CLONING AND SCREENING FOR SEQUENCE VARIATION IN *CSN1S2* GENE IN SELECTED INDIAN GOAT BREEDS

THESIS SUBMITTED TO THE
ICAR-NATIONAL DAIRY RESEARCH INSTITUTE, KARNAL
(DEEMED UNIVERSITY)
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR
THE AWARD OF THE DEGREE OF

MASTER OF VETERINARY SCIENCE

IN

ANIMAL BIOCHEMISTRY

By

SAMENI DEEPIKA

B.V.Sc. & A.H.

ANIMAL BIOCHEMISTRY DIVISION
ICAR-NATIONAL DAIRY RESEARCH INSTITUTE
(DEEMED UNIVERSITY)
KARNAL-132001 (HARYANA), INDIA

2019

Regn. No. 17-M-AB-07

CLONING AND SCREENING FOR SEQUENCE VARIATION IN *CSN1S2* GENE IN SELECTED INDIAN GOAT BREEDS

BY

SAMENI DEEPIKA

Thesis Submitted to the

ICAR-NATIONAL DAIRY RESEARCH INSTITUTE, KARNAL

(Deemed University)

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF VETERINARY SCIENCE

IN

ANIMAL BIOCHEMISTRY

Approved by:

(Langwan 27/07/19)
EXTERNAL EXAMINER

Sujeena 27/7/19
Dr. SUNITA MEENA
Major Advisor

Members of Advisory Committee

1. **Dr. (Mrs.) Suman Kapila**
   Principal Scientist, Animal Biochemistry Division ..... Suman Kapila 27/7/19

2. **Dr. Suneel Kumar Onteru**
   Senior Scientist, Animal Biochemistry Division ..... O. Suneel kumar

3. **Dr. Sachinandan De**
   Principal Scientist, Animal Biotechnology Division ..... SDe 27/7/2019

4. **Dr. I.D. Gupta**
   Principal Scientist, Animal Genetics and Breeding Division (Joint Director's Nominee) 27/7/19

ANIMAL BIOCHEMISTRY DIVISION
ICAR-NATIONAL DAIRY RESEARCH INSTITUTE
(DEEMED UNIVERSITY) KARNAL -132001
(HARYANA)

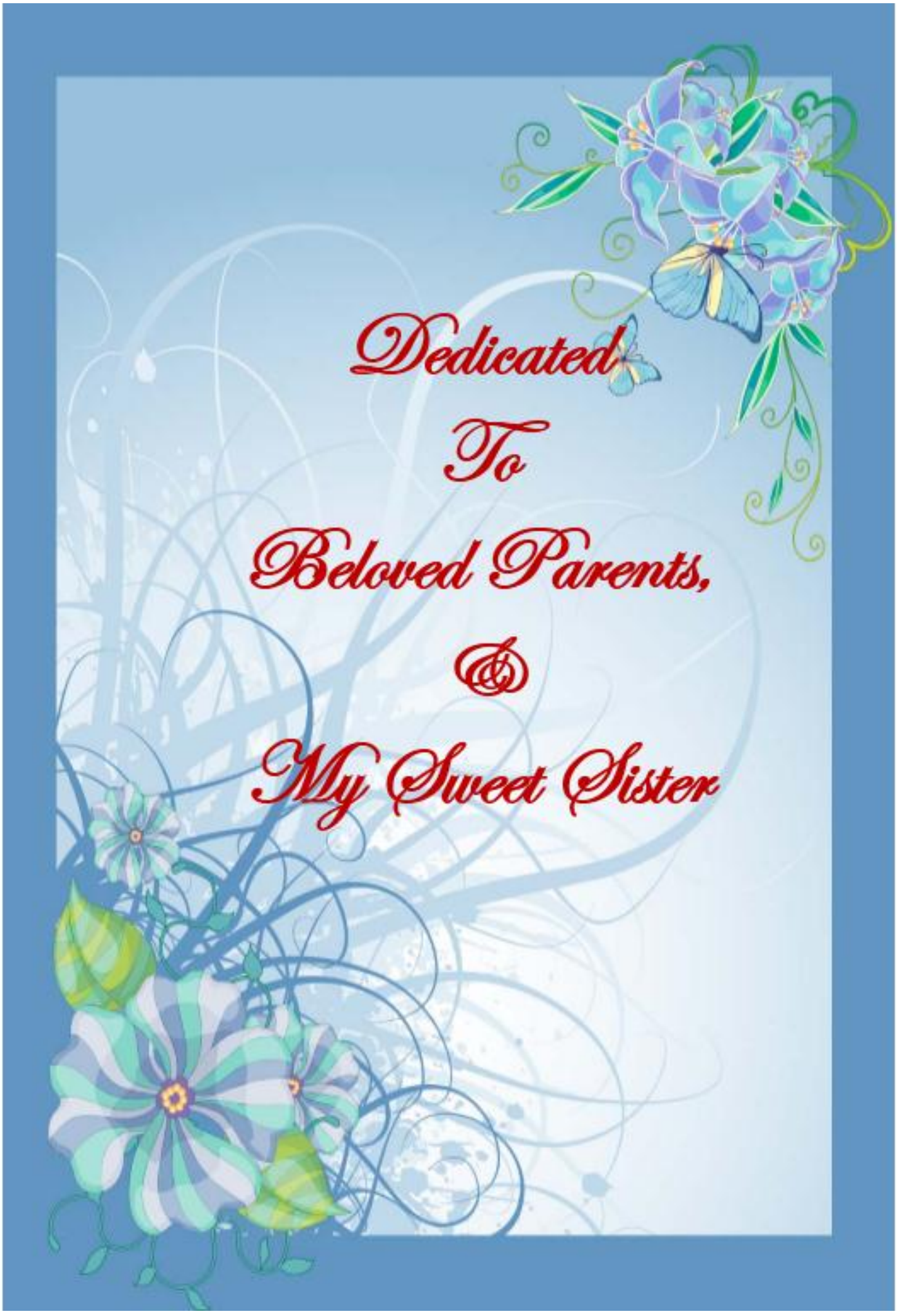
Dr. SUNITA MEENA
Scientist

CERTIFICATE

This is to certify that the thesis entitled "CLONING AND SCREENING FOR SEQUENCE VARIATION IN *CSN1S2* GENE IN SELECTED INDIAN GOAT BREEDS" submitted by SAMENI DEEPIKA, in partial fulfillment of the requirement for award of the degree of MASTER OF VETERINARY SCIENCE in ANIMAL BIOCHEMISTRY of the ICAR-National Dairy Research Institute (Deemed University), Karnal (Haryana), India, is a bonafide research work carried out by her under my supervision and guidance and no part of the thesis has been submitted for any other degree or diploma.

Dr. SUNITA MEENA
(Major Advisor)

Dated:


Dedicated
To
Beloved Parents,
&
My Sweet Sister

Acknowledgements

First of all, I thank the Almighty who was always there to help me to sustain the endurance during my present investigation by providing me such a good opportunity and making me to contact good & right people in the right place.

I express my esteemed sense of gratitude to my advisory committee member ***Dr. Sachinandan De****, Principal Scientist, Animal biotechnology Division, NDRI for his guidance. This piece of work is the result of his astute guidance, close supervision, keen interest, critical appraisal, judicious planning of the project, calm, understanding and providing me with invaluable access to all facilities required which not only inspired me to bring this project to a successful completion but also left certain experience which will guide me throughout my life.*

I wish to thank my major advisor, ***Dr. Sunita Meena****, Scientist, Animal Biochemistry Division, NDRI. I could not have imagined having a better advisor and mentor for my M.V.Sc. I am thankful for her inspiration, encouragement, care and support throughout the work.*

I express my sincere sense of gratitude to the members of my advisory committee, ***Dr.(Mrs) Suman Kapila****, Principal Scientist, Animal biochemistry Division;* ***Dr.Suneel Kumar Onteru****, Senior Scientist, Animal Biochemistry Division and* ***Dr.I.D.Gupta****, Principal Scientist, Animal Genetics and Breeding Division, for their valuable suggestions, critical evaluation and encouragement throughout the research period.*

. I am also thankful to ***Dr. R.R.B. Singh****, Director & Vice Chancellor, N.D.R.I. for providing me the necessary facilities required for my study and research and keeping my financial worries at bay in the form of ICAR-JRF and N.D.R.I.-Institutional fellowship.*

I am thankful to ***Dr. Dheer Singh****, Head of the Division, Animal Biochemistry, for providing infrastructural facilities and office staff for helping in smooth flow through official procedures.*

Thanks to all faculty members of Animal Biochemistry Division for their co-operation and support.

I would like to thank ***Dr. Amar Singh Meena****, Scientist, Animal Genetics and breeding Division, CSWRI, Avilanagar, for his invaluable help which helped me in the completion of my research work*

I wish to record my heartfelt indeptedness to ***Devika Gautam*** *mam for her unconditional help, inspiring and benevolent guidance, care and sustained support throughout the research work. Thankyou so much mam.*

I am wholeheartedly thankful to ***Ashutosh*** *sir and* ***Murtaza Ali*** *sir for their valuable assistance, productive discussions and constant support that helped me in the completion of my research work.*

I express my sincere thanks to my labmates, ***Mayank Roshan sir, Jatinder sir, Sacchidananda bera and Parmanand*** *for their unconditional help that enabled me to complete my research work successfully.*

My sincere and very special thanks to ***Sunil,*** *whose incredible ideas and care has contributed a lot for successful completion of my research work.*

I thank my dear friends ***Loren, Madhu, Prachurya, Lopu and Elizabeth*** *for making the burden of my research work lighter by constant help throughtout my research period.*

I thank my fellow batchmates ***Dipanjan, Sravanthi, Rosalin, Amrita, Sikha, Nancy, Erica, Divya, Vaishali, Yuvraj, Javed, Prasanth and Akash*** *for always being there for each other.*

I express my thanks to lab assistants, ***Rohit*** *and* ***Prem ji,*** *for their help throughout the research work.*

All languages fall short of to express my undoubtfull gratitude, indebtedness, love and affection to my beloved Family. The blessings of my grandmothers ***S.Yasodha*** *and* ***S.Suryakantham*** *and the love and care of my father* ***S.Venugopal Rao,*** *mother* ***S.Gnaneshwari*** *and sister* ***S.Harika*** *who constantly educated, guided and moulded me into the present position and whose boundless love, unparalleled affection, dedicated efforts, encouragement and moral support is a constant source of motivation for me in shaping up my career.*

End is inevitable for any kind of work. Though acknowledging is an endless task, I end saying infinite thanks to all those whom I am able to recall here and also to those whom I might have left unknowingly.

