox & P.L.H. McSweeney), pp. 277-317. Kluwer Academic/Plenum Publishers, New York.

Martin P., Szymanowska M., Zwierzchowski L. & Leroux C. (2002) The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reproduction Nutrition Development* **42,** 433-59.

Meisel H. (1998) Overview on milk protein-derived peptides. *International Dairy Journal* **8,** 363-73.

Moioli B., D' Andrea M. & Pilla F. (2007) Candidate genes affecting sheep and goat milk quality. *Small Ruminant Research* **68,** 179-92.

Montgomery G.W. & Sise J.A. (1990) Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research* **33,** 437-41.

Park Y.W., Juarez M., Ramos M. & Haenlein G.F.W. (2007) Physico-chemical characteristics of goat and sheep milk. *Small Ruminant Research* **68,** 88-113.

Picariello G., Rignanese D., Chessa S., Ceriotti G., Trani A., Caroli A. & Di Luccia A. (2009) Characterization and genetic study of the ovine α_{s2}-casein (*CSN1S2*) allele *B*. *Protein Journal* **28,** 333-40.

Pirisi A., Piredda G., Papoff C.M., Di Salvo R., Pintus S., Garro G., Ferranti P. & Chianese L. (1999) Effects of sheep α_{s1}-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. *Journal of Dairy Research* **66,** 409-19.

Prinzenberg E.M., Jianlin H. & Erhardt G. (2008) Genetic Variation in the κ-Casein Gene (*CSN3*) of Chinese Yak (Bos grunniens) and Phylogenetic Analysis of *CSN3* Sequences in the Genus Bos. *Journal of Dairy Science* **91,** 1198-203.

Recio I., Ramos M. & Amigo L. (1997) Study of the polymorphism of ovine α_{s1}- and α_{s2}-caseins by capillary electrophoresis. *Journal of Dairy Research* **64,** 525-34.

Shin H.J., Park H.Y., Jeong S.J., Park H.W., Kim Y.K., Cho S.H., Kim Y.Y., Cho M.L., Kim H.Y., Min K.U. & Lee C.W. (2005) STAT4 expression in human T cells is regulated by DNA methylation but not by promoter polymorphism. *Journal of Immunology* **175,** 7143-50.

Threadgill D.W. & Womack J.E. (1990) Genomic analysis of the major bovine milk protein genes. *Nucleic Acids Research* **18,** 6935-42.

Yeh C.Y., Yang R., Boyle T., Ye Z.-H. & Moa J. (1997) POPGENE-The user-friendly freeware for population genetic analysis. http://www.ualberta.ca/~fyeh/.

Supporting information

```
                                421                             449                 469      480
CSN1S2*A        (GU169092)      TGTTCTCCTTCTAGGATTGGAAAGATTTTTAGTTCTAAGTTAAAAAGCGTACACTTTCTT
CSN1S2*C and D (GU169094)      ****************************A******************A***********

                                481                 500                                 540
CSN1S2*A        (GU169092)      CCATAAACATTTGAATTTACTGGGGAAATATTTTCACTCCGAGAATCACTTCTTAAAAGT
CSN1S2*C and D (GU169094)      ******************T****************************************

                                541       550
CSN1S2*A        (GU169092)      AAGAAACATG----------------------------------------------------
CSN1S2*C and D (GU169094)      **********TAAGAATTAAAGATATTAAAATTTAAAATATTAAAAAAATTAAAAATAAA
                                          551                                             600

                                                                                551    559
CSN1S2*A        (GU169092)      ----------------------------------------------------GAACCATAC
CSN1S2*C and D (GU169094)      AAAAATAAATAAAAATTTTAAAAATGTAGGGGAAAAAAAAAAGAAACATG*********
                                601                                             650    659
```

Figure S1 Alignment of intron 2 sequence of *CSN1S2*A* and *CSN1S2*C/D* including the 100 bp insertion within alleles *C* and *D*. Detail of GenBank accession numbers GU169092 and GU169094.

Stars mark identical nucleotides. Dashes mark missing nucleotides in *CSN1S2*A*, not carrying the insertion. Numbering above the sequences corresponds to GenBank accession number GU169092 and italic numbering below the sequences corresponds to GenBank accession number GU169094. Single bold nucleotides mark three of five SNPs associated with the 100 bp insertion.

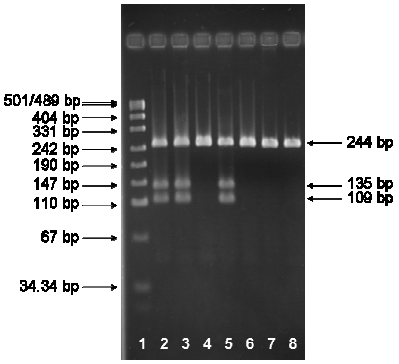


Figure S2 Agarose gel electrophoresis of P*si*I-digested PCR products for the detection of ovine *CSN1S2*C*.

Lane 1: marker; Lanes 2, 3, and 5: *CSN1S2*C* heterozygous; Lanes 4, 6, 7, and 8: *CSN1S2 non-C/non-C*.

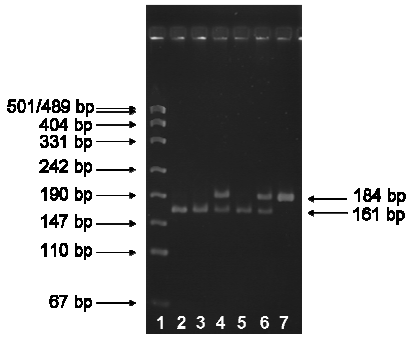


Figure S3 Agarose gel electrophoresis of *Hpy188*III-digested ACRS-PCR products for the detection of ovine *CSN1S2*D*.

Lane 1: marker; Lanes 2, 3, and 5: *CSN1S2 non-D/non-D*; Lanes 4 and 6: *CSN1S2*D* heterozyous; Lane 7: uncut PCR product or *CSN1S2 D/D*. The small 23 bp fragment resulting from the digestion in *CSN1S2 non-D* alleles is not visible.

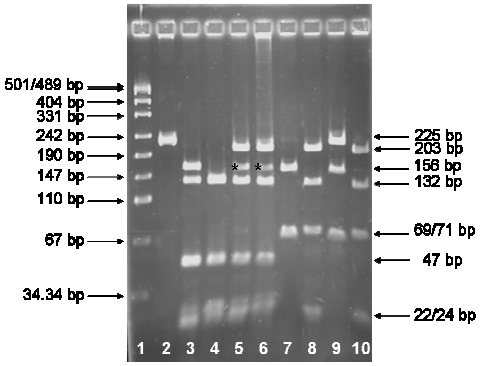


Figure S4 Agarose gel electrophoresis of *Rsa*I- and/or *Bme1390*I-digested PCR products for the detection of ovine *CSN1S2*B* and A_{322}>C_{322}-SNP.

Lane 1: marker; Lane 2: uncut PCR product; Lanes 3–6 *Rsa*I/*Bme1390*I-double digestion: Lane 3: *CSN1S2 BC*, Lane 4: *CSN1S2 BB*, Lanes 5 and 6: *CSN1S2 AB*, *BD*, or *BG*; Lanes 7 and 9 *Bme1390*I-RFLP-analysis: Lane 7: *CSN1S2* with C_{322}/C_{322} (= *BB*, *BC*, or *CC*); Lane 9: *CSN1S2* with A_{322}/C_{322} (= *AB*, *AC*, *BD*, *CD*, *BG*, *CG*); Lanes 8 and 10 *Rsa*I-RFLP-analysis: Lane 8 and 10: *CSN1S2 B/non-B*. *After digestion of samples with *CSN1S2 AB*, *BD*, or *BG* an additional fragment of 156 bp without diagnostic relevance occurs.

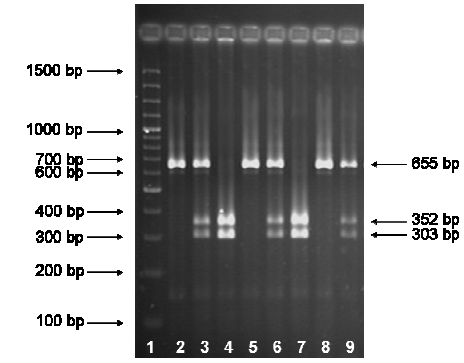


Figure S5 Agarose gel electrophoresis of *Hph*I-digested PCR products for the detection of ovine *CSN1S2*G*.

Lane 1: marker; Lane 2: uncut PCR product; Lanes 3, 6, and 9: *CSN1S2*G* heterozygous; Lanes 4 and 7: *CSN1S2 GG*; Lanes 5 and 8: *CSN1S2 non-G/non-G*.

6. Association study between milk protein variants and milk performance traits in East Friesian Dairy sheep

I. J. Giambra*, E. Stam[#], F. Atema[#], E. Schuiling[†], H. Brandt*, and G. Erhardt*

*Department of Animal Breeding and Genetics, Justus-Liebig-University, 35390 Gießen, Germany

[#]De Kooihoek Farm, 7245 VE Laren gld, The Netherlands

[†]Fokwaarde+, Op't Eekt, 8097 PC Oosterwolde, The Netherlands

Published in:

Proceedings of 9th World Congress on Genetics Applied to Livestock Production, 36 (2010)

Submitted: 19.02.2010

Accepted: 27.05.2010

Introduction

Detection and mapping of genes with an influence on economic important traits in dairy sheep is a prerequisite to improve dairy sheep production systems. In dairy cattle and goats genetic polymorphisms of milk proteins show associations to quantitative and qualitative parameters in milk (Martin *et al.*, 2002; Caroli *et al.*, 2009), while in sheep association studies are rare and the results are controversial (Amigo *et al.*, 2000; Barillet, 2007).

So far different alleles/phenotypes are identified for ovine α_{s1}-casein (α_{s1}-CN A-F, H, I) (Amigo *et al.*, 2000; Ceriotti *et al.*, 2005; Giambra *et al.*, 2010a, b), α_{s2}-casein (α_{s2}-CN A-D) (Picariello *et al.*, 2009; Giambra *et al.*, 2010c), β-lactoglobulin (β-LG A-C) (Bell and McKenzie, 1967; Erhardt, 1989a), and α-lactalbumin (α-LA A, B) (Schmidt and Ebner, 1972). Compared to bovine and caprine κ-casein (κ-CN) this fraction is monomorphic at protein level in sheep, while some single nucleotide polymorphisms without relation to protein variability were demonstrated (Ceriotti *et al.*, 2004; Feligini *et al.*, 2005).

Casein genes are organized as a tightly linked cluster on ovine chromosome 6 (Threadgill and Womack, 1990; Bevilacqua *et al.*, 2006) and therefore the estimation of associations between casein variants and milk production traits can be improved by considering the entire casein haplotype instead of the single genes (Sacchi *et al.*, 2005).

The aim of this study was to determine the allele frequencies for milk proteins by isoelectric focusing and to estimate genotype and haplotype effects on milk production traits within Dutch East Friesian Dairy sheep.

Material and methods

Sheep milk samples. Milk samples of 76 Dutch East Friesian Dairy sheep (OMS) were collected during routine milking.

Isoelectric focusing (IEF). Separation and identification of milk proteins were done by IEF of skimmed milk samples according to Erhardt (1989b) and Giambra *et al*. (2010c).

Milk performance traits. For all tested lactations (in detail 260 lactations) of the 76 sheep the Fokwaarde+ organisation provided milk recording data, in particular milk yield (milkkg240), fat and protein % with a mean of 479 kg milk, 5.49% fat and 5.14% protein in the standardised 240 days lactation.

Statistical analyses. Allele and genotype frequencies were calculated with program PopGene V 1.31 (Yeh *et al.*, 1997). A χ^2-test was performed to test the goodness of fit to Hardy-Weinberg equilibrium expectations for the distribution of genotypes.

Haplotype frequencies for α_{s1}- and α_{s2}-CN and the occurrence of linkage disequilibrium were estimated with EH software (Xie and Ott, 1993), considering only alleles with

frequencies > 0.05. For association studies data were analysed with the GLM procedure of SAS 9.1 (SAS Institute, Inc., Cary, NC, USA). Dependent variables in the analysis were milk, fat, and protein yield as well as fat and protein percent derived from routine evaluation. Beside the separately analysed milk protein genotypes the number of lactation (1, 2, 3, ≥4) and the current class of season (year and month of lambing) were included as fixed effects in the model.

Results and discussion

Isoelectric focusing. IEF lead to the identification of α_{s1}-CN C and H, α_{s2}-CN A, B, and C, as well as β-LG A and B, while κ-CN and α-LA were monomorph. Genotype and allele frequencies are shown in Table 1.

Table 1: Genotype and allele frequencies for α_{s1}-CN, α_{s2}-CN, α-LA, and β-LG in Dutch East Friesian Dairy sheep determined by IEF (n = 76)

	Genotype							Allele			
	AA	AB	AC	BB	BC	CC	CH	A	B	C	H
α_{s1}-CN	-	-	-	-	-	0.84	0.16	-	-	0.92	0.08
α_{s2}-CN	0.16	0.58	0.01	0.22	0.03	-	-	0.45	0.53	0.02	-
α-LA	1.00	-	-	-	-	-	-	1.00	-	-	-
β-LG	0.56	0.41	-	0.03	-	-	-	0.77	0.23	-	-

Milk protein allele frequencies are similar to those reported in Germany for OMS (Giambra *et al.*, 2010c). There was a good agreement between the observed genotype frequencies and those expected on the basis of Hardy-Weinberg equilibrium in all milk proteins.