Date: *(Sameni Deepika)*

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND SYMBOLS

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
bp	Base pair
PBS	Phosphate Buffer Saline
SSCP	Single strand conformation polymorphism
PAGE	Polyacrylamide Gel Electrophoresis
TAE	Tris Acetate EDTA
TBE	Tris Borate EDTA
min	Minute
sec	Second
°C	Degree Celsius
NBAGR	National Bureau of Animal Genetic Resources
FAO	Food and Agriculture Organization
nt	Nucleotide
µL	Micro liter
mL	Milli liter
g	Gram
xg	Relative centrifugal force
rpm	Revolutions per minute
cDNA	Complementary deoxyribonucleic acid
Fig	Figure
CDS	Complete coding sequence
%	Percent

ABSTRACT

India has ranked second in goat population and first in goat milk production. Goat milk is highly nutritious and mostly consumed by rural people. Protein content on average is 3.4% where casein constitutes 80% of whole protein in milk. In goat milk casein protein include αs1, αs2, β and κ casein variants which are encoded as *CSN1S1*, *CSN2*, *CSN1S2* and *CSN3* casein gene which is clustered from 250 kb to 350 kb on chromosome no. 6. Caprine casein genes exhibit an unusual and extensive polymorphism that affects milk quality and composition. There are 17, 8, 12 and 23 alleles of *CSN1S1*, *CSN1S2*, *CSN2* and *CSN3* gene respectively. The *CSN1S2* gene has 8 alleles such as A, B, C, D, E, F, 0 and G with three different level of yield of αs2 casein protein. A, B, C, E and F alleles are associate with normal level of αs2-casein protein, D allele is defective allele and 0 allele is null allele responsible for absence of αs2-casein protein. The G allele was not studied at genetic level. Indian goat breeds might have different types of casein protein having different biological functions. To our knowledge the complete coding region of *CSN1S2* gene in Indian goat breeds is not been reported. The present study was initiated with the objective of cloning of complete *CSN1S2* gene in five Indian goat breeds where complete coding region of that gene was completed in six Indian goat breeds (Sirohi, Marwari, Jakhrana, Bakarwali, Osmanabadi and Chegu) and one non-descript goat of Jammu. Three reported alleles (A, B and C) were found among these breeds. Along with reported alleles seven new alleles were also found which show substitution at nucleotide and amino acid level. Allelic variations were found in 4^{th}, 5^{th}, 11^{th}, 15^{th} and 16^{th} exons of *CSN1S2* gene. The percentage of variation was higher in 16^{th} exon of Indian goat breeds. The second objective was to screen the possible sequence variation in *CSN1S2* gene in which exon 11 and exon 16 regions were included. The primers set for exon 11 were designed in such a way that if D allele is present then product size will be decreased accordingly as 106 nt deletion is present in D allele. A total 140 goats of eleven different breeds were screened but no deletion was found in exon 11. To screen the sequence variation at exon 16, primers were designed to amplify the exon 16 from genomic DNA. SSCP was performed in 10% non-denaturing acrylamide gel. Out of 60 animals screened two animals showed a different band pattern in which one of them belongs to Bakarwali and the other belongs to Marwari breed of goat. On sequencing in these two animals a novel allele was found due to variation in sequences in exon 16.

सारांश

भारत बकरी की आबादी में दूसरे और बकरी के दूध उत्पादन में प्रथम स्थान पर है। बकरी का दूध अत्यधिक पौष्टिक होता है और ज्यादातर ग्रामीण लोग इसका सेवन करते हैं। दूध में प्रोटीन की मात्रा औसतन 3.4% होती है जिसमें 80% केसीन प्रोटीन होता है। बकरी के दूध के αs1, αs2, β और k केसीन प्रोटीन *CSN1S1*, *CSN2*, *CSN1S2* और *CSN3* केसीन जीन से बनता है जो की क्रोमोसोम 6 के 250 से 350 के. बी. के क्षेत्र में उपस्थित है । बकरी के जीन्स की असामान्य और व्यापक बहुरूपता दूध की गुणवत्ता और संरचना को प्रभावित करती है। क्रमशः *CSN1S1*, *CSN1S2*, *CSN2* और *CSN3* जीन के 17, 8, 12 और 23 प्रकार मौजूद हैं। *CSN1S2* जीन में ए , बी, सी, डी, ई, एफ, 0 और जी G जैसे 8 प्रकार हैं, जो की αs2 केसीन प्रोटीन की तीन अलग-अलग स्तर के उत्पाद की ज़िम्मेदार हैं। ए, बी, सी, ई और एफ एलील αs2 केसीन प्रोटीन के सामान्य स्तर के साथ संबद्ध हैं, डी एलील दोषपूर्ण एलील है और 0 एलील अशक्त एलील है जो αs2-कैसिइन प्रोटीन की अनुपस्थिति के लिए जिम्मेदार है। जी एलील का आनुवंशिक स्तर पर अध्ययन नहीं किया गया है। भारतीय बकरी की नस्लों में केसीन प्रोटीन के विभिन्न प्रकार के कारण विभिन्न जैविक कार्य हो सकते हैं। अब तक भारतीय बकरी की नस्लों में *CSN1S2* जीन के पूर्ण कोडिंग क्षेत्र की जानकारी नहीं है। वर्तमान अध्ययन, पांच भारतीय बकरी की नस्लों में पूर्ण *CSN1S2* जीन के क्लोनिंग के उद्देश्य से शुरू किया गया था, लेकिन जीन का पूर्ण कोडिंग क्षेत्र छह भारतीय बकरी की नस्लों (सिरोही, मारवाड़ी, जखराना, बकरवाली, उस्मानाबादी और चेगू) और एक जम्मू की गैर विवरण नस्ल में पूरा किया गया। इन नस्लों में ए, बी और सी एलील पाए गए। रिपोर्ट किए गए एलील्स के साथ-साथ सात नए एलील भी पाए गए जो न्यूक्लियोटाइड और एमिनो एसिड स्तर पर प्रतिस्थापन दर्शाते हैं। 4, 5, 11, 15 और 16 एक्सॉन में भिन्नताएं पाई गईं। भारतीय बकरी की नस्लों के 16वें एक्सॉन में सबसे अधिक भिन्नतायें पाई गयी। दूसरा उद्देश्य में *CSN1S2* जीन में स्थित 11वें और 16वें एक्सॉन में संभावित अनुक्रम भिन्नता को प्रदर्शित करना था। एक्सॉन 11 के लिए प्राइमरों को इस तरह से डिजाइन किया गया था कि अगर डी एलील मौजूद है तो उत्पाद का आकार उसी हिसाब से कम हो जाएगा क्योंकि डी एलील में 106 nt विलोपन मौजूद है। ग्यारह अलग-अलग नस्लों की कुल 140 बकरियों की जांच की गई, लेकिन एक्सॉन 11 में कोई विलोपन नहीं पाया गया। एक्सॉन 16 में अनुक्रम भिन्नता को जानने के लिए, प्राइमर को जीनोमिक डीएनए से एक्सॉन 16 को परिवर्धित करने के लिए बनाया गया। एस.एस.सी.पी. को 10% नॉन-डेनेट्यूरिंग एक्रिलामाइड जेल में किया गया। 60 जानवरों में से दो जानवरों ने अलग बैंड पैटर्न दिखाया, जिसमें से एक बकरवाली का है और दूसरा मारवाड़ी नस्ल का है। अनुक्रमण किए जाने पर इन दो जानवरों में 16 वें एक्सॉन में भिन्नता के कारण एक नया एलील पाया गया था।

CHAPTER –1

Introduction

INTRODUCTION

Goat was the first farm animal to be domesticated around 8000 B.C. (Hatziminaoglou and Boyazoglu, 2004). In developing countries, goats were mostly reared by the poor farmers in rural areas for production of milk and meat. They make a valuable contribution to the poor due to their low management expenditure, greater adaptability and small size (Aziz, 2010). The total world goat population is 1034.4 million where Asia ranks first in world goat population of around 50.47% (FAOSTAT 2017). Indian goat population is 135.17 million and India ranks first in goat milk production among all the countries in the world (19th Livestock census 2012). Goat milk is highly nutritious containing various macro and micro nutrient, along with bio active compounds which provide health protection. Goat milk with lower concentration of αs1 casein is hypoallergenic as compared to cow milk and is suitable for infants (Potocnik *et al*., 2011).

The major milk protein is casein comprising around 80% of the total milk protein which is insoluble and rest 20% is whey protein that is soluble. Casein genes are the members of secretory calcium binding phosphoprotein which is composed of group of acidic, proline rich, phosphoprotein present as a micelle structure and calcium aiding in colloidal suspension of milk (Kawasaki *et al*., 2011 and Rijnkels, 2002). The casein gene is clustered in 250 kb genomic region in chromosome number 6 of goat in the order of *CSN1S1, CSN2, CSN1S2* and *CSN3.* The *CSN1S1, CSN2, CSN1S2* and *CSN3* code for 199, 207, 208 and 192 amino acids respectively. Polymorphism in casein gene leads to difference in the quality and the quantity of casein protein (Marletta *et al*., 2007). Genetic polymorphism leads to diversity and is inherited from generation to generation. Variations in gene sequence occurred due to transition and transversion of nucleotides which results into different protein types. Although casein genes are not polymorphic, till date in goat, 17, 12, 8 and 23 alleles of *CSN1S1, CSN2, CSN1S2* and *CSN3* genes respectively are reported. Polymorphism of *CSN1S1* and *CSN3* was studied in Indian caprine and bovine breeds (Gautam *et al*., 2018; Gupta *et al.,* 2009; Rachagani and Gupta, 2008). There is no study about screening of variation and polymorphism in coding region of *CSN1S2* gene in Indian goat breeds.

CSN1S2 gene is 18.5 kb long which is composed of eighteen exons, ranging from 21 to 266 nucleotides, where 1st, 17th and 18th exons are noncoding (Marletta *et al*., 2007). The open reading frame is 669 nucleotides flanked by 5' and 3'untranslated regions of 42 and 302 nucleotides, respectively (Bouniol, 1993). The eight alleles are reported (A, B, C, D, E, F, 0 and G). Except allele G, all other alleles are studied both at nucleotide and amino acid level. These alleles are associated with αs2 casein concentration where allele A, B, C, E and F can produce around 2.5 g/L of milk. Due to the deletion of 106 nucleotides in allele D the casein concentration reduced to 1.5g/L and non-detectable amount of αs2 casein is found in null (0) allele. In Indian goat breeds (Jamunapari, Marwari, Sirohi, Beetal, Ganjam and local goats) it was observed that among A, B and 0 alleles A allele is found with highest frequency whereas the Ganjam goats were associated with higher 0 allele (Rout *et al*., 2010 and Sharma *et al*., 2017). There was no information about *CSN1S2* gene regarding full length sequences of coding region. Thus the present work was initiated with following two objectives:

1. **Cloning of complete *CSN1S2* gene in five Indian goat breeds**
2. **Screening of possible sequence variation in *CSN1S2* gene in five Indian goat breeds**

CHAPTER –2

Review of Literature

2.1 Goat husbandry in India

Goat was the first livestock to be domesticated around 8000 B.C. and was found in the region of Ganj Darech, Kermanshah valley, Iran which is also the homeland of one wild goat species (Hatziminaoglou and Boyazoglu, 2004). Goat is expected to be first milking animal exploited by humans (Scott, 1986). These animals had a great impact on livelihood of ancient people during the first known civilization in the area of Mesopotamia (Middle East). Goats have great adaptability capacity because of which they are sustained in different environmental conditions and disseminated all over the world. They are multipurpose animals and useful to man throughout ages due to their productivity, small size, and non-competiveness (Aziz, 2010). The world total goat population is 1034.4 million. The largest population of goats was found in Asian continent accounting 50.47% of world total population followed by Africa having 40.86% (Table 2.1). Asia ranks first in goat milk and meat production (FAOSTAT 2017).

Table 2.1: The number of goats, goat's milk and meat production in different continents of the world (FAOSTAT 2017).

Region	Number (millions) of Goats	Goat milk production (million tonnes)	Goat meat production (million tonnes)
Asia	522.1	10.6	4.1
Africa	422.7	4.5	1.4
Northern America	2.6	0.03	0.008
Central America	8.9	0.16	0.040
South America	21.4	0.31	0.07
Caribbean	4.0	0.25	0.01
Europe	19.2	2.9	0.11
Oceania	4.0	0.000044	0.03
World	1034.4	18.65	5.9

In India, the population of indigenous breeds along with non-descript goats is 135.17 million which account 26.40% of total livestock population in the country and it is 13.06% of whole world goat population. About 26.97% are pure breed, 11.77% are graded breed and remaining 61.26% are non-descript breeds (19th Livestock census 2012). There are 34 indigenous breeds besides its non-descript animals in India registered under NBAGR (Breed registration committee meeting, 2018, http://www.nbagr.res.in/registeredbreed.html). There are 351 breeds in the world of which 23 are Indian breeds which were well recognized and studied (Mandal *et al*., 2014). These breeds are defined according to their geographical position, morphological characteristics and production performance. Indian goats are distributed in four geographical regions namely temperate Himalayan region, northwestern region, southern peninsular region and eastern region (Table 2.2) They are divided into large, medium and small breeds according to body size and also classified as milk, meat and dual type based on production performance. These breeds show variation in fecundity; production like milk, meat and fiber; disease resistance; drought-ability and heat tolerance (Acharya *et al*., 1982).

In developing countries, goats contribute a significant role in livelihood especially of the poor in rural regions. Goats are mainly reared by small and medium farmers as cost of maintaining goats is very less (Aziz, 2010). Many breeding programs were established but were followed in developed region but not able to reach the rural areas properly. Crossbreeding and inter-mixing the breeds with local breeds is threat to Indian goat population. In 2014, Mandal and co-authors had mentioned that threatened breeds are Jamunapari, Beetal, Jakhrana and Surti in northwestern reion, Sangamneri, Osmanabadi, Malabari and Attappady Black in the southern region, Ganjam in eastern region, and Chegu and Changthangi in the temperate Himalayan region. The conservative measure like in-situ and ex-situ conservation methods had been adopted to maintain and increase the population.

2.2 Importance of goat milk

India ranks first in goat milk production in the world (Table 2.3) producing 6.2 million tonnes of goats' milk from 36.13 million milk producing animals (FAOSTAT 2017). Goat milk and the dairy products of different ruminants is a source of major and minor nutrients which play a significant role in human health and nutrition (Ceballos *et al*., 2009). Due to its specific composition, it is used as raw material for manufacturing

food for infants and elderly population in specific condition (Park, 2017). The composition of milk varies depending on the stage of lactation, feeding, breed and environmental condition. The total solid content of goat milk (Granadian goats) is higher in comparison to cow milk (Holstein Friesian cows). αs1 casein protein content of goat milk is very less when compared to the cow milk and the digestibility is good in case of goat milk than that of cow (Haenlein, 2004; Ceballos *et al*., 2009).

Table 2.2: Distribution of thirty four Goat Breeds in different regions of India (NBAGR)

Temperate Himalayan region	Northwestern region	Southern peninsular region	Eastern region
Gaddi	Jamunapari	Sangamneri	Ganjam
Changthangi	Marwari	Osmanabadi	Bengal Goat
Chegu	Zalawadi	Kannai Adu	Assam Hill
Bhakarwali	Beetal	Malabai	Sumi-Ne
Pantja	Kutchi	Attappady	
	Sirohi	Konkan Kanyal	
	Barbari	Berari	
	Mehsana	Bidri	
	Surti	Nandidurga	
	Jakhrana	Teressa	
	Gohilwadi	Salem Black	
	Kahmi	Kodi Adu	
	Rohilkhandi		

The saturated medium chain fatty acids are higher in goat milk because of which milk fat is easily digestible when hydrolyzed by action of salivary pregastric lipase (Ceballos *et al*., 2009). Thus goat milk is excellent source of energy, easy to digest and equally has fast oxidative metabolism (Leyton *et al*., 1987; Aurousseau *et al*., 1989; Vela´zquez *et al*., 1996; Matsuo and Takeuchi, 2004). n-6 and n-3 PUFA are higher in

content but ratio of n-6:n-3 was lesser in case of goat milk comparing to cow milk. The mineral composition of Fe, Mb, Na and Zn in goat milk is very similar to that of cow milk however there is higher amount of Ca, K, Cl and Mn in goat milk (Haenlien, 2004). Despite similar content of Fe which is naturally low in milk, the bioavailability of it is high in goat milk than that of bovine milk. Goat milk has a higher share of nucleotides that contribute to an enhanced Fe absorption in the intestine (Claeys *et al*., 2014; Bar1owska *et al*., 2011; Park, 2017).

Table 2.3: List of top ten goat milk producing countries of the world, quantity of milk they are producing and number of producing animals present in respective countries (FAOSTAT 2017)

Country	Milk production (Million tonnes)	Number (millions) of producing animals
India	6.2	36.13
Bangladesh	1.11	30.11
Sudan	1.10	18.95
Pakistan	0.84	8.4
France	0.60	0.84
Greece	0.56	4.5
Turkey	0.52	4.96
Spain	0.50	1.26
Mali	0.40	18.01
Somalia	0.37	6.24

Kefir is a product of milk which was prepared with cow and goat milk, had almost same titratable acidity of nearly 1% lactic acid; lactose content of cow milk was lower than that of goat milk and in case of percentage of lactose hydrolysis, the cow and goat milk showed 21.4 and 22.4 % of hydrolysis respectively. The proteolysis accounted for 7.2 and 7.5% and fat hydrolysis was 6.5 and 4.2% fat hydrolysis in case of cow milk kefir and goat milk kefir respectively (Manthani *et al*., 2018). The predominant factor for sensitization to milk allergy was due to presence of abundant amount of αs1-casein in

ruminant casein except in some goats similar to that of human and equine casein (Bar1owska *et al*., 2011; Malacarne *et al*., 2002; Potocnik *et al*., 2011).

Table 2.4: Composition of goat and cow milk. (Jenness (1980); Strzalkowska *et* al., 2010; Ceballos *et al*., 2009; Park, 2017)

Components (%)	**Goat**	**Cow**
Protein	2.3-4.4	2.5-4.1
Lactose	4.0-5.1	4.4-5.3
Calcium	0.085-0.198	0.125
Water	86-87	86-87
Cholesterol	0.0169-0.018	0.025-0.038
Fat	3.0-6.0	3.2-5.34
(% of total fatty acids)		
C4:0	1.27	3.84-4.43
C6:0	3.28	2.28-2.31
C8:0	3.68-3.94	1.69-1.72
C10:0	6.54-11.07	3.36-3.42
C12:0	4.45-7.64	3.83
C14:0	9.92-11.92	11.24-12.13
C16:0	25.64-26.4	32.24-33.35
C18:0	6.66-9.92	9.97-12.06
C18:1	18.76-24.17	19.41-23.35
C18:2	2.04-2.72	2.41
C18:3	0.53	0.25-0.42
SFA	70.42	69.38-71.24
UFA	28.8	24.1-30.62
MUFA	24.46-25.67	25.56-27.73
PUFA	4.08-4.67	2.89-3.20

2.3 Goat casein and its biological value

Goat milk contains major and minor proteins. There are two major categories of milk proteins: insoluble proteins (the casein family) and soluble proteins (whey proteins). Caseins protein includes αs1, αs2, β, and κ-caseins, while the whey proteins are α-lactalbumin and β-lactoglobulin. αs1, αs2 and β casein are calcium sensitive protein having greater solubility in the presence of calcium whereas k is not. Milk also comprises important minor proteins, such as serum albumin, immunoglobulins, lactoferrin, transferrin, calcium-binding protein, prolactin and folate-binding protein. The non-protein nitrogen component is 5.8% of overall nitrogen (Park, 2017; Prosser *et al*., 2008). Caseins are the major protein fraction of milk and their function is to transport calcium phosphate in milk and provide the newborns with a source of calcium and phosphorus for bone formation as well as to contribute to the requirement for amino acids (Stewart *et al*., 1987).

Goat milk proteins have many biological activities due to the presence of bioactive peptides released from parent protein by digestive enzymes. Antioxidant peptides can prevent oxidative deterioration of foods and prolong the shelf life. Ahmed *et al*., 2015 reported that the hydrolysates of fractionated whey and casein proteins from goat milk exhibited potent superoxide anion ($O_2^{\cdot -}$) scavenging activity as investigated using the natural xanthine/xanthine oxidase system. Several antioxidant peptides were derived from both caseins and whey proteins in which β-casein and β-lactoglobulin plays major role in antioxidant activity, respectively (Ahmed *et al*., 2015). The milk of Prisca goat breeds were having highest Total Antioxidant Capacity compared to other species such as donkey and cow (Simos *et al*., 2011). Goat milk contains inhibitory peptides for Angiotensin-converting enzyme belong to β-lactoglobulin, β-casein and κ-casein protein and hence, goat milk can be used as natural inhibitors of hypertension (Ibrahim *et al*., 2017). Hypoallergic property of goat milk was observed when they digested with alcalase to produce goat milk protein hydrolyzates, which decrease the level of histamine and TNF-α from HMC-1 cells (human mast cell line). Therefore, the GMPH can be used to replace the cow milk in case allergenicity and to develop foods for infants and convalescent (Bia *et al*.,2015 and Jung *et al*., 2016).

2.4 Genetic Characterization of Indian goat germplasm

Dixit *et al*., 2012 studied genetic diversity and relationship between Indian goat breeds by using 25 microsatellites markers where they observed a total of 412 alleles and they concluded that Kanniadu and Sirohi breeds are most diverse breeds followed by Osmanabadi being the least diverse breed. The genetic distance tended to be lowest between Jamunapari and Marwari and widest between Kanniadu and Zalawadi. Most of the loci in almost all breeds were homozygous. Three major mtDNA (mitochondrial DNA) lineages were identified in all continent and the most common was lineage A in all continents, lineage B was dominant in Indian subcontinent, Mongolia, and Southeast and lineage C was found in few goats from Mongolia, Switzerland, and Slovenia (Luikart *et al*., 2001). Joshi *et al*., (2004) found C lineage was also present in Indian goats and two other lineages D & E in Barbari breed. There are 34 registered goat breeds in India which are known for their utility (Meat/Milk/fiber). Most of the breeds were reared mainly for meat in addition to their other own specific performance like milk yield from Beetal, Jakhrana, Jamunapari, Surti and Bakarwali, pashmina from Chegu, Gaddi, Changthangi and fecundity from Black Bengal (Verma, 2018).

2.5 Genetic polymorphism and mutation detection by PCR-SSCP method

A polymorphism is an allelic variation in the genetic code of a particular gene sequence leading to change in the protein sequence in case of non-silent variation at genetic level. At least two different forms or variants of a gene sequence exist in a single population. Genetic polymorphism leads to diversity and inherited generation to generation as no single form has fully advantage or disadvantage over the others in terms of natural selection Variation in gene sequence occurred due to transition and transversion of nucleotides which result into SNPs, sequence repeats, insertions, deletions and recombination. Genetic polymorphisms may be the result of chance processes or may have been induced by external agents (Ford, 1940).

SSCP (Single Strand Conformation Polymorphism) is a technique which relies on the principle of electrophoretic mobility of single stranded DNA molecules in a non-denaturing polyacrylamide gel which is highly dependent on its molecular size and structure. They are widely used in biomedical research, where the separation of DNA molecules differ by one or more nucleotides on the basis of chemical or physical properties rather than by size alone (Gasser *et al*., 2006). Orita *et al*., 1989, developed

this method to analyze the single strand DNAs on neutral polyacrylamide gel electrophoresis to detect DNA polymorphisms and found that there was mobility shift by single base substitution. Single-stranded conformation polymorphisms (SSCPs) result from mutations detectable by conventional restriction fragment length polymorphism (RFLP) analysis as well as single nucleotide substitution that RFLP analysis would not detect. From observations of 30 unknown mutations, Hayashi (1991) estimated PCR-SSCP analyses sensitivity (probability of detecting at least one strand shifted) of more than 99% for 100 to 300 bp fragments and 89% for 300 to 450 bp fragments. In addition to sensitivity, the SSCP technique is technically simpler than mismatch technique such as RNase digestion chemical cleavage, GC clamping and temperature gradient electrophoresis (Keen and Inglehearn, 1996; Michaud *et al*., 1992).