Estimated α_{s1}-/α_{s2}-CN-haplotype frequencies are presented in Table 2, whereas χ^2-test showed a significant association (linkage disequilibrium) between α_{s1}-CN and α_{s2}-CN ($P \leq 0.05$), mainly caused by linkage between α_{s1}-CN H and α_{s2}-CN B.

Table 2: Frequency of α_{s1}-/α_{s2}-CN-haplotypes (n = 73) [a]

	Haplotype (α_{s1}-/α_{s2}-CN)			
Frequency	CA	CB	HA	HB
Expected[b]	0.43	0.49	0.04	0.04
Observed	0.46	0.46	0.00	0.08

[a] For haplotype evaluation, α_{s2}-CN C (allele frequency < 0.05) was not considered (n = 3).
[b] Expected haplotype frequency under independence hypothesis.

Association studies. For the association studies the genotypes α_{s2}-CN AC and BC, as well as β-LG BB were excluded (frequency < 0.10). Due to the linkage between α_{s1}-CN H and α_{s2}-CN B for animals genotyped as α_{s1}-CN CH and α_{s2}-CN AB the α_{s1}-/α_{s2}-CN-haplotype CA/HB was assumed. For all other animals haplotypes could definitely be reconstructed.

Highly significant effects of number of lactation on yield traits ($P \leq 0.01$) and of class of season on all traits ($P \leq 0.001$) were identified and confirm environmental influences on milk production traits in sheep (Ruiz *et al.*, 2000).

Genotype and haplotype effects of milk protein polymorphisms on milk production traits are demonstrated in Table 3. The results show a highly significant positive association between α_{s1}-CN CH and fat percent and yield, while in different Italian sheep breeds (Amigo *et al*., 2000) the predominant α_{s1}-CN CC was associated with higher fat, protein, and casein content. We assume that α_{s1}-CN H did not occur in these Italian sheep breeds, as they used comparable techniques and conditions. Missing of α_{s1}-CN alleles A and D in the typed OMS seems to be caused by indirect selection, due to the advantage of α_{s1}-CN C and H. Furthermore, α_{s1}-CN CH showed a tendency to higher milk yield also described by Wessels *et al*. (2004).

Within α_{s2}-CN we identified a significant positive association between α_{s2}-CN AB and protein yield, not described before. However, similar studies with East Friesian Dairy sheep kept on different farms in Germany found significant relationships between α_{s2}-CN genotypes and milk yield (Horstick *et al.*, 2002) or fat kg (Wessels *et al*., 2004).

The identified significant positive effects of β-LG AA on fat and protein % confirm results of Rampilli *et al.* (1997). On the other side the described positive effects of β-LG BB on milk yield (Amigo *et al.*, 2000) could not be approved as the frequency of this genotype was low within the population. As shown in Table 3 significant haplotype effects could be identified on fat %, protein %, fat and protein yield. In detail haplotype CB/HB showed association with highest fat and protein percent and yield respectively.

Table 3: Estimates of genotype and haplotype (α_{s1}-/α_{s2}-CN) effects on milk production traits (least square means and standard errors) in Dutch East Friesian Dairy sheep

Genotype/Haplotype		Trait milkkg240	fat %	protein %	fat kg	protein kg
α_{s1}-CN	Sign.[a]	n.s.	***	n.s.	**	n.s.
(260 lactations)	CC	499±5	5.50±0.03	5.17±0.02	27.4±0.3	25.8±0.3
	CH	516±14	5.83±0.09	5.16±0.06	30.0±0.8	26.6±0.7
α_{s2}-CN	Sign.	n.s.	n.s	n.s.	n.s.	**
(253 lactations)	AA	478±10	5.58±0.07	5.12±0.04	26.6±0.6	24.4±0.5
	AB	505±6	5.52±0.04	5.19±0.03	27.9±0.4	26.2±0.3
	BB	503±9	5.51±0.06	5.18±0.04	27.7±0.5	26.0±0.5
β-LG	Sign.	n.s.	**	**	n.s.	n.s.
(252 lactations)	AA	498±7	5.61±0.04	5.23±0.03	27.9±0.4	26.0±0.3
	AB	503±7	5.46±0.04	5.11±0.03	27.5±0.4	25.7±0.3
α_{s1}-/α_{s2}-CN	Sign.	n.s.	**	**	**	**
(253 lactations)	CA/CA	478±10	5.58±0.06	5.12±0.04	26.7±0.6	24.5±0.5
	CA/CB	505±6	5.51±0.04	5.21±0.03	27.8±0.4	26.3±0.3
	CA/HB	509±17	5.74±0.11	5.05±0.07	29.0±1.0	25.7±0.9
	CB/CB	499±10	5.41±0.06	5.14±0.04	27.0±0.6	25.6±0.5
	CB/HB	523±22	6.00±0.14	5.37±0.09	31.6±1.3	27.9±1.1

[a]Sign. = Significance level: n.s. = $P > 0.05$; ** = $P \leq 0.01$; *** = $P \leq 0.001$.

These are to our knowledge the first association studies including ovine casein haplotypes while in cattle numerous studies are already available (Caroli *et al.*, 2009).

The possibility to identify genotypes with α_{s1}-CN allele H by IEF (Giambra *et al.*, 2010c) and by DNA-tests (Giambra *et al.*, 2010b) opens the chance to screen further dairy sheep breeds for the occurrence of this allele and to confirm its positive effects. As the caseins are organized within a cluster (Threadgill and Womack, 1990; Bevilacqua *et al.*, 2006) these effects should be estimated using casein haplotypes or composite casein genotypes instead of separate genotypes. In the case of consistent results this will open the possibility to consider the casein locus as marker in breeding programs like in goat (Sanchez *et al.*, 2005).

Conclusion

The observation of significant effects of casein genotypes and haplotypes in Dutch East Friesian Dairy sheep should be confirmed in other dairy breeds. The haplotype that improved fat and protein percentage was also associated with positive effects on yield of fat and protein. If the observed effects are real, the casein complex offers the use in marker assisted selection in dairy sheep.

References

Amigo, L., Recio, I., and Ramos, M. (2000). *Int. Dairy J.,* 10:135-149.

Barillet, F. (2007). *Small Ruminant Res.,* 70:60-75.

Bell, K., and Mckenzie, H.A. (1967). *Biochim. Biophys. Acta,* 147:123-134.

Bevilacqua, C., Helbling, J.C., Miranda, G.*, et al.* (2006). *Reprod. Nutr. Dev.,* 46:567-578.

Caroli, A.M., Chessa, S., and Erhardt, G.J. (2009). *J. Dairy Sci.,* 92:5335-5352.

Ceriotti, G., Chessa, S., Bolla, P.*, et al.* (2004). *J. Dairy Sci.,* 87:2606-2613.

Ceriotti, G., Chiatti, F., Bolla, R.*, et al.* (2005). *Ital. J. Anim. Sci.,* 4:64-66.

Erhardt, G. (1989a). *Anim. Genet.,* 20:197-204.

Erhardt, G. (1989b). *J. Anim. Breed. Genet.,* 106:225-231.

Feligini, M., Vlaco, S., Curik, V.C.*, et al.* (2005). *J. Dairy Res.,* 72:317-321.

Giambra, I.J., Chianese, L., Ferranti, P.*, et al.* (2010a). *Int. Dairy J.,* 20:195-202.

Giambra, I.J., Chianese, L., Ferranti, P.*, et al.* (2010b). *J. Dairy Sci.,* 93:792-795.

Giambra, I.J., Jäger, S., and Erhardt, G. (2010c). *Small Ruminant Res.,* in press.

Horstick, A., Hamann, H., Erhardt, G.*, et al.* (2002). *Züchtungskunde,* 74:135-143.

Martin, P., Szymanowska, M., Zwierzchowski, L.*, et al.* (2002). *Reprod. Nutr. Dev.,* 42:433-459.

Picariello, G., Rignanese, D., Chessa, S.*, et al.* (2009). *Protein J.,* 28:333-340.

Rampilli, M., Cecchi, F., Giuliotti, L.*, et al.* (1997). In *Proc. IDF Seminar Milk Protein Polymorphism*, volume 2, pages 311-315.

Ruiz, R., Oregui, L.M., and Herrero, M. (2000). *J. Dairy Sci.,* 83:2709-2719.

Sacchi, P., Chessa, S., Budelli, E.*, et al.* (2005). *J. Dairy Sci.,* 88:1561-1568.

Sanchez, A., Ilahi, H., Manfredi, E.*, et al.* (2005). *J. Anim. Breed. Genet.,* 122:21-29.

Schmidt, D.V., and Ebner, K.E. (1972). *Biochim. Biophys. Acta,* 263:714-720.

Threadgill, D.W., and Womack, J.E. (1990). *Nucleic Acids Res.,* 18:6935-6942.

Wessels, G., Hamann, H., Erhardt, G.*, et al.* (2004). *Berl. Münch. Tierärztl. Wochenschr.,* 117:414-419.

Xie, X., and Ott, J. (1993). *Am. J. Hum. Genet.,* 53 (Suppl.):1107.

Yeh, C.Y., Yang, R., Boyle, T.*, et al.* (1997). http://www.ualberta.ca/~fyeh/.

7. General discussion of the results

This study focused on the analysis of milk proteins in sheep on a biochemical and molecular genetic way to get a more complete picture about ovine milk protein variability, to uniform nomenclature, to develop DNA-based tests, and to start association studies with milk performance traits.

7.1. Protein, mRNA and DNA analyses

The combination of screening milk protein fractions by IEF, followed by further protein analyses such as polyacrylamide gel electrophoresis, immunoblotting, and mass spectrometry in the case of α_{s1}-CN, and molecular genetic mRNA and DNA sequence studies lead to the detection of various polymorphisms within ovine α_{s1}- and α_{s2}-CN genes and proteins.

IEF is a common method for screening milk protein variability within the six main ruminants' milk proteins and helped to detect new milk protein variants within cattle and goat in the past (Erhardt, 1993, 1996; Erhardt *et al.*, 1998; Erhardt *et al.*, 2002). In this study IEF was used for typing α_{s1}-, α_{s2}-, and κ-CN as well as α-LA and β-LG in several sheep breeds of different purposes. We confirmed α_{s1}-CN alleles A, C, D, and X, which we named H, and identified a new variant named I. In accordance with Chessa *et al.* (2003) we could not differentiate between α_{s1}-CN B and C. Therefore, it is necessary to clarify if α_{s1}-CN B, described by Chianese *et al.* (1996), is really an existing *CSN1S1* variant. The α_{s1}-CN alleles E and F (Chianese *et al.*, 1996; Pirisi *et al.*, 1999) were not found, possibly due to breed specific occurrence of these protein variants in Italian sheep breeds, but were considered throughout the identification and nomination of new variants. Within α_{s2}-CN alleles A and B, named by Chessa *et al.* (2003), were demonstrated in all breeds included in this study. Additionally, two new pattern forms, α_{s2}-CN C and D, were described.

Allele frequencies of the new α_{s1}-CN variants H and I and α_{s2}-CN C and D are generally low but they occurred in common and not only in rare or special breeds. In detail, α_{s1}-CN H seems to be characteristically for East Friesian Dairy sheep, an important dairy breed all over the world (Haenlein and Wendorff, 2006), whereas α_{s1}-CN I was identified only in Gray Horned Heath, mainly used for preservation of the heath in specific areas of Germany. Wessels *et al.* (2004) were the first who described α_{s1}-CN H, which was also detected in our study in several flocks of East Friesian Dairy sheep held in Germany and the Netherlands. Indirect selection over the last years did not cause a loss or an increase of α_{s1}-CN H. Therefore, this allele seems to be manifested in East Friesian Dairy sheep.

The α_{s2}-CN allele C is more common and appeared in four out of six sheep breeds of different purposes. On the other hand α_{s2}-CN D occurred only in Merinoland sheep, especially used for lamb and meat production. Also ovine *LGB*C*, bovine *CSN1S1*F* and *CSN3*G1* (Caroli *et al.*, 2009) or caprine *CSN1S1*H* and *CSN2*A1* (Sacchi *et al.*, 2005; Caroli *et al.*, 2007b), for example, are actually detected only in special breeds and can be used to identify relationships between different breeds in the particular species and to reconstruct evolutionary ways like in Pinzgauer cattle (Caroli *et al.*, 2010).

The fraction of β-CN was illustrated and clearly separated by IEF but more complex techniques such as mass spectrometries are necessary to define protein variability. Therefore, to date, no screening system is available for typing ovine β-CN and *CSN2* was not included in genetic analyses. Further protein analyses and additional molecular genetic approach should be made to get a more complete overview of ovine β-CN variability like in cattle and goat (Caroli *et al.*, 2006; Caroli *et al.*, 2009; Küpper *et al.*, 2010).