2.6 Structure of Casein gene cluster

The casein genes are tightly linked and clustered on DNA segment of about 250 kb mapped on chromosome no. 6 in cattle and goat (Fig. 2.1) in the following order: CSN1S1 (coding for s1-casein), CSN2 (β-casein), CSN1S2 (s2-casein) and CSN3 (k-casein) (Rijnkels, 2002).

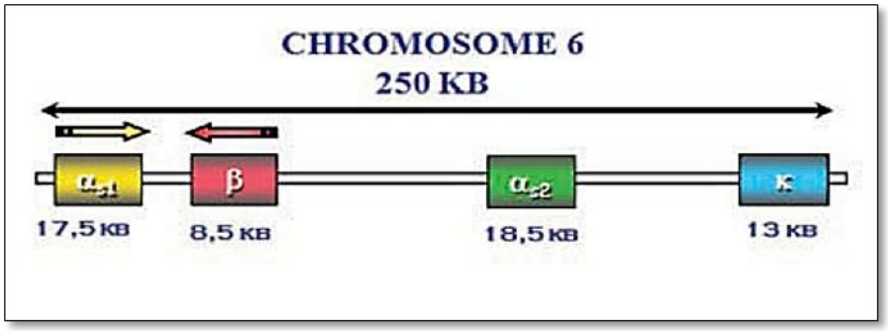


Fig. 2.1 Genomic organisation of the bovine/goat casein locus (Caroli *et al*., 2009)

The *CSN1S1* and *CSN2* genes are 12 kb apart and convergently transcribed (Leroux and Martin, 1996).The *CSN1S1, CSN2, CSN1S2* and *CSN3* code for 199, 207, 208 and 192 amino acids respectively. Phosphorylation of protein was present at serine residues, high in αs2 casein followed by αs1, β and k casein (Martin *et al*., 2002). The *CSN1S1, CSN2* and *CSN1S2* genes are evolutionary related to each other and share common 5' flanking regulatory motifs, though the organization at 5' flanking region of *CSN3* gene is different but the expression pattern is similar (Groenen and Poel, 1994; Coll *et al*., 1995; Martin *et al*., 2002). The genomic organization of casein gene is

conserved in mammals despite of some difference in distance separating the casein genes.

2.7 Genetic polymorphism of casein gene cluster

Polymorphism in casein gene leads to difference in the quality of protein and the quantity of casein protein varies due to difference in gene expression (Marletta *et al*., 2007). Domestic goats (Capra hircus) from Europe, Africa, the Far East, and the Near East and wild goat (Capra aegagrus) populations shared a substantial number of casein SNP, from 36.1% (CSN2) to 55.1% (*CSN1S2*). Around 50% of the casein SNP were shared by 2 or more populations, and 18 to 44% were shared by all populations when casein variations are compared among bezoars and four domestic populations (Gaun *et al*., 2019). *CSN1S1* and *CSN3* genes are highly polymorphic compared to *CSN1S2* and *CSN2* gene (Caravaca *et al*., 2009).

The goat *CSN1S1* gene and its promoter comprise of 16785 bp (base pairs) including 1138 bp of exonic region and 15647 bp of intronic region. The gene contains 19 exons (Ramunno *et al.,* 2005). There were 17 allelic variants formed due to transition, transversion, deletion and insertion of nucleotide sequences in the coding region of *CSN1S1* gene associated with difference in both qualitative and quantitative differences. They have four different types of efficiency in synthesis. The alleles A, B1, B2, B3, B4, C, H, L and M are responsible for a high amount of αs1-casein (~3.5 g/L per allele), the alleles E and I with an intermediate amount (~1.1 g/L per allele), and the alleles D, F and G with a low amount (~0.45 g/L per allele) (Vacca *et al*., 2009). The alleles 01, 02 and N are the null alleles (Bevilacqua *et al*., 2002). Therefore, milk produced by goats with different *CSN1S1* genotypes shows a variable amount of αs1-casein, ranging from 7 g/L in strong allele homozygous goats, to 0.9 g/L and 0 g/L in weak and null homozygotes, respectively. The full characterization of *CSN1S1* F and N alleles was done at genomic DNA level and observed 118 polymorphic sites in N, F and A *CSN1S1* alleles along with a deletion of the cytosine at the 23rd nucleotide of the 9th exon of N and F (Ramunno *et al*., 2005).

There are twelve genetic variants of *CSN2* been reported (A1, A2, A3, B, C, D, E, F, H1, H2, I, G) in cattle (Caroli *et al*., 2006) and at least nine allele (A, A1, B, C, C1, D, E, and 0, and 01) are present in goats (Caroli *et al*., 2006; Chessa *et al*., 2008, Vacca

et al., 2009). The polymorphism study of buffalo β-casein gene showed two major alleles for coding region and four major variants of promoter region by PCR-SSCP analysis.

In case of *CSN1S2* gene till date eight alleles have been identified which are responsible for three different expression level αs2 casein : A, B (Boulanger *et al*., 1984), C (Bouniol *et al*., 1994), E (Lagonigro *et al*., 2001) and F (Ramunno *et al*., 2001(b)) alleles are associated with normal content of protein about 2.5g/L whereas D allele associated with reduced content of protein 1.5g/L (Ramunno *et al*., 2001(b)) and 0 allele (Ramunno *et al*., 2001(a)) associated with non-detectable amount of αs2 casein in goat milk. G allele has been identified at protein level by isoelectric focusing (IEF) associated with normal content of αs2 casein (Erhardt *et al.*, 2002).

The open reading frame of *CSN3* gene is of 579 bp coding for 21 amino acids of signal peptide and 171 amino acids of mature protein (Gautam *et al*., 2019). It was found that k casein gene has total twenty three variants at protein level named from A to W (Vacca *et al*., 2009; Gautam *et al*., 2019) and thirty seven variants at DNA level. Simultaneous discrimination of CSN3 alleles A B C D and E was obtained by PCR-SSCP analysis (Chessa *et al*., 2003). Recently, Gautam and coauthors found new alleles (P, Q, R, S, T, U, V and W) by PCR-SSCP analysis of amplified region of 437-bp fragment of exon 4 in the goat *CSN3* gene containing amino acid residues 23–145 of Indian goat breeds. Allele A is prevalent in Indian breeds like Barbari, Beetal and Surti, whereas allele B was found more frequent in Barbari. The genotype having AB and BB are associated with high level of k casein in goat milk than with AA genotype (Caravaca *et al*., 2009).

2.8 Exon skipping

Alternative splicing or differential splicing is a post transcriptional regulatory process where gene expression of a gene leads to multiple proteins (Stamm *et al*., 2005) This is a defective processing effecting of primary transcripts. There are two process which leads to decrease in peptide size (i) usage of cryptic splicing site induced by error in the junction sequence leading to a sort of slippage of the spliceosome and (ii) invoked by defect in the consensus sequences, either at the 5’ and/or 3’ splice junctions or at the branch point, or both (Martin *et al*., 2002). F and D allelic variants were found in *CSN1S1* gene of goat due to mutation away from the splice site recognition like deletion in exon 9 and two insertion in downstream intron (3 and 11 bp in length) involved in the

base pairing interaction with intron 5' splice site which might consequently be less accessible to U1 snRNA (Leroux *et al*., 1992). Casual exon skipping is responsible for missing of amino acids which changes the net charge showing different isoelectric focusing pattern. The exon skipping is not related to the allelic variability. Boisnard *et al*., 1991 found two non-allelic forms of ovine αs2 casein proteins produced by different mRNAs which differ by deletion of 34-42 aminoacids and reported that, this occurred in all individuals analysed, indicating it as stabilized mechanism for production of multiple proteins. In Equidae family, due to alternative splicing, there are two isoforms of *CSN1S2* gene, one with 221 amino acids and the other smaller with 168 amino acids (Cieslak *et al*., 2016). The structural variation may influence the allergenocity (Ballabio *et al*., 2011). Matéos *et al*., 2009 reported that the equine *CSN1S1* protein had isoforms like full length protein, proteins with exon 7 and 14 missings and variants having missed both 7 and 14 exons. All these isoforms lacked Gln residue coded by CAG codon present at the beginning of exon 11 which was spliced by cryptic splice site mechanism of alternative splicing (Smith *et al*., 1993). The deletion of CAG codon of 11th exon is common to other species including ewe, goat, cow and water buffalo (Matéos *et al*., 2009). Exon skipping is common in *CSN1S1* and *CSN1S2* transcripts due to presence of many small size exons. Ramunno *et al*., 2001(a) found exon skipping of 6th and 11th exon of *CSN1S2* transcripts in goat which were homologous to null allele.

2.9 Structure of *CSN1S2* gene

The *CSN1S2* gene is 18.5 kb long which is composed of eighteen exons, ranging from 21 to 266 nucleotides (Fig. 2.2), where 1st, 17th and 18th exons are noncoding (Marletta *et al*., 2007). The open reading frame is 669 nt flanked by 5' and 3'untranslated regions of 42 and 302 nt, respectively (Bouniol, 1993).

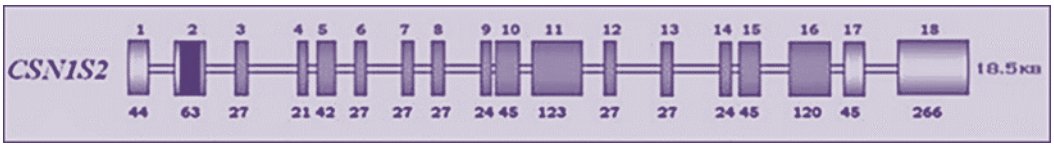


Fig. 2.2 Structural organization of *CSN1S2* gene with exon number and its size

2.10 Polymorphism of *CSN1S2* gene

Boulanger *et al*., 1984 reported B allele by acid gel electrophoresis at pH 3 using milk samples of Alpine and Saanen goats. It differ from the A allele by single nucleotide polymorphism at 10th nt of 9th exon where a G to A transition occurred resulting in

change in amino acid, glutamine to lysine at 64th position of mature protein. C allele was identified by phenotyping by isoelectric focusing which has slightly more acidic isoelectric point and genotyping by cloning and sequencing using milk of French dairy breeds, Alpine and Saanen. There is transversion of 545th nt, A to T affecting 167th codon where lysine is substituted by isoleusine (Bouniol *et al*., 1994). A non-detectable amount of αs2 casein was observed by Ramunno *et al.*, 2001(a) which is due to presence of *0* allele. They performed SDS-PAGE analyses of 441 individual skimmed milk samples of goats of Southern Italy and then performed RP-HPLC of samples that are having A/A, A/0 and 0/0 genotype. By sequencing the cDNA synthesized from RNA of animal having 0/0 allele revealed a G to A transition at 80th nt of 11th exon leading to a stop codon. This mutation has eliminated a NcoI restriction site. F and D allele were identified as novel alleles by Ramunno *et al*., 2001(b) in goats reared in province of Naples. They found that F allele was characterized by a G to A transition by sequence analyses at 13th nt of 3rd exon changing codon of valine to codon of isoleucine (both are apolar amino acid) at 7th position of mature protein. This mutation affects ALW26I restriction site which act as a marker for PCR-RFLP. D allele was associated with decreased αs2 casein where there is deletion of 106 nt spanning the last 11 nt of 11th exon and 95 nt of next intron. This deletion resulted in missing of 122, 123 and 124 amino acid of mature protein and had Asn amino acid instead of Thr at 121 position. Transversion of 83rd base from C to G at 16th exon was identified as E allelic variant by screening 279 goats of different breeds of Italy (Lagonigro *et al*., 2001) where proline was replaced by arginine at 197th position of mature protein. This mutation alters the restriction site for NlaIII which was used for identification of carrier of E allele by PCR-RFLP method (Lagonigro *et al*., 2001). A, B, C and E alleles were simultaneously identified by Chessa *et al*., 2008) by at that time newly approached method, PCR-SSCP where they did two PCR, one is to amplifiy a fragment of exon 9 containing mutation to distinguish allele from A, C, and E, and second PCR to amplify the fragment of exon 16 containing two mutation for C and E alleles. They validated the results by analyzing the 37 goat samples at protein and DNA level by isoelectric focusing and PCR-RFLP methods respectively. Erhardt *et al*., 2002, reported a new allele, as G based on phenotypic analysis by IEF but there was no study at genomic level and no information about position of change in nucleotide of aminoacid sequences.

A and F alleles were present in two chinese dairy goat breeds (Xinong Saanen and Guanzhong) where A allele had higher frequency than allele F and no other alleles obsereved in these breeds (Yue *et al*., 2013). In South Africa, dairy breeds (Sannen, British Alpine and Toggenburg) A, B, C and F alleles were found, where the frequency of A allele and AA genotype being high. AC genotype was high in case Sannen breed of South Africa and three unique genotypes, BB, BF and CF were also found but with lower frequency (Grobler *et al*., 2017). In Vallesana, Roccaverano, Jonica, Garganica, and Maltese breeds of Italy A, B, C, E, F and 0 alleles were found. A and C alleles were dominant and 0 as least frequent allele (Sacchi *et al*., 2005). Four variants A, B, C and G were present in some Italian goat breeds (Saanen, Camosciata, Verzaschese, Orobica, Frisa, and Bionda dell'Adamello) identified at protein level by IEF with highest frequency of A allele and AA genotype (Chiatti *et al*., 2005). On characterization of goat milk protein of Indian goat breeds (Jamnapari, Barbari and Jakhrana, Marwari, Sirohi, Beetal, Ganjam and local goats) and it was observed that A, B and 0 alleles were present in population with A allele having highest frequency except in Ganjam goats which was having higher 0 allele (Rout *et al*., 2010 and Sharma *et al*., 2017).

CHAPTER –3

Materials & Methods

MATERIALS & METHODS

3.1 Chemicals and Kits

Acetic acid, Acrylamide, agarose, ammonium persulfate (APS), diethylpyrocarbonate (DEPC), bisacrylamide, Phosphate buffered saline solution (PBS), TRIZOL, chloroform, isopropanol, ethanol, tris base, Ethylene diamine tetra acetic acid (EDTA), Sodium hydroxide (NaOH), sodium dodecyl sulfate (SDS), potassium acetate, isoamylalcohol, ethidium bromide, Boric acid, Phenol, formaldehyde, ammonium chloride, potassium carbonate, formamide, 37% formaldehyde, Silver nitrate, Sodium thiosulphate, sodium carbonate, bromophenol blue, xylene cyanol, Glacial acetic acid, methanol, Sodium Phosphate dibasic, Potassium phosphate monobasic, Hydrochloric acid (HCl), Potassium chloride (KCl), Sodium chloride (NaCl), sodium acetate, β mercaptoethanol. All the above chemicals were procured from Sigma-Aldrich, St. Louis, USA.

RevertAid First Strand cDNA synthesis kit (Thermo Fisher Scientific, USA), Macherey-Nagel kit (Clonetech, Japan), InsTAclone™ PCR Cloning Kit (Thermo Fisher Scientific, USA).

3.2 Materials and Instruments

All tips, glasswares, disposable plastic tubes (15 and 50 mL) were obtained from TARSONS products Pvt.Ltd 31, Shakespeare sarani, Kolkata, ABDOS labtech Pvt. Ltd., New Delhi and Ambion. Pipettes used were ACCUPIPETES. Reagents like DNTPs were obtained from FERMENTUS which is used in PCR reaction. Kapa Taq-polymerase enzyme used in PCR reaction was obtained from Sigma.

Incubator (Yorco Sales Pvt. Ltd., India), centrifuge (High Speed Refrigerated Centrifuge, table top, Sigma, USA), gel documentation system (GelDoc X^R Biorad, Germany), spectrophotometer ND-1000 (NANODROP, Thermo Scientific, USA), shaking incubator (Eppendorf® New Brunswik™ Innova 44®, Polymerase chain reaction machine (Veriti® Thermal Cycler, Applied Biosystems®, USA).

3.3 DNA isolation from milk by phenol-chloroform method

DNA isolation was done by phenol chloroform (Lahiri and Numberger, 1991) Milk samples of goats belonging to Gaddi breed from Mandi district of Himachal

Pradesh were obtained in sterile tubes containing formaldehyde and brought under refrigeration conditions and processed immediately or stored at -80°C till further processing.

Reagents:

a) **1x PBS (pH 7.4):** 2 g of NaCl, 0.05 g of KCl, 0.02875g $Na_2HPO_4.7H_2O$ (Sodium Phosphate dibasic) and 0.05 g of KH_2P0_4 (Potassium phosphate monobasic) were dissolved in 100 mL of millipore water. The pH was adjusted to 7.4 with 1N HCl and total volume was made to 250 mL with distilled water. This solution was sterilized by autoclave and stored at room temperature.

b) **1M Tris stock solution (pH 8.0):** 121.1 g of Tris base was dissolved in 800 mL of distilled water. The pH was adjusted to 8 with 1.0N HCl and total volume was made to 1000 mL with distilled water. This solution was sterilized by autoclave and stored at room temperature.

c) **0.5M EDTA stock solution (pH 8.0):** 186.1 g of disodium EDTA•$2H_2O$ was dissolved in 800 mL of distilled water by using magnetic stirrer. The pH was adjusted to 8 with 1N NaOH and total volume was made to 1000 mL with distilled water. This solution was sterilized by autoclave and stored at room temperature.

d) **1M NaCl stock solution:** 58.4 g of NaCl was dissolved in 800 mL of distilled water and final volume was made upto 1000 mL. This solution was autoclaved and stored at room temperature.

e) **10x TENS buffer:** 500 µL of 1M Tris stock solution, 100 µL of 0.5M EDTA stock solution, 2.5 mL of 10% SDS and 1M of NaCl stock solution were added together and final volume was made to 50 mL with distilled water so that final concentration of Tris, EDTA, NaCl were 10mM, 1mM and 100mM respectively. This solution was stored at room temperature.

f) **10% SDS:** 10g of SDS was dissolved in distilled water and the volume was made to 100 mL. This solution was stored at room temperature in a glass container.

g) **Proteinase k (10mg/mL):** 100 mg proteinase k was mixed in 10 mL of TE buffer.

h) Tris saturated Phenol:

1. Crystalized phenol was melted at 68°C for 10 to 15 minutes and 100 mL of melted phenol was transferred to a brown bottle. Then 0.1% of 8-hydroxy Quinoline was added, colour changes to yellow.
2. Tris base (0.5M), pH~10.5, was added to equal volume of phenol and kept on magnetic stirrer for several minutes. Then kept still until two phase separates. The upper phase with tris is removed. Repeat this step.
3. Then equal volume of (0.5M) Tris base, pH 8.0 until the pH of the phenol phase is >7.8
4. Add 0.2% β mercaptoethanol and mix.
5. Finally, 0.1 M Tris base (pH 8.0) is added to about 1/3rd volume of phenol and store at 4°C.

i) Chloroform:Isoamylalcohol (24:1): 24mL of chloroform was added to 1mL of Isoamylalcohol and mixed well.

j) 3M Sodium acetate: 246.1g of sodium acetate was added in 500 mL of millipore water. The pH was adjusted to 5.3 with glacial acetic acid and final volume was made to 1000 mL with millipore water and filter-sterilized.

k) 70% Ethanol: 35mL volume of 100% ethanol to 15mL volume of millipore water. It was always freshly prepared and kept at -20°C before use.

l) Tris-EDTA (TE) Buffer: 500 µL of 1M Tris stock solution and 100 µL 0.5M of EDTA were mixed together. The pH was adjusted to 8 with 1.0N HCl and final volume was made to 50mL by adding autoclaved MilliQ water. This solution was kept at room temperature.

m) 50x TAE buffer: 121g of Tris Base, 93g of EDTA (0.5M) and 28.5ml of glacial acetic acid was added and final volume made up to 500ml with MilliQ then autoclaved.

n) 1x TAE: for 500ml of 1x TAE**,** 10ml of 50x TAE is mixed with 490ml of MilliQ and used.

Procedure:

a) Milk was skimmed at 3000 rpm for 10 min at 4°C and fat was removed. This step was repeated until no fat remains.

b) The pellet that was formed at the bottom containing somatic cells was washed twice by adding 1 mL of PBS, pellet was mixed well by pipetting in and out then centrifuged at 3000 rpm for 10 min at 4°C.

c) Finally, the pellet was dissolved properly and dissolved properly in 500 µL of TENS buffer and kept at 55°C in dry bath for 3 h.

d) Then 200µL of proteinase k was added and kept at 37°C in incubator overnight.

e) The cellular debris and proteins were removed by adding 1 mL of Phenol: chloroform (1:1) and mixed gently by inverting. Thereafter, it was centrifuged at 5000 rpm for 10 min at 25°C.

f) The aqueous layer was taken into a separate 2 mL tube to which equal amount of chloroform: isoamylalcohol (24:1) was added, mixed and centrifuged at 5000 rpm for 10 min at 25°C.

g) The upper layer was taken in a separate tube to which $1/10^{th}$ of its volume of 3M sodium acetate was added and mixed well.

h) The DNA was precipitated by adding equal volume of chilled isopropanol and mixed gently until a visible thread appeared.

i) The DNA thread was transferred in a new 1.5 mL tube and washed twice by adding 1 mL of 70% ethanol.

j) DNA pellet was air dried and it was dissolved in 50 µL of TE buffer by keeping at 65°C for 1 h.

k) The purity of DNA was checked by running 2 µL of DNA in 0.8% agarose gel at 80 volts for 30 min and quantified by NANODROP.

3.4 DNA isolation from blood by phenol-chloroform method

5 mL of blood samples were collected using EDTA coated vacutainers as anticoagulant from jugular vein of goat. The number of breeds and number of animals

from each breed were mentioned in Table 4.6. The blood samples were brought under refrigeration and processed immediately or kept at -20°C till further processing.

Reagents:

a) **1x RBC lysis buffer**, **1.5M NH_4Cl, 100 nM $KHCO_3$ and 10nM Na_4EDTA:** 4g of NH_4Cl, 0.5g $KHCO_3$ and 0.0185g of Na_4EDTA were dissolved in 250 mL of distilled water and final volume was made to 500 mL. This solution was autoclaved and stored at room temperature.

b) **DNA Extraction buffer, 1M Tris buffer (pH 8), 3M NaCl and 0.5M EDTA (pH 8.0)** 0.5ml of 1M Tris buffer (pH 8), 4ml of 3M NaCl and 0.2 ml of 0.5M EDTA was added and final volume made up to 50ml with MilliQ .

c) **Normal saline:** 0.9g of NaCl was dissolved in 100 ml MilliQ, autoclaved and used subsequently.

Proteinase k, 10% SDS, Tris saturated phenol, Chloroform: Isoamylalcohol (24:1), 3M sodium acetate, TE buffer reagents compositions were mentioned above under reagents of section 3.3.