Within ovine κ-CN no variation could be demonstrated, being in accordance with further studies (Moioli *et al.*, 1998; Chessa *et al.*, 2003). Interspecies comparisons demonstrated that *CSN3* possesses the greatest degree of conservation among the casein genes, which might be related to its essential function as a casein micelle stabiliser (Alexander *et al.*, 1988). The great importance of sheep milk for ages to make cheese may have played a crucial role in the ovine *CSN3* molecular evolution (Ceriotti *et al.*, 2004) and may also be a reason for its highly fixed amino acid structure and the low polymorphic degree. Perhaps there has been an indirect selection towards the prevalent *CSN3* form leading to good cheese making abilities and/or advantageous utilisation of nutrients for lambs in non-dairy breeds. As discussed for bovine species, strong positive selection results in low levels of polymorphism (Ward *et al.*, 1997).

IEF-analyses of ovine whey proteins revealed monomorphism within ovine α-LA and illustrated all currently known β-LG variants within the sheep breeds analysed. Inclusion of breeds within this study not analysed before legitimates the conclusion that probably the main protein forms of these milk proteins are known. Low polymorphic degree of α-LA in sheep (Schmidt and Ebner, 1972; Erhardt, 1989a) is comparable with cattle (Farrell *et al.*, 2004) and goat (Cosenza *et al.*, 2003) and is probably caused by its important function in lactose synthetase complex (Brew and Grobler, 1992; Martin *et al.*, 2002). Constant monitoring of milk protein variation in different cattle breeds was recommended as an essential practice aiming to avoid an increase of unfavourable alleles (Caroli *et al.*, 2009). Therefore, including the whey proteins in future studies, concerning ovine milk protein variability and their associations to production traits seems to be advisable to detect emerging variants.

Furthermore, mRNA- and DNA-based analyses should be promoted in these ovine whey proteins, especially in ovine α-LA.

In addition, to our knowledge, the first ovine casein haplotype studies were made within this IEF-study and identified seven α_{s1}-/α_{s2}-CN-haplotypes while haplotypes CA and CB were most frequent. This is explainable by the high frequencies of the corresponding casein alleles, being in accordance with past studies (Amigo *et al.*, 2000; Chessa *et al.*, 2003). The identification of the new casein alleles lead to the occurrence of new haplotypes like α_{s1}-/α_{s2}-CN-haplotypes CC and HB and supplemented the possible degree of allele and genotype variability. Casein haplotype studies gained importance in cattle and goat milk protein studies over the last years (Caroli *et al.*, 2006; Nilsen *et al.*, 2009), due to the tight linkage of the casein genes and the resulting linked inheritance (Threadgill and Womack, 1990; Bevilacqua *et al.*, 2006). This study revealed significant linkage between *CSN1S1* and *CSN1S2* and was a first step to reach the level of knowledge already achieved in goat for example (Caroli *et al.*, 2006; Küpper *et al.*, 2010). Construction and usage of complete casein haplotypes require the knowledge of variability of all casein loci, whereas also in cattle and goat fragmentary casein haplotypes or composite genotypes (Sacchi *et al.*, 2005; Heck *et al.*, 2009; Bonfatti *et al.*, 2010) are analysed. This is mainly caused by the occurrence of individual monomorphic caseins or studies restricted to certain caseins. Inclusion of ovine *CSN2* and *CSN3* in future research should be made, if variability can be detected within these caseins on protein or DNA level, to complete casein haplotype studies.

Further protein analyses as well as mRNA/cDNA sequence studies detected causal mechanisms for the becoming of ovine protein patterns α_{s1}-CN H and I. Mass spectrometry of purified, tryptic digested, and dephosphorylated α_{s1}-CN proteins identified lower mass values of α_{s1}-CN H and I in comparison to C''. These lower masses are due to the missing of amino acid residues caused by exon skipping in both cases and identified by mRNA sequencing. Both methods are therefore able to identify these modifications in the transcription process, whereas most approaches are based on protein studies (Ferranti *et al.*, 1998; Ferranti *et al.*, 2001), and mRNA studies are rarer (Passey *et al.*, 1996). In addition, protein analyses and mRNA sequencing verified the occurrence of non-allelic protein forms within ovine α_{s1}-CN in all analysed protein variants, confirming their wide distribution within this ovine casein (Ferranti *et al.*, 1998; Ferranti *et al.*, 2001), not only in α_{s1}-CN A, C, and D.

The mRNA, used for the sequence analyses within this study, was extracted from milk somatic cells as described by Boutinaud *et al.* (2002). For this reason no udder tissue samples were required and the necessity of slaughtering animals was avoided. Sequence studies of the

reverse transcribed mRNA (cDNA) made it possible to get the complete coding sequence of *CSN1S1* and *CSN1S2*, both characterised by small exons and large non-coding introns (Rijnkels, 2002), in one step. This is easier compared to the usage of DNA and the amplification of various single PCRs for the individual exons (Lühken *et al.*, 2009).

The results of this study show the necessity to analyse the complete pathway between DNA and protein including transcription and translation processes, as they cause high degrees of variability in ovine α_{s1}-CN. DNA sequence studies were necessary to explain mRNA and resulting protein variability, to identify the reasons of alternative splicing in *CSN1S1*H* and *CSN1S1*I*. For this the collected milk samples were partly also used as a DNA source (Lühken *et al.*, 2009), although DNA yield was often lower in comparison with extractions from blood, as described by Chessa *et al.* (2007) before. However, DNA extraction using milk, for example collected during sampling for milk performance testing, as starting material can be an alternative for typing great numbers of lactating ewes, as it is more comfortable in comparison to collect further blood samples. Alternatively milk can directly be used as template in PCR reactions as described by Chessa *et al.* (2010), allowing to skip the DNA extraction step.

The sequencing of DNA identified two kinds of sequence differences, causing the disruption of splice sites and with this exon skipping during the serial splice reactions. Within *CSN1S1*H* a deletion combined with an insertion and in *CSN1S1*I* a single SNP interrupted the complementarity of intron splice donor signals with the U1 small nuclear RNA. In bovine and caprine casein variants exon skipping is generated by SNPs within the following introns (Leroux *et al.*, 1992; Mohr *et al.*, 1994; Grosclaude and Martin, 1997). The identified sequence differences of ovine *CSN1S1*H* therefore seem to be a new reason for alternative splicing in ruminant milk proteins. Exon skipping resulted in both proteins in a lack of eight amino acids, of which three are with negative net charge (2 Asp and 1 Glu in α_{s1}-CN H and 1 Asp and 2 Glu in α_{s1}-CN I), explaining the more alkaline isoelectric point of both protein variants in comparison to α_{s1}-CN C. Due to this, and due to the loss of phosphorylated serin residues in α_{s1}-CN I, the electrophoretic behaviour in IEF is clarified.

Additionally, DNA studies identified further single nucleotide polymorphisms within intron parts of ovine *CSN1S1* not described before, demonstrating that there is still a great potential to identify further variability in future studies.

Also in *CSN1S2* molecular genetic mRNA and DNA sequence studies of alleles *C* and *D* revealed SNPs and insertions of different lengths, partly leading to the further variants *CSN1S2*G* and *A'*, which were not known before and not identifiable by IEF on protein level.

Especially *CSN1S2*G* occurred often, with frequencies of up to 0.383. Therefore, *CSN1S2* also seems to be characterised by a higher degree of polymorphism as thought before, comparable with that of caprine (Ramunno *et al.*, 2001; Sacchi *et al.*, 2005) and higher than in bovine *CSN1S2* (Farrell *et al.*, 2004; Caroli *et al.*, 2009).

The information gained about *CSN1S1* and *CSN1S2* variants were used to establish DNA-based tests to type non-lactating animals, too. Beside IEF a DNA-based test is necessary, for example, to detect *CSN1S2*G* not distinguishable by IEF from *CSN1S2*A*. These tests were based on PCR-RFLP (*CSN1S1*H* and *CSN1S2*B*, *C*, and *G*), ACRS-PCR-RFLP (*CSN1S1*I* and *CSN1S2*D*) or fragment length analysis (*CSN1S1*H*) and are only useable for typing occurrence or non-occurrence of the respective casein allele, like the tests described in literature until now for sheep and other ruminants (Ramunno *et al.*, 1997; Pilla *et al.*, 1998; Prinzenberg *et al.*, 1998; Ceriotti *et al.*, 2004). These DNA-based tests can replace expensive and time-consuming sequencing methods but a next step should be the usage of SNP chip technology for typing ovine milk protein variability.

The commercial Illumina OvineSNP50 Bead Chip contains only few SNPs of milk protein genes, not representing the variants and placed within the intron sequences, whereas for *CSN3* and *LGB* no SNPs are included (http://www.livestockgenomics.csiro.au/; https://isgcdata.agresearch.co.nz/). Therefore, usage of this technology for typing sheep milk protein polymorphisms is unsuitable to date and the development of a custom-made SNP array, containing all known ovine milk protein polymorphisms, can be reasonable for the simultaneous and fast typing of a great number of mutations affecting milk protein structure and composition as in cattle (Kaminski *et al.*, 2005; Chessa *et al.*, 2007). In this SNP array complete knowledge about actually described variability within milk proteins, also the small number of known *CSN2* and *CSN3* polymorphisms, should be included. However, the described combination of proteomic and genomic methods should be used in future studies to screen sheep breeds for milk protein variability as DNA-based tests and the chip-technology are not able to detect new variations, only to screen for known polymorphisms.

Moreover, sequence studies, established within this study, can help to complete the whole genome sequence of sheep, which is under progress (https://isgcdata.agresearch.co.nz/), and where casein gene sequences still show big gaps.

7.2. Nomenclature and evolutionary pathways

After definite biochemical and molecular genetic characterisation of ovine *CSN1S1* and *CSN1S2* we proposed to uniform the nomenclature of the different genetic variants, to clarify

their distribution in different sheep breeds. Consequentely the phylogenetic trees of the ovine caseins, as published by Ceriotti *et al.* (2005) for *CSN1S1* for the first time, can be supplemented (Figure 1). *CSN1S1*C'* was declared as the ancestral α_{s1}-CN form due to sequence agreement to cattle and goat (Ceriotti *et al.*, 2005). *CSN1S1*C''* is characterised by a sequence difference concerning amino acid 186 of mature protein shortened by exon 16 or amino acid 194 in the complete sequence of the mature protein, investigated within this study (GenBank Acc. No. FJ440845). *CSN1S1*H* and *I* have both evolved from the most common *CSN1S1*C''*, because their proteins carry Ile_{194}, like those of *CSN1S1*A* and *D*. Therefore, differentiation between *CSN1S1*C'* and *C''* should have happened earlier than the development of *CSN1S1*H* and *I* (Figure 1). *CSN1S1*E* is known to be caused by a deletion of the amino acid residues 70 to 77 (Chianese *et al.*, 2007), but the residual protein sequence is not clarified to date and therefore this allele is not implemented in Figure 1. Further studies, including wild forms of ovine species, should be accomplished to prove or disprove the proposed evolution of *CSN1S1*. Concerning the complete described protein variability of α_{s1}-CN protein and DNA (see 1.3.2.1.), sequence clarification is needed to implement, for example, α_{s1}-CN patterns B and F (Chianese *et al.*, 1996; Pirisi *et al.*, 1999) in this pathway. Comparison tests, to face the occurrence of these alleles and to avoid confusion within nomenclature, are necessary like already done for goat by Prinzenberg *et al.* (2005).

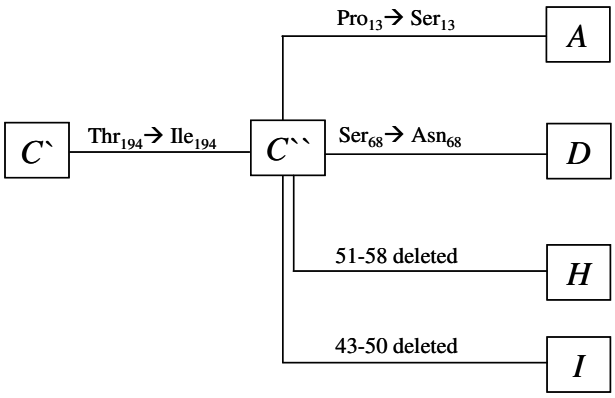


Figure 1. Possible evolutionary pathway of ovine *CSN1S1*, with the characteristic amino acid exchanges or deletions of its variants

Also for ovine *CSN1S2* we generated an evolutionary pathway, which still seems to be fragmentary, due to two amino acid substitutions discriminating *CSN1S2*A'* and *CSN1S2*B* and *C*, respectively. In the past there was a comparable situation in bovine *CSN3*, but a precursor variant could be identified in *Bos grunniens* recently (Prinzenberg *et al.*, 2008). Probably sub forms are/were subsisting in ovine *CSN1S2*A'*, whereby the detection of new

variants and the classification in the phylogenetic pathway can allow a better understanding of evolutionary chains.

The resulting phylogenetic trees of ovine milk proteins can help to understand the history of sheep breeds and their genetic distances and to complete microsatellite and mtDNA-studies (Bruford *et al.*, 2003; Peter *et al.*, 2007). Microsatellite analysis of several European and Middle-Eastern sheep breeds, for example, revealed a greater genetic distance of Gray Horned Heath to 56 other breeds (Peter *et al.*, 2007), which is underlined by the apparently private *CSN1S1* allele *I* identified in this study or the non-occurrence of the *CSN3* T_{237}>C_{237}-SNP (Hartmann, 2010).

7.3. Association studies

Associations between milk protein polymorphisms and milk performance traits are widespreadly known for bovine and caprine species. Therefore, economic gain can be coupled with introduction of milk protein variants and casein haplotypes in animal breeding, for example by influencing cheese making properties or by opening the potential of producing functional food (Caroli *et al.*, 2009). Comparable analyses in sheep are rare, mostly preliminary, and reveal partly contradictory results (Amigo *et al.*, 2000). Inconsistencies in ovine association studies can be explained by breed differences, population size, frequency distribution of the genetic variants, and a failure to consider relationships among animals (Sabour *et al.*, 1996; Giaccone *et al.*, 2000), as well as by single gene analyses and non-inclusion of casein haplotypes, as is necessary due to the linkage of the casein genes (Threadgill and Womack, 1990; Boettcher *et al.*, 2004).

The association study in East Friesian Dairy sheep described above, considering milk performance traits and milk protein polymorphisms, typed by IEF, revealed significant influences of α_{s1}-CN genotype CH on fat content and yield, of α_{s2}-CN AB on protein yield and of β-LG AA on fat and protein content. Additionally, the fractional α_{s1}-/α_{s2}-CN-haplotype HB showed a significant positive association to all milk composition traits. Therefore, *CSN1S1*H* seems to be a favourable allele for cheese making for example, demonstrating also a positive effect in combination with α_{s2}-CN alleles as haplotype.

Protein analysis within this study revealed for the shortened α_{s1}-CN variants H and I, both characterised by missing of eight amino acids, a 74% lower and an identical expression level of *CSN1S1*, repectively, in comparison to *CSN1S1*C''*. Within goat for example, allelic forms of α_{s1}-CN are directly associated with different rates of protein synthesis: the strong alleles are associated with higher milk casein levels, while the weak and null alleles result in

decreased or missing casein contents (Caroli *et al.*, 2006). Abbreviated casein forms like caprine α_{s1}-CN variants F and G are, for example, weak alleles (Leroux *et al.*, 1992; Grosclaude and Martin, 1997). In sheep we could not identify an association between *CSN1S1*H* and lower milk protein content, as described before by Wessels *et al.* (2004), who identified a significantly lower protein content in milk of East Friesian Dairy sheep with α_{s1}-CN genotype CX (= CH, new nomenclature). Otherwise our results can be compared with bovine *CSN1S1*G*: cows carrying this allele produce less α_{s1}-CN but more of the other caseins (Mariani *et al.*, 1995; Farrell *et al.*, 2004). Therefore, whole protein content is not affected due to compensation by higher α_{s2}-, β- and/or κ-CN percentages. To confirm these first results further research is needed within East Friesian Dairy sheep and other important dairy sheep breeds, like Assaf or Lacaune, which are as yet not widespreadly analysed for milk protein variability.

Due to the high importance of cheese production in dairy sheep industry, clotting time and cheese making aptitudes of milk with the new casein variants, as well as with the already known milk protein polymorphisms, should be analysed. Known negative influences of *CSN1S1*D* on these production traits can be a reason for the absence of this allele in German as well as in Dutch East Friesian Dairy sheep in this and an earlier study (Wessels *et al.*, 2004), due to indirect selection.

The influence of *CSN1S1*I*, *CSN1S2*C*, *D*, and *G* on milk production traits should be analysed in future studies, by including animals with these variants. *CSN1S1*I* did not occur within the analysed East Friesian Dairy sheep and allele and genotype frequencies of the new *CSN1S2* alleles were too low for inclusion in association studies. The absence of phosphorylated serin and acidic residues in α_{s1}-CN I or the phosphorylation degree of serin residues within α_{s2}-CN C and D can influence micellar size and curd firming rate like in goat (Tziboula and Horne, 1999), for example, and can therefore show economic relevance.

7.4. Future usage of ovine milk protein polymorphisms

Results of the association studies can be included in ovine breeding programmes to enhance economy of dairy sheep husbandry and sheep milk production all over the world, whereas negative correlations between milk yield and content traits should be considered. When selecting animals for higher milk yield it is important that protein and fat contents do not decrease, especially in ovine breeds whose milk is often transformed into cheese (Ramos *et al.*, 2009). Casein and whey protein polymorphisms can be implemented in genomic selection. As defined by Árnyasi *et al.* (2009) the main objective of their analyses of

quantitative trait loci on OAR6, the chromosome, where the casein genes are lying on (http://www.animalgenome.org/sheep/maps/; de Gortari *et al.*, 1998), was to find and later use molecular genetic markers in selection to improve milk production and composition. In this context a main problem is the inadequate milk recording data base in many European countries without an explicit breeding programme for dairy sheep like in Germany, due to the high costs (Sanna *et al.*, 1997; Astruc *et al.*, 2004). This should be ameliorated to ensure success of implementation of milk proteins in breeding programmes in sheep.

Additionally, typing of local breeds in the least developed countries offers the possibility of increasing productivity also for these sheep in extensive production systems and to ameliorate the supply of the population with protein, for example, to advance their living conditions relying on protein-supply by ovine milk (Haenlein, 2001a; Degen, 2007). It is possible that association studies can generate results opening potential for new purposes of non-dairy breeds and also of sheep breeds which are endangered and/or which carry special milk protein variants.

Moreover, the production of functional food can be profitable for ovine dairy industry and breeding, by generating products with bioactive effect or products with lower allergenic potential. Proteolysis during digestion or food processing of milk proteins may release bioactive or allergenic peptides (Meisel, 1998; Meisel and Bockelmann, 1999; Busse *et al.*, 2002). Functional food containing bioactive peptides derived from bovine milk is commercially available nowadays (Phelan *et al.*, 2009). The appearance, the becoming or the elimination of bioactive peptides can be affected by amino acid exchanges or deletions resulting from milk protein polymorphisms as demonstrated for bovine κ-casein (Weimann *et al.*, 2009), due to the influence on the susceptibility to pepsin hydrolysis. The ovine β-LG variant B, for example, showed a higher pepsin digestion rate than variant A, due to differences in the tertiary structure (El-Zahar *et al.*, 2005).

This study is a first step to allow for such future studies, to open new tools for dairy sheep industry worldwide.

8. Summary

Isoelectric focusing was applied for screening milk protein variants in milk samples from 1,078 animals of six different sheep breeds (Black Faced Mutton sheep, East Friesian Dairy sheep, Gray Horned Heath, Merinoland sheep, Merino Mutton sheep, and Rhön sheep). The known genetic variants of α_{s1}-casein (CN; *CSN1S1*A*, *C*, *D*), α_{s2}-CN (*CSN1S2*A*, *B*), and β-lactoglobulin (LG; *LGB*A*, *B*, *C*) were confirmed. *CSN1S1*C* was predominant in all breeds with frequencies of 0.90 to 1.00. *CSN1S2*A* accumulated in Gray Horned Heath, Merinoland, Merino Mutton, and Rhön sheep (0.73 to 0.93), whereas *CSN1S2*B* was most frequent in Black Faced Mutton and East Friesian Dairy sheep. Within β-LG *LGB*A* was predominant in all breeds with frequencies of 0.54 to 0.80, whereas *LGB*C* showed breed specific distribution within the Merino breeds.

Furthermore, the patterns of α_{s1}-CN X were also detected in East Friesian Dairy sheep and named α_{s1}-CN H. One and two additional patterns were identified within α_{s1}-CN (I) and α_{s2}-CN (C and D), respectively. All animals analysed were monomorph for κ-casein and α-lactalbumin.

In addition, isoelectric focusing identified seven α_{s1}-/α_{s2}-CN-haplotypes within the breeds analysed, while haplotypes CA and CB were most frequent.

Continuative biochemical and molecular genetic analyses of *CSN1S1* confirmed the occurrence of variants *H* and *I*, both caused by alternative splicing events. Within *CSN1S1*H* skipping of exon 8 is caused by a deletion and an insertion within DNA sequence of exon 8 (g.739_742delAAGGinsTTATTTTAATAAA). Additionally, three SNPs both in intron 6 and in intron 7 were identified by DNA sequence analyses, whereas one of the latter ones (g.656T>A) caused an alternative splicing event in *CSN1S1*I* by disrupting the donor splice site of intron 7.

*CSN1S2*C* and *D* are both due to different single nucleotide polymorphisms within exon 7, leading to amino acid exchanges p.Val45Ile and p.Ala48Ser in α_{s2}-CN C and p.Arg46Ser in α_{s2}-CN D, associated with a SNP in exon 2 in both variants. Molecular genetic analyses of *CSN1S2* revealed a further variant *G*, caused by a non-synonymous SNP in exon 15 (p.Arg161His), and a sub form of *CSN1S2*A*, namely *A'*, due to a synonymous A>C mutation in exon 10. Additionally, one SNP both in exon 17 and exon 18, ten SNPs in total in introns 1, 2, 6, and 7, as well as three intron insertions were identified.

The biochemical and molecular genetic results were used to postulate nomenclature for ovine *CSN1S1* and *CSN1S2* variation. Denomination of the new variants was carried out

considering known polymorphisms and their names. In *CSN1S2* we defined also already described but not named variants with *CSN1S2*E* and *F*.

DNA tests were established for *CSN1S1* variants *H* and *I*, as well as for *CSN1S2*B*, *C*, *D*, and *G*, allowing to type sheep independent of age, lactation, and sex.

Additionally, association studies were established between milk protein variants and milk production traits in East Friesian Dairy sheep concerning milk protein genotypes and casein haplotypes. Significant positive associations between α_{s1}-CN CH and fat percentage and yield, between α_{s2}-CN AB and protein yield, and between β-LG AA and fat and protein percentage of sheep milk were identified. Furthermore, α_{s1}-/α_{s2}-CN-haplotype HB was associated with the highest fat and protein percentage and yield, respectively.

9. Zusammenfassung

Die isoelektrische Fokussierung wurde genutzt, um Milchproteinvarianten in Milchproben von 1078 Tieren sechs verschiedener Schafrassen (Schwarzköpfiges Fleischschaf, Ostfriesisches Milchschaf, Graue Gehörnte Heidschnucke, Merinolandschaf, Merinofleischschaf und Rhönschaf) zu bestimmen. Die bekannten genetischen Varianten im α_{s1}-Kasein (CN; *CSN1S1*A*, *C*, *D*), α_{s2}-CN (*CSN1S2*A*, *B*) und im β-Laktoglobulin (LG; *LGB*A*, *B*, *C*) wurden bestätigt. *CSN1S1*C* war in allen Rassen die häufigste Variante mit Allelfrequenzen von 0,90 bis 1,00. *CSN1S2*A* trat gehäuft in den Rassen Graue Gehörnte Heidschnucke, Merinoland-, Merinofleisch- und Rhönschaf auf (0,73 bis 0,93), wobei *CSN1S2*B* in den Schwarzköpfigen Fleischschafen und den Ostfriesischen Milchschafen überwog. Innerhalb des β-LG war *LGB*A* mit Frequenzen von 0,54 bis 0,80 in allen Rassen vorherrschend und *LGB*C* zeigte ein rassespezifisches Vorkommen in den Merinorassen.

Außerdem wurden in Ostfriesischen Milchschafen die Banden des α_{s1}-CN X nachgewiesen und als α_{s1}-CN H bezeichnet. Ein beziehungsweise zwei neue elektrophoretische Bandenmuster wurden im α_{s1}-CN (I) und im α_{s2}-CN (C und D) identifiziert. Alle untersuchten Tiere waren monomorph für κ-Kasein und α-Laktalbumin.

Des Weiteren führte die isoelektrische Fokussierung zur Identifizierung von sieben α_{s1}-/α_{s2}-CN-Haplotypen innerhalb des beschriebenen Tiermaterials, wobei die Haplotypen CA und CB vorherrschend waren.

Weiterführende biochemische und molekulargenetische Analysen des *CSN1S1* bestätigten das Vorkommen der Varianten *H* und *I*, die beide durch alternatives Spleißen bedingt sind. Innerhalb des *CSN1S1*H* ist das Skipping des Exons 8 durch eine Deletion und eine Insertion innerhalb der DNA-Sequenz des Exon 8 begründet (g.739_742delAAGGinsTTATTTTAATAAA). Zusätzlich wurden durch DNA-Sequenzanalysen je drei SNPs im Intron 6 und im Intron 7 detektiert, wobei einer der letztgenannten (g.656T>A) das alternative Spleißen im *CSN1S2*I* begründet, indem er die Donor-Spleißseite des Intron 7 zerstört.

*CSN1S2*C* und *D* sind beide durch unterschiedliche Einzelbasenaustausche innerhalb des Exon 7, die zu den Aminosäureaustauschen p.Val45Ile und p.Ala48Ser im α_{s2}-CN C und p.Arg46Ser im α_{s2}-CN D führen, gekennzeichnet und die in beiden Varianten mit einem SNP im Exon 2 assoziiert sind. Molekulargenetische Analysen des *CSN1S2* identifizierten zusätzlich eine weitere Variante *G*, die durch einen nicht-synonymen SNP im Exon 15 (p.Arg161His) bedingt ist, und eine als *CSN1S2*A'* bezeichnete Unterform des *CSN1S2*A*, die durch eine synonyme A>C Mutation im Exon 10 begründet ist. Weiterhin wurden in den

Exons 17 und 18 jeweils ein SNP und in den Introns 1, 2, 6 und 7 insgesamt zehn SNPs, sowie drei Intron-Insertionen gefunden.