Procedure:

a) Around 5 mL blood was taken in a sterile 50 mL tube and three volumes of RBC lysis buffer was added to the blood and kept for 10 min in ice.

b) The solution was centrifuged at 2000 rpm for 10 min at 4°C and the supernatant was discarded. The step of lysing the RBCs was repeated till there was appearance of light yellowish colour pellet at the bottom of tube.

c) The pellet containing WBC was washed twice with 3 mL of normal saline to remove any clot present along with the pellet and then centrifuged at 2000 rpm for 10 min at 4°C, and supernatant was discarded.

d) To the WBCs, 2 mL of DNA extraction buffer and 200 µL of 10% SDS was added. The tube was tapped well so that pellet was disclosed and maximum cells get exposed to the buffer.

e) The solution was incubated at 37°C for 10 min and then 10 µL of proteinase k (20µg/µL) was added and kept overnight at 55°C or till the cells were lysed properly.

f) Next day, equal volume of Tris saturated phenol was added and mixed properly. Thereafter it was centrifuged at 5000 rpm for 10 min at 25°C.

g) The upper phase of biphasic solution was taken in separate sterile tube to which equal volume of phenol: chloroform:isoamylalcohol (25:24:1) was added mixed well by gently inverting and centrifuged at 5000 rpm for 10 min at 25°C. Again the upper aqueous phase was taken in a separate tube.

h) To remove any remnant of phenol, equal volume of chloroform: isoamylalcohol (24:1) was added and centrifuged at same conditions followed by taking upper aqueous layer into new separate tube.

i) Then DNA was precipitated by first adding 3M sodium acetate of volume $1/10^{th}$ of the aqueous layer and mixed well. Thereafter equal volume of isopropanol was added and mixed.

j) The DNA that got precipitated was pipetted out in a new sterilized tube and washed twice to remove salt contaminants with 1mL of freshly prepared 70% ethanol. After that they were centrifuged at 12000 rpm for 10 min at 4°C.

k) The supernatant was discarded and the DNA pellet was air dried.

l) The dried DNA pellet was dissolved in 50µL of TE buffer and kept at 65°C in dry bath.

m) The purity of DNA was checked by running 2 µL of DNA in 0.8% agarose gel at 80 volts for 30 min using 1x TAE buffer. Then the DNA was quantified using NANODROP.

3.5 Isolation of RNA from milk

Reagents:

a) **PBS-EDTA**: 4g NaCl, 0.1g KCl, 0.575g $Na_2HPO_4.7H_20$, 0,1g KH_2PO_4 and 0.05mM EDTA was added and volume made up to 500ml with MilliQ.

b) **80% ethanol:** 8ml absolute ethanol was added and volume made to 10ml with DEPC water.

Procedure:

Raw milk of volume 10 mL from animals of different goat breeds namely: Sirohi from Karnal (Haryana), Bakarwali from Jammu, Osmanabadi from Solapur

(Maharashtra), Jakhrana and Marwari (Rajasthan), Chegu from Himachal Pradesh and non-descript goat from Jammu was collected and brought at low temperature then processed for RNA isolation.

Milk (10 mL) was centrifuged for 10 min at 4°C, 3000 rpm, fat disc was removed from the top. This step was repeated until no fat remains, supernatant was discarded and the somatic cells in the form of pellet and was washed twice with PBS-EDTA. Then 600 µL of TRIZOL reagent was added to the pellet, vortexed properly and kept at room temperature for five min. Then 200 µL of chloroform was added and centrifuged at 12000 rpm for 10 min at 4°C. The upper aqueous layer was taken into a sterile eppendorf tube to which equal amount of chilled isopropanol was added and kept at -20°C for 10 min followed by centrifugation (12000 rpm for 10 min at 4°C). Supernatant was discarded and pellet was washed twice with 80% ethanol which is freshly prepared with DEPC treated water. Then the pellet was air dried and dissolved in 20 µL of DEPC treated water. The RNA isolated was quantified by running in 1.5% agarose gel stained with ethidium bromide in 1x TAE buffer for 10 min at 100 volts and visualized in gel documentation system, and subsequently quantified by NANODROP.

3.6 Synthesis of complementary DNA (cDNA) from RNA

cDNA was synthesized using RevertAid First Strand cDNA synthesis kit, where two µg of RNA was taken and one µL of random hexamer (100 µM) were added together and incubated at 65°C for five min then snap chilled in ice for five min. Reaction mixture containing four µL of 5x buffer, two µL of dNTPs (10 mM), one µL of ribolock (10 U/µL) and one µL of M-MuLV RT (200 U/µL) were added. cDNA was synthesized in PCR machine, incubated at 25°C for five min, followed by 60 min at 42°C and terminated by heating at 70°C for five min.

3.7 Primer designing

i) For Complete coding sequence (CDS) amplification of *CSN1S2* gene of *Capra hircus*

Based on the available coding sequences of the ovine, caprine *CSN1S2* gene and transcript variants bearing accession no: X03238.1 (Boisnard *et al*., 1991), NM_001285585 (Lagonigro *et al*., 2001, Ramunno *et al*., 2001(a), Ramunno *et al*., 2001(b)) and XM_013964673.2, XM_013964673.2, XM_013964673.2, XM_013964673.2, XM_013964673.2, XM_013964673.2 in the GenBank respectively,

a pair of oligonucleotides primers for *CSN1S2* gene were synthesized (Table 3.1) such that the forward primer, CSN1S2F binds to the 5' end of coding region from starting codon present in the 2nd exon of *CSN1S2* gene (ATG) and the reverse primer, CSN1S2R binds beyond the stop codon present in 17th exon of coding region at 3' end region (Fig. 3.1). Thus these primers amplify the cDNA, and the product size is 716 bp length which consist of 669 bp complete coding region of *CSN1S2* gene.

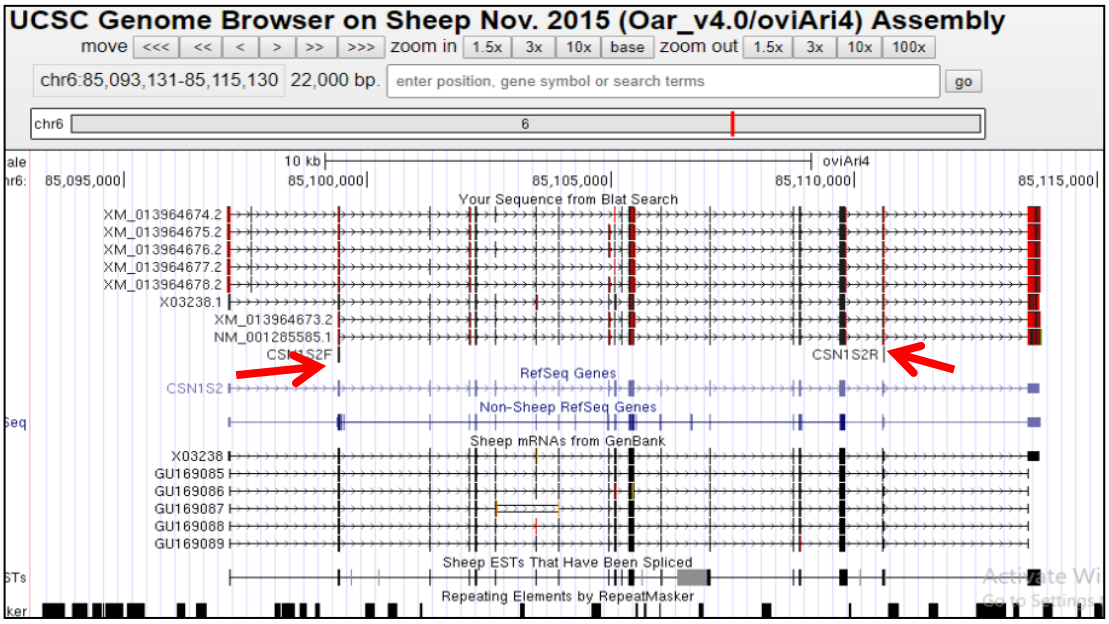


Fig. 3.1: Alignment of *CSN1S2* gene in UCSC Genome Browser. The vertical lines represent the exon and the arrow marks along the horizontal lines represent the intron. CSN1S2F and CSN1S2R are the PCR primers used for the amplification coding region.

ii) For amplification of exon region of *CSN1S2* gene of *Capra hircus* for screening of variation

Two pairs of primers were designed for variation analysis of goat *CSN1S2* gene. First pair of primers (F(11) and R(11)) was for 11th exon region and second pair (F(16) and R(16)) for 16th exon region (Table 3.1). These primers were designed based on available *CSN1S2* sequences of that particular exon region along with flanking intron of cattle, sheep and goat retrieved from GeneBank, NCBI (National Center for Biotechnology Information) and multiple aligned using online available tool, MAFTT version 7. Once the primers were designed, the Tm of selected primer pairs was checked by Oligo Calculator (UC Berkeley MCB) and it was adjusted to 60°C. Finally the primers were blast to check if it aligns with required region of *CSN1S2* gene.

Table 3.1: The list of primers, their sequences and the product size after amplification

Primers	Primer sequence	Product size
CSN1S2F	5'ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCC3'	716 bp
CSN1S2R	5'TCCAGTCCAACCATAACCAGGTAGAAGC3'	
F(11)	5'CCAGTATCTCCAGTATCTGTATCAAGGTC3'	244 bp
R(11)	5,GACAGTTGTTCTACTTTCTAAGTTAAAACTTGTCTC3'	
F(16)	5'GAATAATTATAATTTTCCTAGAAAAAAATCAGCCAGTATTACC3'	166 bp
R(16)	5'ATTAAAAATAAAAGGGAGAACTCACCACATAGGGA3'	

The forward primer, F(11) and reverse primer, R(11) bind at exon 11 region (Fig. 3.2) in a such a way that if deletion is present the size of the amplified product will be 138 bp and in absence of deletion the product will be 244 bp in length. The exon 16 was amplified to a product of size 166 bp by using primers and its binding position was shown in fig. 3.3.

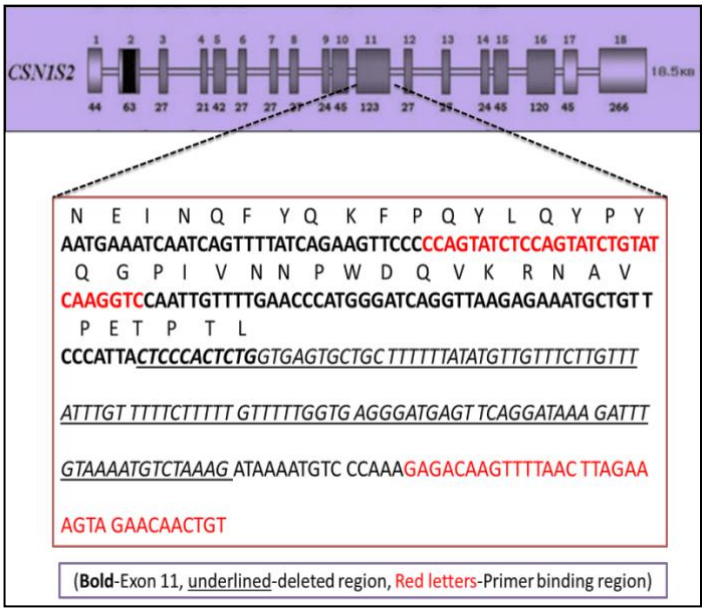


Fig. 3.2: Diagrammatic representation of goat *CSN1S2* allele D where 106 nucleotides are deleted (11 nucleotides from the 11th exon: underlined bold italics in the figure and 95 nucleotides from the 11th intron: underlined italics).