Die biochemischen und molekulargenetischen Ergebnisse wurden genutzt, um die Nomenklatur für die ovinen *CSN1S1* und *CSN1S2* Varianten zu definieren. Die Bezeichnung der neuen Varianten erfolgte unter Berücksichtigung bekannter Polymorphismen und deren Namen. Im *CSN1S2* definierten wir außerdem bereits beschriebene aber nicht benannte Varianten mit *CSN1S2*E* und *F*.

Unter Verwendung der identifizierten DNA Polymorphismen wurden sowohl für die *CSN1S1* Varianten *H* und *I,* als auch für *CSN1S2*B*, *C*, *D* und *G* DNA-Tests entwickelt, die eine Typisierung der Schafe, unabhängig von Alter, Laktation und Geschlecht ermöglichen.

Zusätzlich wurden in Ostfriesischen Milchschafen unter Einbeziehung von Milchproteingenotypen und Kasein-Haplotypen Assoziationsstudien zwischen Milchproteinpolymorphismen und Produktionsmerkmalen durchgeführt. Es wurden signifikant positive Assoziationen zwischen dem α_{s1}-CN Genotyp CH und Fettgehalt und -menge, sowie zwischen α_{s2}-CN AB und Proteinmenge und zwischen β-LG AA und Fett- und Proteingehalt der Schafmilch identifiziert. Zudem war der α_{s1}-/α_{s2}-CN-Haplotyp HB mit dem höchsten Fett- und Proteingehalt sowie mit der höchsten Fett- und Proteinmenge assoziiert.

10. References of the general parts

Addeo F., Moio L., Chianese L., Stingo C., Resini P., Berner I., Krause I., Di Luccia A., Bocca A. (1990) Use of plasmin to increase the sensitivity of the detection of bovine milk in ovine cheese by gel isoelectric focusing of caseins. Milchwissenschaft 45, 708-711.

Alais C., Jollès P. (1967) Isolation, purification, and analysis of two κ-casein-like fractions from sheep casein. J. Dairy Sci. 50, 1555-1561.

Alexander L.J., Stewart A.F., Mackinlay A.G., Kapelinskaya T.V., Tkach T.M., Gorodetsky S.I. (1988) Isolation and characterization of the bovine kappa-casein gene. Eur. J. Biochem. 178, 395-401.

Ali S., McClenaghan M., Simons J.P., Clark A.J. (1990) Characterisation of the alleles encoding ovine β-lactoglobulins A and B. Gene 91, 201-207.

Amigo L., Recio I., Ramos M. (2000) Genetic polymorphism of ovine milk proteins: its influence on technological properties of milk - a review. Int. Dairy J. 10, 135-149.

Anguita G., Martin R., Garcia T., Morales P., Haza A.I., Gonzalez I., Sanz B., Hernandez P.E. (1997) A competitive enzyme-linked immunosorbent assay for detection of bovine milk in ovine and caprine milk and cheese using a monoclonal antibody against bovine beta-casein. J. Food Prot. 60, 64-66.

Anton I., Zsolnai A., Fésüs L., Kukovics S., Molnár A. (1999) Survey of β-Lactoglobulin and α_{s1}-Casein polymorphisms in Hungarian dairy sheep breeds and crosses on DNA level. Arch. Tierz. 42, 387-392.

Árnyasi M., Komlósi I., Lien S., Czeglédi L., Nagy S., Jávor A. (2009) Searching for DNA markers for milk production and composition on chromosome 6 in sheep. J. Anim. Breed. Genet. 126, 142-147.

Arranz J.J., Bayón Y., Gabiña D., de la Fuente L.F., Ugarte E., San Primitivo F. (2001) New developments in the genetic improvement of dairy sheep. Proc. 7th Great Lakes Dairy Sheep Symposium, Eau Claire, Wisconsin, USA, pp. 94-115.

Astruc J.M., Barillet F., Fioretti M., Gabina D., Gootwine E., Mavrogenis A., Romberg F.J., Sanna S., Stefanake E. (2002) Report of the ICAR working group on milk recording of sheep. Proc. 33th Biennal Session of the International Committee for Animal Recording (ICAR), Interlaken, Switzerland.

Astruc J.M., Barillet F., Fioretti M., Gabina D., Gootwine E., Mavrogenis A.P., Romberg F.J., Sanna S.R., Stefanake E. (2004) Report of the working group on milk recording of

sheep. Proc. 34th Biennal Session of the International Committee for Animal Recording (ICAR), Sousse, Tunesia, pp. 315-322.

Astruc J.M., Barillet F., Fioretti M., Gabina D., Gootwine E., Romberg F.J., Sanna S.R. (2007) Report of the ICAR working group on milk recording of sheep. Proc. 35th Biennal Session of the International Committee for Animal Recording (ICAR), Kuopio, Finland, pp. 211-216.

Barillet F., Boichard D. (1987) Studies on dairy production of milking ewes I.- Estimates of genetic parameters for total milk composition and yield. Genet. Sel. Evol. 19, 459-474.

Barillet F., Marie C., Jacquin M., Lagriffoul G., Astruc J.M. (2001) The French Lacaune dairy sheep breed: use in France and abroad in the last 40 years. Livest. Prod. Sci. 71, 17-29.

Barillet F., Arranz J.J., Carta A. (2005) Mapping quantitative trait loci for milk production and genetic polymorphisms of milk proteins in dairy sheep. Genet. Sel. Evol. 37 (Suppl. 1), S109-S123.

Barillet F. (2007) Genetic improvement for dairy production in sheep and goats. Small Ruminant Res. 70, 60-75.

Barroso A., Cañon J., Dunner S. (1996) Preliminary results on the search for variability in ovine κ-casein. Anim. Genet. 27, 40.

Bastos E.C., A. Azevedo, J. Guedes-Pinto, H. (2001) Single strand conformation polymorphism (SSCP) detection in six genes in Portuguese indigenous sheep breed "Churra da Terra Quente". Biotechnol. Agron. Soc. Environ. 5, 7-15.

Bell K., McKenzie H.A. (1967) The whey proteins of ovine milk: β-lactoglobulins A and B. Biochim. Biophys. Acta 147, 123-134.

Bencini R. (2001) Factors affecting the quality of ewe`s milk. Proc. 7th Great Lakes Dairy Sheep Symposium, Eau Claire, Wisconsin, USA, pp. 52-83.

Bevilacqua C., Helbling J.C., Miranda G., Martin P. (2006) Translational efficiency of casein transcripts in the mammary tissue of lactating ruminants. Reprod. Nutr. Dev. 46, 567-578.

Bianchi L., Bolla A., Budelli E., Caroli A., Casoli C., Pauselli M., Duranti E. (2004) Effect of udder health status and lactation phase on the characteristics of Sardinian ewe milk. J. Dairy Sci. 87, 2401-2408.

Boettcher P.J., Caroli A., Stella A., Chessa S., Budelli E., Canavesi F., Ghiroldi S., Pagnacco G. (2004) Effects of casein haplotypes on milk production traits in Italian Holstein and Brown Swiss cattle. J. Dairy Sci. 87, 4311-4317.

Boisnard M., Petrissant G. (1985) Complete sequence of ovine α_{s2}-casein messenger RNA. Biochimie 67, 1043-1051.

Boisnard M., Hue D., Bouniol C., Mercier J.C., Gaye P. (1991) Multiple mRNA species code for two non-allelic forms of ovine α_{s2}-casein. Eur. J. Biochem. 201, 633-641.

Bolla A., Caroli A., Mezzelani R., Rizzi R., Pagnacco G., Fraghì A., Casu S. (1989) Milk protein markers and production in sheep. Anim. Genet. 20, 78-79.

Bonfatti V., Di Martino G., Cecchinato A., Degano L., Carnier P. (2010) Effects of beta-kappa-casein (CSN2-CSN3) haplotypes, beta-lactoglobulin (BLG) genotypes, and detailed protein composition on coagulation properties of individual milk of Simmental cows. J. Dairy Sci. 93, 3809-3817.

Bonsing J., Mackinlay A.G. (1987) Recent studies on nucleotide sequences encoding the caseins. J. Dairy Res. 54, 447-461.

Bouniol C., Printz C., Mercier J.C. (1993) Bovine α_{s2}-casein D is generated by exon VIII skipping. Gene 128, 289-293.

Boutinaud M., Rulquin H., Keisler D.H., Djiane J., Jammes H. (2002) Use of somatic cells from goat milk for dynamic studies of gene expression in the mammary gland. J. Anim. Sci. 80, 1258-1269.

Boyazoglu J., Morand-Fehr P. (2001) Mediterranean dairy sheep and goat products and their quality - A critical review. Small Ruminant Res. 40, 1-11.

Bramanti E., Sortino C., Onor M., Beni F., Raspi G. (2003) Separation and determination of denatured α_{s1}-, α_{s2}-, β- and κ-caseins by hydrophobic interaction chromatography in cows', ewes' and goats' milk, milk mixtures and cheeses. J. Chromatogr. A 994, 59-74.

Brew K., Grobler J.A. (1992) α-Lactalbumin. In: Advanced Dairy Chemistry (ed. by Fox, P.F.), pp. 191-230. Elsevier Science Publishers LTD, London.

Brozos C., Saratsis P., Boscos C., Kyriakis S.C., Tsakalof P. (1998) Effects of long-term recombinant bovine somatotropin (bST) administration on milk yield, milk composition and mammary gland health of dairy ewes. Small Ruminant Res. 29, 113-120.

Bruford M.W., Bradley D.G., Luikart G. (2003) DNA markers reveal the complexity of livestock domestication. Nat. Rev. Genet. 4, 900-910.

Busse P.J., Jarvinen K.M., Vila L., Beyer K., Sampson H.A. (2002) Identification of sequential IgE-binding epitopes on bovine α_{s2}-casein in cow's milk allergic patients. Int. Arch. Allergy Immunol. 129, 93-96.

Caio S., Izquierdo M., Gonzalez J., Hernández F.I., Corral J.M., Pinto L., Rodríguez J., Roa I. (2007) Influence of ovine β-casein phenotype on milk production and composition from Merino ewes. Proc. 5th International Symposium on the challenge to sheep and goats milk sectors, Alghero/Sardinia, Italy, pp. 173-175.

Calavia M.C., Burgos J. (1998) Ovine κ-casein in milk from Lacha Spanish sheep: Heterogeneity and total content. Int. Dairy J. 8, 779-786.

Caroli A., Chiatti F., Chessa S., Rignanese D., Bolla P., Pagnacco G. (2006) Focusing on the goat casein complex. J. Dairy Sci. 89, 3178-3187.

Caroli A., Rignanese D., Chiatti F., Chessa S., Bolla P. (2007a) Occurrence of α_{s2}-casein (*CSN1S2*) B variant in Sarda and Comisana ovine breeds. Proc. 5th International Symposium on the challenge to sheep and goats milk sectors, Alghero/Sardinia, Italy, pp. 46-48.

Caroli A., Chiatti F., Chessa S., Rignanese D., Ibeagha-Awemu E.M., Erhardt G. (2007b) Characterization of the casein gene complex in West African goats and description of a new α_{s1}-casein polymorphism. J. Dairy Sci. 90, 2989-2996.

Caroli A., Chessa S., Erhardt G.J. (2009) Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. J. Dairy Sci. 92, 5335-5352.

Caroli A., Rizzi R., Lühken G., Erhardt G. (2010) Short communication: milk protein genetic variation and casein haplotype structure in the Original Pinzgauer cattle. J. Dairy Sci. 93, 1260-1265.

Carta A., Sanna S.R., Casu S. (1995) Estimating Lactation Curves and Seasonal Effects for Milk, Fat and Protein in Sarda Dairy Sheep with a Test Day Model. Livest. Prod. Sci. 44, 37-44.

Carta A., Casu S., Salaris S. (2009) Invited review: Current state of genetic improvement in dairy sheep. J. Dairy Sci. 92, 5814-5833.

Casoli C., Duranti E., Morbidini L., Panella F., Vizioli V. (1989) Quantitative and compositional variations of Massese sheep milk by parity and stage of lactation. Small Ruminant Res. 2, 47-62.

Castillo V., Such X., Caja G., Casals R., Albanell E., Salama A.A. (2008) Effect of milking interval on milk secretion and mammary tight junction permeability in dairy ewes. J. Dairy Sci. 91, 2610-2619.

Ceriotti G., Chessa S., Bolla P., Budelli E., Bianchi L., Duranti E., Caroli A. (2004) Single nucleotide polymorphisms in the ovine casein genes detected by polymerase chain reaction-single strand conformation polymorphism. J. Dairy Sci. 87, 2606-2613.

Ceriotti G., Chiatti F., Bolla R., Martini M., Caroli A. (2005) Genetic variability of the ovine α_{s1}-casein. Ital. J. Anim. Sci. 4, 64-66.

Chatchatee P., Jarvinen K.M., Bardina L., Beyer K., Sampson H.A. (2001) Identification of IgE- and IgG-binding epitopes on α_{s1}-casein: differences in patients with persistent and transient cow's milk allergy. J. Allergy Clin. Immunol. 107, 379-383.

Chessa S., Bolla P., Dario C., Pieragostini E., Caroli A. (2003) Polimorfismi genetici lattoproteici nella razza ovina gentile di puglia: Monitoraggio mediante focalizzazione isoelectrica. Sci. Tecn. Latt.-Cas. 54, 1-8.

Chessa S., Chiatti F., Ceriotti G., Caroli A., Consolandi C., Pagnacco G., Castiglioni B. (2007) Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms. J. Dairy Sci. 90, 451-464.

Chessa S., Rignanese D., Berbenni M., Ceriotti G., Martini M., Pagnacco G., Caroli A. (2010) New genetic polymorphisms within ovine β- and α_{s2}-caseins. Small Ruminant Res. 88, 84-88.

Chianese L., Garro G., Addeo F., Lopez-Galvez G., Ramos M. (1993) Discovery of an ovine α_{s2}-casein variant. J. Dairy Res. 60, 485-493.

Chianese L., Garro G., Ferranti P., Malorni A., Addeo F., Rabasco A., Pons P.M. (1995) Discrete phosphorylation generates the electrophoretic heterogeneity of ovine *β*-casein. J. Dairy Res. 62, 89-100.

Chianese L., Garro G., Mauriello R., Laezza P., Ferranti P., Addeo F. (1996) Occurrence of five α_{s1}-casein variants in ovine milk. J. Dairy Res. 63, 49-59.

Chianese L., Mauriello R., Ferranti P., Tripaldi C., Taibi L., Dell'Aquila S. (1997) Relationship between α_{s1}-casein variants and clotting capability of ovine milk. Proc. IDF Seminar Milk Protein Polymorphism 2, Palmerston North, New Zealand, pp. 316-323.

Chianese L., Caira S., Garro G., Addeo F. (2007) Primary structure of ovine deleted variant α_{s1}-CN E. Proc. 5th International Symposium on the challenge to sheep and goats milk sectors, Alghero/Sardinia, Italy, pp. 58-60.

Chianese L. (2008) Personal communication: Existence of a further ovine *CSN1S1* variant *G*. Naples, Italy.

Chianese L., Quarto M., Pizzolongo F., Calabrese M.G., Caira S., Mauriello R., De Pascale S., Addeo F. (2009) Occurrence of genetic polymorphism at the α_{s1}-casein locus in Mediterranean water buffalo milk. Int. Dairy J. 19, 181-189.

Corral J.M., Padilla J.A., Izquierdo M. (2010) Associations between milk protein genetic polymorphisms and milk production traits in Merino sheep breed. Livest. Sci. 129, 73-79.

Cosenza G., Gallo D., Illario R., di Gregorio P., Senese C., Ferrara L., Ramunno L. (2003) A MvaI PCR-RFLP detecting a silent allele at the goat alpha-lactalbumin locus. J. Dairy Res. 70, 355-357.

de Gortari M.J., Freking B.A., Cuthbertson R.P., Kappes S.M., Keele J.W., Stone R.T., Leymaster K.A., Dodds K.G., Crawford A.M., Beattie C.W. (1998) A second-generation linkage map of the sheep genome. Mamm. Genome 9, 204-209.

Degen A.A. (2007) Sheep and goat milk in pastoral societies. Small Ruminant Res. 68, 7-19.

Eigel W.N., Butler J.E., Ernstrom C.A., Farrell H.M., Harwalkar V.R., Jenness R., Whitney R.M. (1984) Nomenclature of proteins of cow's milk: Fifth revision. J. Dairy Sci. 67, 1599-1631.

El-Saied U.M., Carriedo J.A., Baro J.A., De La Fuente L.F., San Primitivo F. (1998) Genetic correlations and heritabilities for milk yield and lactation length of dairy sheep. Small Ruminant Res. 27, 217-221.

El-Saied U.M., Carriedo J.A., de la Fuente L.F., San Primitivo F. (1999) Genetic parameters of lactation cell counts and milk and protein yields in dairy ewes. J. Dairy Sci. 82, 639-644.

El-Zahar K., Sitohy M., Choiset Y., Metro F., Haertle T., Chobert J.M. (2005) Peptic hydrolysis of ovine beta-lactoglobulin and alpha-lactalbumin - Exceptional susceptibility of native ovine beta-lactoglobulin to pepsinolysis. Int. Dairy J. 15, 17-27.

Elmaci C., Oner Y., Balcioglu M.S. (2006) Genetic polymorphism of β-lactoglobulin gene in native Turkish sheep breeds. Biochem. Genet. 44, 379-384.

Erhardt G. (1989b) Evidence for a third allele at the β-lactoglobulin (β-Lg) locus of sheep milk and its occurrence in different breeds. Anim. Genet. 20, 197-204.

Erhardt G. (1989a) Genetic polymorphisms of α-lactalbumin and β-lactoglobulin in sheep milk. Anim. Genet. 20, 76-77.

Erhardt G., Godovac-Zimmermann J., Conti A. (1989) Isolation and complete primary sequence of a new ovine wild-type β-lactoglobulin C. Biol. Chem. Hoppe-Seyler 370, 757-762.

Erhardt G. (1993) A new α_{S1}-casein allele in bovine milk and its occurrence in different breeds. Anim. Genet. 24, 65-66.

Erhardt G. (1996) Detection of a new kappa-casein variant in milk of Pinzgauer cattle. Anim. Genet. 27, 105-107.

Erhardt G., Juszczak J., Panicke L., Krick-Saleck H. (1998) Genetic polymorphism of milk proteins in Polish Red Cattle: a new genetic variant of *β*-lactoglobulin. J. Anim. Breed. Genet. 115, 63-71.

Erhardt G., Jäger S., Budelli E., Caroli A. (2002) Genetic polymorphism of goat α_{s2}-casein (*CSN1S2*) and evidence for a further allele. Milchwissenschaft 57, 137-140.

Eurostat: (2010) http://epp.eurostat.ec.europa.eu. 15.03.2010.

FAOstat: (2010) http://www.fao.org/. 15.03.2010.

Farrell H.M., Jr., Jimenez-Flores R., Bleck G.T., Brown E.M., Butler J.E., Creamer L.K., Hicks C.L., Hollar C.M., Ng-Kwai-Hang K.F., Swaisgood H.E. (2004) Nomenclature of the proteins of cows' milk-sixth revision. J. Dairy Sci. 87, 1641-1674.

Feligini M., Vlaco S., Curik V.C., Parma P., Greppi G., Enne G. (2005) A single nucleotide polymorphism in the sheep κ-casein coding region. J. Dairy Res. 72, 317-321.

Ferranti P., Malorni A., Nitti G., Laezza P., Pizzano R., Chianese L., Addeo F. (1995) Primary structure of ovine α_{s1}-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. J. Dairy Res. 62, 281-296.

Ferranti P., Chianese L., Malorni A., Migliaccio F., Stingo V., Addeo F. (1998) Copresence of deleted protein species generates structural heterogeneity of ovine α_{s1}-casein. J. Agric. Food Chem. 46, 411-416.

Ferranti P., Lilla S., Chianese L., Addeo F. (1999) Alternative nonallelic deletion is constitutive of ruminant α_{s1}-casein. J. Protein Chem. 18, 595-602.

Ferranti P., Pizzano R., Garro G., Caira S., Chianese L., Addeo F. (2001) Mass spectrometry-based procedure for the identification of ovine casein heterogeneity. J. Dairy Res. 68, 35-51.

Freitas C., Malcata F.X. (2000) Microbiology and biochemistry of cheeses with Appelation d'Origine Protegee and manufactured in the Iberian Peninsula from ovine and caprine milks. J. Dairy Sci. 83, 584-602.

Fuertes J.A., Gonzalo C., Carriedo J.A., San Primitivo F. (1998) Parameters of test day milk yield and milk components for dairy ewes. J. Dairy Sci. 81, 1300-1307.

Furet J.P., Mercier J.C., Soulier S., Gaye P., Hue-Delahaie D., Vilotte J.L. (1990) Nucleotide sequence of ovine κ-casein cDNA. Nucleic Acids Res. 18, 5286.

Gabina D.F., Arrese F., Arranz J., Beltran de Heredia I. (1993) Average milk yields and environmental effects on Latxa sheep. J. Dairy Sci. 76, 1191-1198.

Gaye P., Gautron J.P., Mercier J.C., Haze G. (1977) Amino terminal sequences of the precursors of ovine caseins. Biochem. Biophys. Res. Commun. 79, 903-911.

Gaye P., Hue-Delahaie D., Mercier J.C., Soulier S., Vilotte J.L., Furet J.P. (1986) Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland. Biochimie 68, 1097-1107.

Gaye P., Hue-Delahaie D., Mercier J.C., Soulier S., Vilotte J.L., Furet J.P. (1987) Complete nucleotide sequence of ovine α-lactalbumin mRNA. Biochimie 69, 601-608.

Giaccone P., Di Stasio L., Macciotta N.P., Portolano B., Todaro M., Cappio-Borlino A. (2000) Effect of β-lactoglobulin polymorphism on milk-related traits of dairy ewes analysed by a repeated measures design. J. Dairy Res. 67, 443-448.

Gootwine E., Goot H. (1996) Lamb and milk production of Awassi and East-Friesian sheep and their crosses under Mediterranean environment. Small Ruminant Res. 20, 255-260.

Groenen M.A., Dijkhof R.J., Verstege A.J., van der Poel J.J. (1993) The complete sequence of the gene encoding bovine αs2-casein. Gene 123, 187-193.

Grosclaude F., Martin P. (1997) Casein polymorphisms in the goat. In: Milk Protein Polymorphism pp. 241-253. International Dairy Federation, Brussels, Belgium.

Haenlein G.F.W. (2001a) Past, present, and future perspectives of small ruminant dairy research. J. Dairy Sci. 84, 2097-2115.

Haenlein G.F.W. (2001b) The nutritional value of sheep milk. Int. J. Animal Sci. 16, 253-268.

Haenlein G.F.W., Wendorff W.L. (2006) Sheep milk. In: Handbook of milk of non-bovine mammals (ed. by Park, Y.W., Haenlein, G.F.W.), pp. 137-194. Blackwell Publishing, Ames, Iowa, USA.

Haenlein G.F.W. (2007) About the evolution of goat and sheep milk production. Small Ruminant Res. 68, 3-6.

Hamann H., Horstick A., Wessels A., Distl O. (2004) Estimation of genetic parameters for test day milk production, somatic cell score and litter size at birth in East Friesian ewes. Livest. Prod. Sci. 87, 153-160.

Hambling S.G., McAlpine AlS., Sawyer L. (1992) β-Lactoglobulin. In: Advanced Dairy Chemistry - 1: Proteins (ed. by Fox, P.F.), pp. 141-190. Elsevier Science Publishers LTD, London.

Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J. (1988) Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene. Nucleic Acids Res. 16, 10379-10380.

Hartmann K. (2010) Genetische Diversität bei europäischen, asiatischen und afrikanischen Schaf- und Ziegenrassen am κ-Kasein-(*CSN3*) und Interleukin-2-Genort (*IL2*). Dissertation Fachbereich Veterinärmedizin, Justus-Liebig-Universität, Gießen.

Hayes H.C., Petit E.J. (1993) Mapping of the β-lactoglobulin gene and of an immunoglobulin M heavy chain-like sequence to homoeologous cattle, sheep, and goat chromosomes. Mamm. Genome 4, 207-210.

Hayes H.C., Popescu P., Dutrillaux B. (1993) Comparative gene mapping of lactoperoxidase, retinoblastoma, and α-lactalbumin genes in cattle, sheep, and goats. Mamm. Genome 4, 593-597.

Heck J.M., Schennink A., van Valenberg H.J., Bovenhuis H., Visker M.H., van Arendonk J.A., van Hooijdonk A.C. (2009) Effects of milk protein variants on the protein composition of bovine milk. J. Dairy Sci. 92, 1192-1202.

Hinrichs J. (2004) Mediterranean milk and milk products. Eur. J. Nutr. 43 Suppl 1, I/12-17.

Hobbs A.A., Rosen J.M. (1982) Sequence of rat α- and γ-casein mRNAs: evolutionary comparison of the calcium-dependent rat casein multigene family. Nucleic Acids Res. 10, 8079-8098.

Holt C., Sawyer L. (1988) Primary and predicted secondary structures of the caseins in relation to their biological functions. Protein Eng. 2, 251-259.

http://www.animalgenome.org/sheep/maps/: NAGRP Sheep Genome Coordination Program. 12.05.2010.

http://www.ernaehrung.de: Deutsches Ernährungsberatungs- und -informationsnetz. 25.03.2010.

http://www.livestockgenomics.csiro.au/: Ovine Genome Assembly v1.0. 15.06.2010.

http://www.waap.it/sheep_enquiry/: Dairy sheep milk enquiry-On-line database. 06.04.2010.

https://isgcdata.agresearch.co.nz/: OAR True Sheep Chromosomes ver.1.0. ISGC Data. 02.09.2010.

Ikonen T., Ahlfors K., Kempe R., Ojala M., Ruottinen O. (1999) Genetic parameters for the milk coagulation properties and prevalence of noncoagulating milk in Finnish dairy cows. J. Dairy Sci. 82, 205-214.

Imam Ghali M., Saidi-Mehtar N., Guerin G. (1991) Sheep gene mapping: additional DNA markers included (CASB, CASK, LALBA, IGF-1 and AMH). Anim. Genet. 22, 165-172.

Jaayid T.A. (2006) Genetic polymorphism of β-lactoglobulin (β-Lg) in two Russian sheep breeds. Egyptian Journal of Sheep, Goat and Desert Animal Science 1, 7-14.