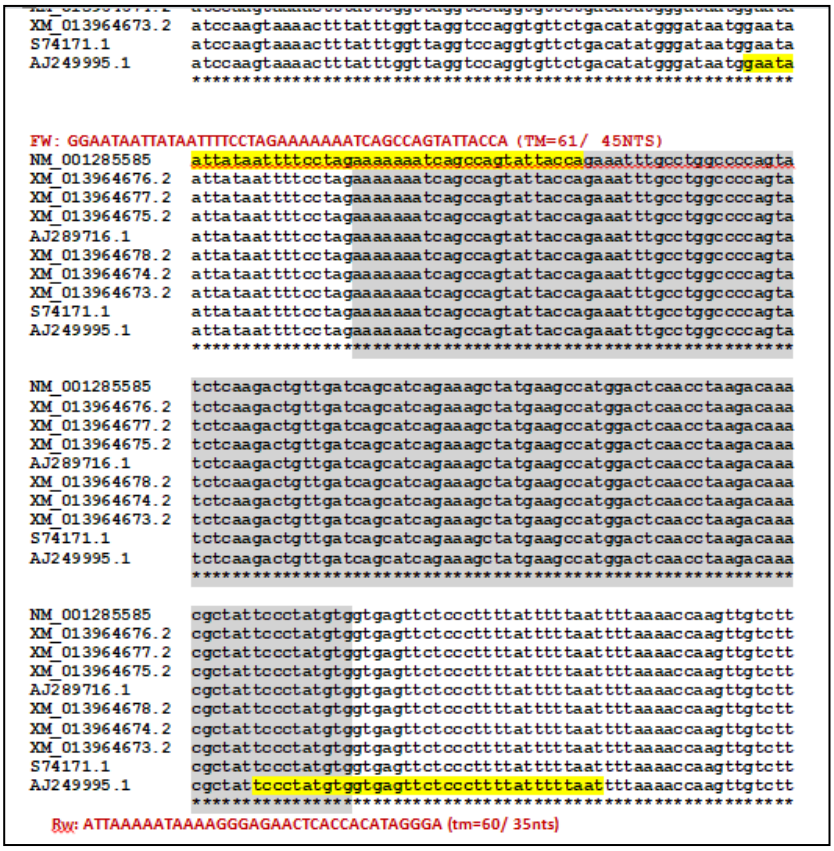


Fig. 3.3: Multiple alignment of *CSN1S2* exon 16 region. The position of primer (F(16) & R(16)) binding to the exon 16 of *CSN1S2* gene are highlighted in yellow colour. The grey colour marked region is exon 16 and 5'& 3' flanking regions are intron regions.

3.8 Polymerase Chain Reaction

The entire coding region of 669 nucleotides of Caprine *CSN1S2* gene from cDNA, exon 11 and exon 16 region from genomic DNA was amplified by PCR. All PCR amplifications were performed in 25 µL reaction volume where the components of reaction mixture and their final concentrations were given in the Table 3.2. The components were mixed properly and amplification was done by PCR under the thermo cycling parameters were mentioned in Table 3.3. The amplified product were resolved by electrophoresis on 1.5% agarose gel in case of *CSN1S2* gene and on 2.5% agarose gel in case of exon regions, stained with ethidium bromide in 1x TAE buffer at 100 volts and visualized in gel documentation system against 100bp DNA ladder.

Table 3.2: Components of PCR reaction mixture and their concentration

Components	CDS of *CSN1S2* gene	Exon 11 and exon 16 region
10X Taq polymerase buffer	2.5µL	2.5 µL
10mM dNTPs	200 µM	200 µM
Forward primer	10pM	10pM
Reverse prime	10pM	10pM
Taq DNA polymerase	1.25 U	0.625 U
Template	2 ng of cDNA	50 ng of genomic DNA
Nuclease free water	To make final 25µL	To make final 25µL

Table 3.3: The PCR conditions, their temperature and duration for amplification of CDS of *CSN1S2* gene and exon 11 and exon 16 region of *CSN1S2* gene

Conditions	Temperature	Duration			Cycles
		CDS of *CSN1S2* gene	Exon 11	Exon 16	
Initial denaturation	95°C	3 min	3 min	3 min	1
Denaturation	95°C	30 sec	15 sec	20 sec	30
Annealing	55°C	20 sec	10 sec	15 sec	
Extension	72°C	30 sec	15 sec	20 sec	
Final Extension	72°C	7 min	4 min	4 min	1

3.9 Purification of PCR product from agarose gel

PCR product containing DNA was purified from gel using Macherey-Nagel kit. The precisely cut gel (specific amplified band) under UV light was first weighed. Gel binding buffer was added (200 µL/100 mg gel) and kept at 50°C for 5 min until the gel is dissolved in the buffer. Then the solution was brought to room temperature and poured in to the Mini-spin column and centrifuged at 1100xg for 1min, the flow-through in the

collection tube was discarded. The column tube containing DNA was washed twice with 700 µL of 70% ethanol by centrifugation at 1100xg for 1 min and the flow through was discarded. The PCR product was eluted in 20 µL of nuclease free water. The eluted product was confirmed by running in 1.5% agarose gel for 10 min in 1x TAE buffer under 100 volts and quantified by NANODROP.

3.10 Ligation of purified PCR product in a vector

Ligation was done using Thermo Scientific™ InsTAclone™ PCR Cloning Kit. The purified PCR product was ligated to linear vector pTZ57R/T with the help of *E.coli* T4 DNA ligase. The ligation reaction consisted of 250µg of insert (purified PCR product), 82.5 µg of vector pTZ57R/T, 2.4U of T4 DNA ligase and 3µL of 5x T4 DNA ligase buffer. The volume was made to 15 µL with nuclease free water. Then the ligation mixture was incubated overnight at 4°C.

3.11. Transformation

3.11.1 Preparation of LB media and LB plates

Luria Bertani (LB) broth (2.5 g of LB broth powder in 100 mL of Millipore) and LB agar (30 g of LB agar powder in 1000 mL of Millipore water along with 1.5% of agar) were prepared and autoclaved. Around 20 mL of LB agar media was poured into petri dishes and allowed to solidify. Thereafter, both the agar plates and the broth were supplemented with ampicillin (50 mg/mL).

3.11.2 Transformation by heat shock method

Competent *E.coli* cells (XL-1 blue MRF' strain) stored at -80°C, were allowed to thaw. 10 µL of ligated product was added to 100 µL of competent cell and kept for 10 min in ice. After incubation for 10 min in ice, *E.coli* cells were given heat shock by keeping in water bath for 90 sec maintained at 48°C and immediately kept in ice for 10 min. For growth of the cells, 600 µL of super optimal broth (SOB) and 100 µL of 1M filter-sterilized glucose was added and incubated at 37°C for 45 min to 1h in a shaking incubator (180 rpm). Then tubes were spun for more than 2 min and the concentrate from tubes (around 100 µL) was spread onto LB agar plates supplemented with 50g/mL of ampicillin. The plates were incubated overnight at 37°C in an incubator.

3.11.3 Cloning of the transformed cells

The single isolated clones grown on plates were picked with the help of a sterile tooth pick and grown in 5 mL of LB broth and incubated at 37°C for overnight at constant shaking at 360 rpm.

3.12 Plasmid isolation

Reagents:

A. Alkaline lysis solution I: 50mM glucose, 25mM Tris and 10mM EDTA

0.45 g of glucose, 250 µL of 1M Tris and 1mL 0.5M of EDTA were mixed in 20 mL of Millipore water and pH is adjusted to 8 by using 1N NaOH, then final volume was made to 50 mL. It was autoclaved and stored at 4°C.

B. Alkaline lysis solution II: 0.2N NaOH and 1% (w/v) SDS

Freshly prepared before use by adding 1mL 10N NaOH and 5mL of 10% SDS and final volume was made to 50 mL with Millipore water and used at room temperature.

C. Alkaline lysis solution III: 5M Potassium acetate and glacial acetic acid

60 mL of 5M Potassium acetate (49.07g potassium acetate in 100mL H2O) and 11.5 mLof glacial acetic acid were added together and made final volume to 100mL by milllipore water. Stored at 4°C

D. Tris-EDTA (TE) Buffer: (refer 3.3)

Procedure:

a. The overnight grown culture (transformed cells) was centrifuged at 8000 rpm for 10 min at room temperature to pellet down the grown cells and the supernatant was discarded.

b. The pellet was resuspended in 300 µL of alkaline lysis solution I by pipetting in and out or vortexed well so that bacterial cells get resuspended to the maximum and the suspension was transferred to 2 mL eppendorf tubes.

c. Then 400 µL of alkaline lysis solution II was added to the suspension and gently mixed by inversion and incubated at room temperature for 5 min.

d. 300 µL of ice cooled alkaline lysis solution III was added for precipitation, then gently invert mixed and immediately kept in ice for 20 min.

e. After the incubation, centrifugation was done at 12000 rpm for 10 min at 4°C to settle down the precipitate and the supernatant containing the plasmid was taken in a fresh 1.5 mL eppendorf tube without taking any precipitate.

f. The plasmid was precipitated by adding equal volume of chilled isopropanol to the supernatant taken and centrifuged at 12000 rpm for 10 min at 4°C.

g. The pellet was washed twice with 70% ethanol and centrifuged at 12000 rpm for 10 min at 4°C, supernatant was discarded and pellet was air dried.

h. The air dried pellet was dissolved in 50 µL of TE buffer by incubating at 65°C for 1 h. RNA contamination was removed by adding 2 µL RNase A solution and incubated at 37°C for 1 h.

i. Then purification of plasmid was done by adding phenol and chloroform in 1:1 ratio to equal volume of the plasmid sample, mixed well and centrifuged at 3000 rpm for 10 min at 25°C.

j. After centrifugation, the upper layer was taken into a fresh tube and to that equal amount of chloroform was added, centrifuged and then again upper aqueous layer was taken in new Eppendorf tube. To the aqueous upper layer, $1/10^{th}$ of its volume, 3M sodium acetate was added and mixed well.

k. Then plasmid was precipitated by equal volume of isopropanol and centrifuged at 12000 rpm for 10 min at 4°C.

l. The pellet was washed with 70% ethanol by centrifuging at 12000 rpm for 10 min at 4°C, supernatant was discarded and pellet was air dried.

m. The air dried pellet was dissolved in 50 µL of TE buffer by incubating at 65°C for 1 h and stored at -20°C for further use.

3.13 Screening of recombinant clones by PCR

The plasmids isolated above were screened for the presence of insert by PCR using 10pM each same set of primers (primer names), 200 µM dNTPs, 1x Taq DNA polymerase buffer, 0.75 units of Taq DNA polymerase and 30ng/ml of plasmid DNA as a template. Final volume was made 25uL with nuclease free water. PCR amplification

conditions are mentioned in table (Table 3.3). Amplified products were resolved by electrophoresis on agarose gel stained with ethidium bromide, run in TAE buffer for 20 min and visualized in gel documentation system against 100bp DNA ladder. The selected clones were sent for sequencing.

3.14 Editing and analysis of sequences

a) The nucleotide sequences which were obtained after sequencing were scanned for any presence of vector DNA by using VecScreen tool of NCBI and trimmed it if present.

b) Then all the sequences were Blast to check matching with references sequences of GenBank.

c) If they are matching then the alignment was checked and in case of reverse alignment, the sequences were reverse complemented by using reverse complement tool in sequence manipulation suite.

d) In case of CDS of *CSN1S2* gene, the coding region of all were selected by selecting starting codon (ATG) and stop codon (TAA) and other sequences were trimmed. In case of exon 16, 5' and 3' ends of exon 16 were marked and other flanking sequences were trimmed.

e) The nucleotide sequences were translated to amino acid sequences by using online available tool, translate of sequence manipulation suite.

f) All the nucleotide sequences were multiple aligned along with the reference sequences (Acc no-NM_ 001285585.1, S74171, AJ249995, AJ289716, AJ289715) using MAFTT version 7 which is online available.

g) Similarly, all the protein sequences after translating were multiple aligned with the reference sequences in GenBank.

h) Any variations in the sequences were highlighted and their positions were noted.