Jandal J.M. (1996) Comparative aspects of goat and sheep milk. Small Ruminant Res. 22, 177-185.

Jollès J., Fiat A.M., Schoentgen F., Alais C., Jollès P. (1974) The amino acid sequence of sheep κ_A-casein. II. Sequence studies concerning the κ_A-caseinoglycopeptide and establishment of the complete primary structure of the protein. Biochim. Biophys. Acta 365, 335-343.

Jones W.K., Yu-Lee L.Y., Clift S.M., Brown T.L., Rosen J.M. (1985) The rat casein multigene family. Fine structure and evolution of the β-casein gene. J. Biol. Chem. 260, 7042-7050.

Kaminski S., Ahman A., Ruść A., Wójcik E., Malewski T. (2005) MilkProtChip - a microarray of SNPs in candidate genes associated with milk proteins biosynthesis - development and validation. J. Appl. Genet. 46, 45-58.

King J.W.B. (1966) The casein of sheep milk. Proc. 10th European Conference of Animal Blood Groups and Biochemical Polymorphism, Paris, France, pp. 427-431.

Kolde H.J., Braunitzer G. (1983) The primary structure of ovine β-lactoglobulin: 1. Isolation of the peptides and sequence. Milchwissenschaft 38, 18-20.

Küpper J., Chessa S., Rignanese D., Caroli A., Erhardt G. (2010) Divergence at the casein haplotypes in dairy and meat goat breeds. J. Dairy Res. 77, 56-62.

Law A.J.R. (1992) Quantitative fractionation of ovine casein by cation-exchange FPLC. Milchwissenschaft 47, 279-282.

Legarra A., Ugarte E. (2001) Genetic parameters of milk traits in Latxa dairy sheep. Anim. Sci. 73, 407-412.

Leroux C., Mazure N., Martin P. (1992) Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat α_{s1}-casein transcripts. Structural organization of the relevant gene. J. Biol. Chem. 267, 6147-6157.

Lévéziel H., Metenier L., Guerin G., Cullen P., Provot C., Bertaud M., Mercier J.C. (1991) Restriction fragment length polymorphism of ovine casein genes: close linkage between the α_{s1}-, α_{s2}-, β- and κ-casein loci. Anim. Genet. 22, 1-10.

Lühken G., Caroli A., Ibeagha-Awemu E.M., Erhardt G. (2009) Characterization and genetic analysis of bovine α_{s1}-casein *I* variant. Anim. Genet. 40, 479-485.

Maddox J.F., Davies K.P., Crawford A.M., Hulme D.J., Vaiman D., Cribiu E.P., Freking B.A., Beh K.J., Cockett N.E., Kang N., Riffkin C.D., Drinkwater R., Moore S.S., Dodds K.G., Lumsden J.M., van Stijn T.C., Phua S.H., Adelson D.L., Burkin H.R., Broom J.E., Buitkamp J., Cambridge L., Cushwa W.T., Gerard E., Galloway S.M.,

Harrison B., Hawken R.J., Hiendleder S., Henry H.M., Medrano J.F., Paterson K.A., Schibler L., Stone R.T., van Hest B. (2001) An enhanced linkage map of the sheep genome comprising more than 1000 loci. Genome Res. 11, 1275-1289.

Mahé M.F., Miranda G., Queval R., Bado A., Zafindrajaona P.S., Grosclaude F. (1999) Genetic polymorphism of milk proteins in African Bos taurus and Bos indicus populations. Characterization of variants α_{s1}-Cn H and κ-Cn J. Genet. Sel. Evol. 31, 239-253.

Manfredi E., Ricordeau G., Barbieri M.E., Amigues Y., Bibé B. (1995) Génotype caséine α_{s1} et sélection des boucs sur descendance dans les races Alpine et Saanen. Genet. Sel. Evol. 27, 451-458.

Mariani P., Summer A., Anghinetti A., Senese C., Di Gregorio P., Rando P., Serventi P. (1995) Effetti dell'allele α_{s1}-Cn G sulla ripartizione percentuale delle caseine α_{s1}, α_{s2}, β e κ in vacche di razza Bruna. Ind. Latte 31, 3-13.

Martin P., Leroux C. (1994) Characterization of a further α_{s1}-casein variant generated by exon skipping. Proc. 24th International Society of Animal Genetics Conference, Prague, Czech Republic, p. 86.

Martin P., Szymanowska M., Zwierzchowski L., Leroux C. (2002) The impact of genetic polymorphisms on the protein composition of ruminant milks. Reprod. Nutr. Dev. 42, 433-459.

Martin P., Ferranti P., Leroux C., Addeo F. (2003) Non-bovine caseins: Quantitative variability and molecular diversity. In: Advanced Dairy Chemistry Volume 1: Proteins (ed. by Fox, P.F., McSweeney, P.L.H.), pp. 277-317. Kluwer Academic/Plenum Publishers, New York.

McKenzie H.A. (1996) alpha-Lactalbumins and lysozymes. EXS 75, 365-409.

Meisel H. (1998) Overview on milk protein-derived peptides. Int. Dairy J. 8, 363-373.

Meisel H., Bockelmann W. (1999) Bioactive peptides encrypted in milk proteins: proteolytic activation and thropho-functional properties. Anton. Leeuw. Int. J. G. 76, 207-215.

Mele M., Conte G., Serra A., Buccioni A., Secchiari P. (2007) Relationship between beta-lactoglobulin polymorphism and milk fatty acid composition in milk of Massese dairy ewes. Small Ruminant Res. 73, 37-44.

Mercier J.C., Haze G., Gaye P., Hue D. (1978a) Amino terminal sequence of the precursor of ovine β-lactoglobulin. Biochem. Biophys. Res. Commun. 82, 1236-1245.

Mercier J.C., Haze G., Gaye P., Petrissant G., Hue D., Boisnard M. (1978b) Amino terminal sequence of the precursor of ovine α-lactalbumin. Biochem. Biophys. Res. Commun. 85, 662-670.

Mercier J.C., Gaye P., Soulier S., Hue-Delahaie D., Vilotte J.L. (1985) Construction and identification of recombinant plasmids carrying cDNAs coding for ovine α_{s1}-, α_{s2}-, β-, κ-casein and β-lactoglobulin. Nucleotide sequence of α_{s1}-casein cDNA. Biochimie 67, 959-971.

Mercier J.C., Vilotte J.L. (1993) Structure and function of milk protein genes. J. Dairy Sci. 76, 3079-3098.

Moatsou G., Samolada M., Katsabeki A., Anifantakis E. (2004) Casein fraction of ovine milk from indigenous Greek breeds. Lait 84, 285-296.

Mohr U., Koczan D., Linder D., Hobom G., Erhardt G. (1994) A single point mutation results in A allele-specific exon skipping in the bovine α_{s1}-casein mRNA. Gene 143, 187-192.

Moioli B., Pilla F., Tripaldi C. (1998) Detection of milk protein genetic polymorphisms in order to improve dairy traits in sheep and goats: A review. Small Ruminant Res. 27, 185-195.

Moioli B., D' Andrea M., Pilla F. (2007) Candidate genes affecting sheep and goat milk quality. Small Ruminant Res. 68, 179-192.

Moody D.E., Pomp D., Newman S., MacNeil M.D. (1996) Characterization of DNA polymorphisms in three populations of hereford cattle and their associations with growth and maternal EPD in line 1 herefords. J. Anim. Sci. 74, 1784-1793.

Morand-Fehr P., Boyazoglu J. (1999) Present state and future outlook of the small ruminant sector. Small Ruminant Res. 34, 175-188.

Moreno F.J., Recio I., Olano A., Lopez-Fandino R. (2000) Chromatographic characterization of ovine κ-casein macropeptide. J. Dairy Res. 67, 349-359.

Moroni P., Pisoni G., Varisco G., Boettcher P. (2007) Effect of intramammary infection in Bergamasca meat sheep on milk parameters and lamb growth. J. Dairy Res. 74, 340-344.

Mroczkowski S., Korman K., Erhardt G., Piwczynski D., Borys B. (2004) Sheep milk protein polymorphism and its effect on milk performance of Polish Merino. Arch. Tierz. 47, 114-121.

Nilsen H., Olsen H.-G., Hayes B., Sehested E., Svendsen M., Nome T., Meuwissen T., Lien S. (2009) Casein haplotypes and their association with milk production traits in Norwegian Red cattle. Genet. Sel. Evol. doi:10.1186/1297-9686-41-24.

Nudda A., Feligini M., Battacone G., Macciotta N.P.P., Pulina G. (2003) Effects of lactation stage, parity, β-lactoglobulin genotype and milk SCC on whey protein composition in Sarda dairy ewes. Ital. J. Anim. Sci. 2, 29-39.

Otaviano A.R., Lima A.L.F., Laureano M.M.M., Sena J.A.D., de Albuquerque L.G., Tonhati H. (2008) Beta-casein gene polymorphism permits identification of bovine milk mixed with bubaline milk in mozzarella cheese. Genet. Mol. Biol. 31, 902-905.

Othmane M.H., Carriedo J.A., San Primitivo F., De La Fuente L.F. (2002a) Genetic parameters for lactation traits of milking ewes: protein content and composition, fat, somatic cells and individual laboratory cheese yield. Genet. Sel. Evol. 34, 581-596.

Othmane M.H., De La Fuente L.F., Carriedo J.A., San Primitivo F. (2002b) Heritability and genetic correlations of test day milk yield and composition, individual laboratory cheese yield, and somatic cell count for dairy ewes. J. Dairy Sci. 85, 2692-2698.

Othmane M.H., Carriedo J.A., De la Fuente F., San Primitivio F. (2002c) Factors affecting test-day milk composition in dairy ewes, and relationships amongst various milk components. J. Dairy Res. 69, 53-62.

Pandya A.J., Ghodke K.M. (2007) Goat and sheep milk products other than cheeses and yoghurt. Small Ruminant Res. 68, 193-206.

Pariset L., Cappuccio I., Ajmone-Marsan P., Bruford M., Dunner S., Cortes O., Erhardt G., Prinzenberg E.M., Gutscher K., Joost S., Pinto-Juma G., Nijman I.J., Lenstra J.A., Perez T., Valentini A., Consortium E. (2006) Characterization of 37 breed-specific single-nucleotide polymorphisms in sheep. J. Hered. 97, 531-534.

Park Y.W., Haenlein G.F.W. (2006) Overview of Milk of non-bovine mammals. In: Handbook of milk of non-bovine mammals (ed. by Park, Y.W., Haenlein, G.F.W.), pp. 3-10. Blackwell Publishing, Ames, Iowa, USA.

Park Y.W., Juarez M., Ramos M., Haenlein G.F.W. (2007) Physico-chemical characteristics of goat and sheep milk. Small Ruminant Res. 68, 88-113.

Passey R., Glenn W., Mackinlay A. (1996) Exon skipping in the ovine α_{s1}-casein gene. Comp. Biochem. Phys. B 114, 389-394.

Pellegrini O., Remeuf F., Rivemale M., Barillet F. (1997) Renneting properties of milk from individual ewes: Influence of genetic and non-genetic variables, and relationship with physicochemical characteristics. J. Dairy Res. 64, 355-366.

Peter C., Bruford M., Perez T., Dalamitra S., Hewitt G., Erhardt G. (2007) Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. Anim. Genet. 38, 37-44.

Phelan M., Aherne A., FitzGerald R.J., O'Brien N.M. (2009) Casein-derived bioactive peptides: Biological effects, industrial uses, safety aspects and regulatory status. Int. Dairy J. 19, 643-654.

Picariello G., Rignanese D., Chessa S., Ceriotti G., Trani A., Caroli A., Di Luccia A. (2009) Characterization and genetic study of the ovine α_{s2}-casein (*CSN1S2*) allele *B*. Protein J. 28, 333-340.

Pilla F., Bevilacqua C., Leroux C., Fraghi A., Martin P. (1998) Genotyping of α-s1 casein in sheep. Anim. Genet. 29, 472-473.

Pirisi A., Piredda G., Papoff C.M., Di Salvo R., Pintus S., Garro G., Ferranti P., Chianese L. (1999) Effects of sheep α_{s1}-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. J. Dairy Res. 66, 409-419.

Prinzenberg E.M., Anglade P., Ribadeau-Dumas B., Erhardt G. (1998) Biochemical characterization of bovine α_{s1}-casein F and genotyping with sequence-specific primers. J. Dairy Res. 65, 223-231.

Prinzenberg E.M., Erhardt G. (1999) Molecular genetic characterization of ovine β-lactoglobulin C allele and detection by PCR-RFLP. J. Anim. Breed. Genet. 116, 9-14.

Prinzenberg E.M., Gutscher K., Chessa S., Caroli A., Erhardt G. (2005) Caprine κ-casein (CSN3) polymorphism: new developments in molecular knowledge. J. Dairy Sci. 88, 1490-1498.

Prinzenberg E.M., Jianlin H., Erhardt G. (2008) Genetic Variation in the κ-Casein Gene (*CSN3*) of Chinese Yak (Bos grunniens) and Phylogenetic Analysis of *CSN3* Sequences in the Genus Bos. J. Dairy Sci. 91, 1198-1203.

Provot C., Persuy M.A., Mercier J.C. (1989) Complete nucleotide sequence of ovine β-casein cDNA: inter-species comparison. Biochimie 71, 827-832.

Provot C., Persuy M.A., Mercier J.C. (1995) Complete sequence of the ovine β-casein-encoding gene and interspecies comparison. Gene 154, 259-263.

Rahmann G. (2001) Milchschafhaltung im ökologischen Landbau. Bad Dürkheim: Stiftung Okologie & Landbau.

Ramos A.M., Matos C.A.P., Russo-Almeida P.A., Bettencourt C.M.V., Matos J., Martins A., Pinheiro C., Rangel-Figueiredo T. (2009) Candidate genes for milk production traits in Portuguese dairy sheep. Small Ruminant Res. 82, 117-121.

Rampilli M., Cecchi F., Giuliotti L., Cattaneo T.M.P. (1997) The influence of β-lactoglobulin genetic polymorphism on protein distribution and coagulation properties in milk of

massesse breed ewes. Proc. 2nd IDF Seminar Milk Protein Polymorphism, Palmerston North, New Zealand, pp. 311-315.

Ramunno L., Cosenza G., Rando A., Macciotta N.P.P., Pappalardo M., Masina P. (1997) Identification of carriers of the Welsh CASA1 variant using an allele-specific PCR method. Anim. Genet. 28, 154-155.

Ramunno L., Cosenza G., Pappalardo M., Longobardi E., Gallo D., Pastore N., Di Gregorio P., Rando A. (2001) Characterization of two new alleles at the goat *CSN1S2* locus. Anim. Genet. 32, 264-268.

Raynal-Ljutovac K., Lagriffoul G., Paccard P., Guillet I., Chilliard Y. (2008) Composition of goat and sheep milk products: An update. Small Ruminant Res. 79, 57-72.

Reale S., Campanella A., Merigioli A., Pilla F. (2008) A novel method for species identification in milk and milk-based products. J. Dairy Res. 75, 107-112.

Recio I., Fernández-Fournier A., Martín-Álvarez P.J., Ramos M. (1997) β-lactoglobulin polymorphism in ovine breeds: influence on cheesemaking properties and milk composition. Lait 77, 259-265.

Recio I., de la Fuente M., A., Juárez M., Ramos M. (2009) Bioactive components in sheep milk. In: Bioactive components in milk and dairy products (ed. by Park, Y.W.), pp. 83-104. Wiley-Blackwell, Ames, Iowa, USA.

Richardson B.C., Mercier J.C. (1979) The primary structure of the ovine β-caseins. Eur. J. Biochem. 99, 285-297.

Rijnkels M. (2002) Multispecies comparison of the casein gene loci and evolution of casein gene family. J. Mammary Gland Biol. Neoplasia 7, 327-345.

Rodriguez E., Martin R., Garcia T., Hernandez P.E., Sanz B. (1990) Detection of cows' milk in ewes' milk and cheese by an indirect enzyme-linked immunosorbent assay (ELISA). J. Dairy Res. 57, 197-205.

Ruiz R., Oregui L.M., Herrero M. (2000) Comparison of models for describing the lactation curve of Latxa sheep and an analysis of factors affecting milk yield. J. Dairy Sci. 83, 2709-2719.

Sabour M.P., Lin C.Y., Lee A.J., McAllister A.J. (1996) Association between milk protein genetic variants and genetic values of Canadian Holstein bulls for milk yield traits. J. Dairy Sci. 79, 1050-1056.

Sacchi P., Chessa S., Budelli E., Bolla P., Ceriotti G., Soglia D., Rasero R., Cauvin E., Caroli A. (2005) Casein haplotype structure in five Italian goat breeds. J. Dairy Sci. 88, 1561-1568.

Sanchez A., Ilahi H., Manfredi E., Serradilla M. (2005) Potential benefit from using the α_{s1}-casein genotype information in a selection scheme for dairy goats. J. Anim. Breed. Genet. 122, 21-29.

Sanna S.R., Carta A., Casu S. (1997) (Co) variance component estimates for milk composition traits in Sarda dairy sheep using a bivariate animal model. Small Ruminant Res. 25, 77-82.

Schmidt D.V., Ebner K.E. (1972) Multiple forms of pig, sheep and goat α-lactalbumin. Biochim. Biophys. Acta 263, 714-720.

Sevi A., Taibi L., Albenzio M., Muscio A., Annicchiarico G. (2000) Effect of parity on milk yield, composition, somatic cell count, renneting parameters and bacteria counts of Comisana ewes. Small Ruminant Res. 37, 99-107.

Snowder G.D., Glimp H.A. (1991) Influence of Breed, Number of Suckling Lambs, and Stage of Lactation on Ewe Milk-Production and Lamb Growth under Range Conditions. J. Anim. Sci. 69, 923-930.

Staiger E.A., Thonney M.L., Buchanan J.W., Rogers E.R., Oltenacu P.A., Mateescu R.G. (2010) Effect of prolactin, β-lactoglobulin, and κ-casein genotype on milk yield in East Friesian sheep. J. Dairy Sci. 93, 1736-1742.

Storry J.E., Alistair S.G., Millard D., Owen A.J., Ford G.D. (1983) Chemical composition and coagulation properties of renneted milks from different breeds and species of ruminants. J. Dairy Res. 50, 215-229.

Threadgill D.W., Womack J.E. (1990) Genomic analysis of the major bovine milk protein genes. Nucleic Acids Res. 18, 6935-6942.

Trujillo A.J., Casals I., Guamis B. (2000) Analysis of major ovine milk proteins by reversed-phase high-performance liquid chromatography and flow injection analysis with electrospray ionization mass spectrometry. J. Chromatogr. A 870, 371-380.

Tziboula A., Horne D.S. (1999) The role of α_{s1}-casein in the structure of caprine casein micelles. Int. Dairy J. 9, 173-178.

Ugarte E., Ruiz R., Gabina D., de Heredia I.B. (2001) Impact of high-yielding foreign breeds on the Spanish dairy sheep industry. Livest. Prod. Sci. 71, 3-10.

Vincenzetti S., Polidori P., Mariani P., Vita A. (2008) Protein fraction characterization of sheep milk from the Comisana breed. Vet. Res. Commun. 32, S179-181.

Ward T.J., Honeycutt R.L., Deri J.N. (1997) Nucleotide sequence evolution at the κ-casein locus: Evidence for positive selection within the family bovidae. Genetics 147, 1863-1872.

Weimann C., Meisel H., Erhardt G. (2009) Bovine κ-casein variants result in different angiotensin I converting enzyme (ACE) inhibitory peptides. J. Dairy Sci. 92, 1885-1888.

Wessels G., Hamann H., Erhardt G., Distl O. (2004) Genotypeffekte von Milchprotein-Polymorphismen auf die Milchleistung beim Ostfriesischen Milchschaf. Berl. Münch. Tierärztl. Wochenschr. 117, 414-419.

Wilson L.L., Varela-Alvarez H., Hess C.E., Rugh M.C. (1971) Influence of energy level, creep feeding and lactation stage on ewe milk and lamb growth characters. J. Anim. Sci. 33, 686-690.

Zumbach B., Peters K.J. (2008) Voraussetzungen für Zuchtprogramme und wirksame Zuchtwertschätzverfahren in Europa. Proc. Internationale Milchschaftagung, Berlin, Germany, pp. 51-64.

Acknowledgement

First of all I would like to thank Prof. Dr. G. Erhardt for giving the topic of this thesis, for his assistance and support during the whole process of research, and for the suggestions and the fast corrections in the writing process of the manuscripts and the whole thesis.

Moreover, I am very grateful to Prof. Dr. Dr. habil. E. Usleber and Prof. Dr. V. Dzapo for their willingness to accept the role as second and third referee.

I express my gratitude to all colleagues of the department of animal breeding and genetics for the good teamwork, the pleasant time, and the enjoyable coversations. Especially, I would like to thank Gesine and Christina for their valuable advices concerning the realisation of lab methods and Horst for his guidance and assistance in statistical analyses. Sincere thanks goes to Sylvia for the introduction to the world of IEF, the help with the interpretation, and for the assistance in allocating samples. Julia and Daniela, I would like to thank for their help in handling bigger and smaller computer problems, and Helen for reading and correcting big parts of the thesis. Furthermore, I am grateful to all other "lab-ladies" and doctoral candidates for their help in handling the samples, as well as for the tips to restrain some unmanageable lab methods.

Additionally, I give my thanks to the staff of the research station "Oberer Hardthof", particularly to Anja, Stefan, and Ralf, for making the collection of milk samples possible.

Moreover, I would like to thank the employees of the research station Iden and the sheep breeders B. and R. Althoff, Steimel-Weroth; F. Atema, VE Laren gld Niederlande; R. and R. Bergmann, Betheln; M. Dors, Lehrte; W. Görne, Bennewitz; H.-U. Hartmann, Uetze-Obershagen; Heidschnuckenhof Jeversen, Wietze-Jeversen; B. Hucke, Hardegsen; M. Ise, Gießen-Kleinlinden; J. Jarick, Kolkwitz-Kackrow; O. Junker-Matthes, Bad Berleburg-Diedenhausen; W. Rompf, Breitscheid; G. Rüpke, Winsen (Aller); S. Prokasky, Mengerskirchen; E. Prunzel-Ulrich, Landolfshausen; S. and K. Schäfer, Marienheide-Scharde; Dr. D. Schnehage, Koldingen, and Verein Schäferhof Neuenkirchen, Neuenkirchen for the possibility to analyse milk from their sheep and for the allocation of all necessary data.

Special thanks I express to my family and my two- and four-legged friends for their support during the whole time of study and dissertation. My sincere thanks goes to my boyfriend Raffi, who was always good-humoured during these times and who gave me motivation and made me glad with diverse poems.

Danksagung

Zunächst danke ich Herrn Prof. Dr. G. Erhardt für die Bereitstellung des Dissertationsthemas, für die Betreuung während der Bearbeitung, die vielen Anregungen und die schnellen Korrekturen bei der Erstellung der Manuskripte und der kompletten Doktorarbeit.

Weiterhin bin ich den Herren Prof. Dr. Dr. habil. E. Usleber und Prof. Dr. V. Dzapo für ihre Bereitschaft das Amt des zweiten beziehungsweise dritten Gutachters zu übernehmen sehr dankbar.

Allen Mitarbeitern und Mitarbeiterinnen des Instituts für Tierzucht und Haustiergenetik danke ich für die gute Zusammenarbeit, die angenehme Zeit und die vielen netten Gespräche. Im Besonderen danke ich Gesine und Christina für die wertvollen Ratschläge bei der Durchführung der Laborarbeit und Horst für die Beratung und Betreuung bei den statistischen Analysen. Herzlich gedankt sei auch Sylvia für die Einführung in die Welt der IEF und die Hilfe bei der Auswertung und Probennahme. Julia and Daniela möchte ich für ihre Hilfe bei größeren und kleineren Computerproblemen und Helen für die Durchsicht und Korrektur großer Teile der Arbeit danken. Auch allen anderen „Labordamen“ und Doktoranden/Doktorandinnen danke ich für die Hilfe bei der Bewältigung der Proben, sowie für die vielen Tipps, die dazu beitrugen manche widerspenstige Labormethode doch zu bändigen.

Außerdem danke ich den Mitarbeitern der Lehr- und Forschungsstation „Oberer Hardthof“, vor allem Anja, Stefan und Ralf, ohne die die Probennahme nicht möglich gewesen wäre. Weiterhin bedanke ich mich bei den Mitarbeitern der Lehr- und Versuchsanstalt Iden und bei den Schafzüchtern B. und R. Althoff, Steimel-Weroth; F. Atema, VE Laren gld Niederlande; R. und R. Bergmann, Betheln; M. Dors, Lehrte; W. Görne, Bennewitz; H.-U. Hartmann, Uetze-Obershagen; Heidschnuckenhof Jeversen, Wietze-Jeversen; B. Hucke, Hardegsen; M. Ise, Gießen-Kleinlinden; J. Jarick, Kolkwitz-Kackrow; O. Junker-Matthes, Bad Berleburg-Diedenhausen; W. Rompf, Breitscheid; G. Rüpke, Winsen (Aller); S. Prokasky, Mengerskirchen; E. Prunzel-Ulrich, Landolfshausen; S. und K. Schäfer, Marienheide-Scharde; Dr. D. Schnehage, Koldingen und dem Verein Schäferhof Neuenkirchen, Neuenkirchen für die Möglichkeit Milch ihrer Schafe untersuchen zu dürfen und die Bereitstellung aller nötigen Daten.

Ein besonderer Dank geht natürlich auch an meine Familie und meine zwei- und vierbeinigen Freunde für ihre Unterstützung während der gesamten Studien- und Dissertationszeit. Mein herzlichster Dank gilt meinem Freund Raffi, der in diesen Zeiten immer gut gelaunt an meiner Seite stand und mich mit verschiedensten Gedichten motiviert und erfreut hat.

Ich erkläre:

Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe.
Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.
Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

édition scientifique
VVB LAUFERSWEILER VERLAG

VVB LAUFERSWEILER VERLAG
STAUFENBERGRING 15
D-35396 GIESSEN

Tel: 0641-5599888 Fax: -5599890
redaktion@doktorverlag.de
www.doktorverlag.de

ISBN: 978-3-8359-5804-3