3.15 Polymerase Chain Reaction-Single Stranded Conformation Polymorphism (PCR-SSCP) (Hayashi, 1991)

Reagents:

a) Acrylamide-Bis-acrylamide (37.5:1): 37.5g acrylamide and 1g N, N'-methylenebisacrylamide

b) **SSCP gel loading buffer:** 9.5 mL of formamide, 200 µL of 0.5 M NaOH, 10 mg of xylene cyanol, 10mg of bromophenol blue and 400 µL of 0.5 M EDTA (pH 8.0) were added together and vortex properly.

c) **5X TBE buffer (Tris Borate EDTA buffer pH 8.0):** 27g of Tris base and 13.7g of Boric acid were added to 300 mL of millipore water and mixed well using magnetic stirrer till they get dissolved and the final volume was made to 500 mL by millipore water. 10 mL of 0.5 M EDTA (pH8.0) was added and mixed well.

d) **Ammonium per sulphate (APS) (10%):** 1g of APS dissolved in 10 mL of millipore water.

e) **Fixative solution:** 60 mL of glacial acetic acid, 250 mL of methanol, 100 µL of 37% formaldehyde solution were added together and the final volume was made to 500 mL with millipore water. The solution was kept in amber bottle and stored in dark with no exposure to light.

f) **Silver staining solution:** 0.2g of $AgNO_3$ and 150 µL of 37% formaldehyde solution dissolved in 200 mL of millipore water. The solution was kept in amber bottle and stored in dark with no exposure to light.

g) **Developer solution:** 7.5 g of sodium carbonate, 187.5 µL of 37% formaldehyde solution and 0.001g of sodium thiosulphate were dissolved in 250 mL of millipore water in amber bottle and stored in dark with no exposure to light.

h) **Stop solution:** 80 mL of glacial acetic acid and 250 mL of methanol added together and final volume was made to 500 mL by millipore water.

3.15.1 Single Stranded Conformation Polymorphism (SSCP) gel preparation

1. Casting glass plates were cleaned with anionic surfactant and distilled water to remove any grease from the surface. The plates were placed together by keeping spacer in between them at both the adjacent sides. Then plates were sealed with the help of microporous tape to all the three sides of the glass plates.

2. Acrylamide gel mix (10%) was prepared by adding 6.7 mL of acrylamide:bisacrylamide (37.5:1), 1ml of glycerol, 150 µL of 10% APS in 12.3mL of 1x TBE buffer and at last 10 µL of TEMED was added.

3. The gel mixture was mixed thoroughly and poured carefully between the gap of glass plates immediately avoiding any air bubble within the gel and the comb was inserted.

4. The gel was kept at room temperature till it polymerized then kept at 4°C for 30 min before electrophoresis.

3.15.2 Sample preparation and gel electrophoresis

1. The sample was denatured with SSCP gel loading solution in 1:2 ratio at 95°C for 5 min followed by snap chilling in ice before loading.

2. The comb from the gel cast was removed carefully and the wells were cleaned properly with distilled water. Then the plates were placed into the casting apparatus and kept inside the buffer tank.

3. The loading dye was poured into wells and pre run was done for 20 min at 110 volts. The denatured samples were loaded into the wells carefully. Electrophoresis was performed at 110 volts in 0.5x TBE gel running buffer at 4°C for 17 hours.

3.15.3 SSCP Gel staining

1. The gel was removed from the glass plates carefully and it was fixed in fixative solution for 30 min at room temperature with gentle shaking and the gel was washed quickly with distilled water for 2 min at room temperature with gentle shaking.

2. Staining was done with silver nitrate staining solution for 45 min at room temperature with gentle shaking.

3. The gel was rinsed quickly twice with distilled water for 2 min at room temperature with gentle shaking.

4. The bands were visualized in developer solution at room temperature with gentle shaking.

5. Developing reaction was stopped by adding stop solution for at least 30 minutes.

6. The gel was rinsed with distilled water for at least 30 minutes at room temperature.

CHAPTER –4

Results and Discussion

The present study was initiated with the objective of cloning and sequencing of complete *CSN1S2* gene and then to explore the possible sequence variation in coding region of *CSN1S2* gene in Indian goat breeds. The following are the results of the work done under these objectives.

In objective 1, RNA was isolated from milk of different breeds of goat, cDNA was synthesized, *CSN1S2* gene was amplified and cloned. The *CSN1S2* gene was screened and sequenced. The sequences of different goat breeds were compared and analyzed.

4.1 Coding region of *CSN1S2* gene of six Indian goat breeds:

Milk from six different goat breeds namely Sirohi, Bakarwali, Osmanabadi, Marwari, Jakhrana and Chegu (Table 4.1) from different agroclimatic regions of India were collected and total RNA (Fig. 4.1) was isolated by TRIZOL reagent. Approximately 6 µg of total RNA was isolated from 10 ml of fresh milk.

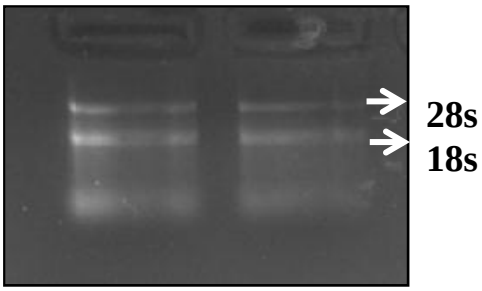


Fig. 4.1: Agarose gel electrophoresis of RNA (isolated from goat milk somatic cell by Tri-zol reagent. Lane 1 and 2 represent bands of RNA resolved in 1.5% agarose gel, run in 1x TAE buffer for 10 mins at 100 volts).

cDNA was synthesized from 2µg of total RNA. Taking cDNA as template, the complete coding region of caprine *CSN1S2* gene was amplified by polymerase chain reaction at 55°C annealing temperature using primers, CSN1S2F and CSN1S2R. The amplified product was 716 bp length and was confirmed by agarose gel electrophoresis (Fig. 4.2). The 716 bp band was cut out from agarose gel for purification of the DNA. The purified 716 bp DNA was ligated to pTZ57R/T vector in 3:1 ratio and cloned by transformation into XL-1 blue MRF' strain of *E.coli.* A total of ten recombinant colonies

(Fig. 4.3) were selected from each animal and plasmids were isolated for screening of insert DNA by PCR using same set of primers for amplification of CDS of *CSN1S2* gene (Fig. 4.4). The colonies with specific band at 716 bp in agarose gel were sent for capillary sequencing.

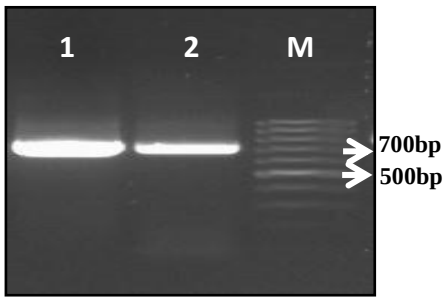


Fig. 4.2: Agarose gel electrophoresis of PCR amplified products of *CSN1S2* gene from cDNA, run in 1.5% agarose gel, 1x TAE buffer for 15 min at 100 volts. Lane 1 and 2: amplified fragment of ~716 bp products. Lane M: 100 bp double stranded DNA marker.

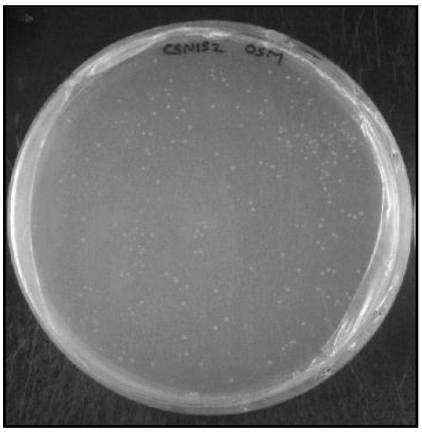


Fig. 4.3: LB agar plates showing recombinant colonies grown overnight at 37°C in an incubator. Agar plates are supplemented with ampicillin (50 mg/mL).

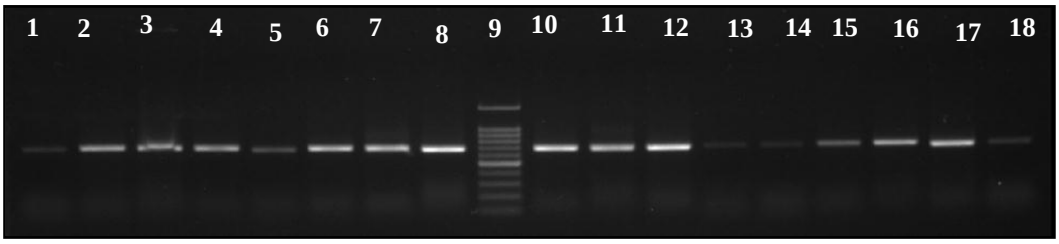


Fig. 4.4: Agarose gel (1.5%) electrophoresis to screen individual clones for insert. Lane 1-17: clones with positive insert (~716 bp), Lane 9: 100 bp linear double stranded DNA marker. Run in 1x TAE buffer for 15 mins at100 volts.

All the raw sequences were screened for any vector sequences present and trimmed them. Then only the coding region was selected by locating its start and stop codon and analyzed by comparing to the reference sequences present in Genbank. Reference nucleotide sequences were retrieved from GenBank for accession no: X65160, NM_001285585.1, S74171, AJ249995, AJ289715, AJ289716. Sequences belonging to accession no: X65160 and S74171 were identified in French goat breeds, Alpine and Saanen (Bouniol, 1993; Bouniol *et al*., 1994) subsequently named as allele A and C respectively .The accession no: AJ289715, AJ249995 and AJ289716 were identified from goats of Southern Italy and named as allele 0, E and F (Ramunno *et al*., 2001(a); Lagonigro *et al*., 2001; Ramunno *et al*., 2001(b) respectively) (Table 4.2). The amino acid sequences having accession no, NP_001272514.1, AAB32166.1, CAB59920.1 and CAB94236.1 corresponding to A, C, F and E alleles respectively were retrieved and compared with amino acid sequences translated from nucleotide sequences of the *CSN1S2* gene of selected Indian breeds below.

The results of sequence analyses were explained for each breed separately in the following paragraphs.

4.1.1 Bakarwali

Bakarwali is a well-known local breed of Jammu & Kashmir region reared for both meat and milk production. A total of two animals were included in the present study. RNA was isolated from milk, collected from these two animals and further experiment was continued as mentioned above. For one R'T-PCR cloning experiment, ten colonies (recombinant clones) were selected. Four were sequenced out of total twenty clones (ten clones of each animal) in which one clone (bkws2) is from one animal and

other three (S2BK4, S2BK2, S2BK5) are of another animal of same breed. The coding region of all the nucleotide sequences and the translated protein sequences were multiple aligned with the sequences present in GenBank. The sequences of four clones of two animals were analyzed and three different variants were found from which two were reported earlier and one was a new variant. In the bkws2 clone, there was transversion of A to T at 545th position of coding region of *CSN1S2* gene leading to substitution of Lys_{167} to Ile_{167} when compared to the sequence of allele A (Acc. no- X65160; NM_001285585.1) reported by Bouniol *et al*., 1994 as C allele. The S2BK4 variant was found to be B allele as there is transition of G to A at 235th position in comparison to allele A (Boulanger *et al*., 1984; Bouniol, 1993) which was a non-silent variation with substitution from Glu_{64} to Lys_{64} but there was a silent transition at 505th nucleotide position from C to T which was also found in S2BK5 variant. A new variant was found in clone, S2BK2 which has similar sequence as that of S2BK4 apart from transition of C to T at 326th position which leads to substitution of Pro_{94} to Leu_{94} which is not present in reported alleles. At the same time, exon skipping was also observed in three clones, bkw and S2BK2 in which exon 6 was skipped and in S2BK5 clone, exon 9 was skipped.

Table 4.1: List of Indian goat breeds selected from different regions of India and number of animals from each breed

S.No	Breed	No of animals	Region
1	Sirohi	2	Karnal, Haryana
2	Bakarwali	2	Jammu
3	Osmanabadi	2	Solapur, Maharashtra
4	Marwari	5	Rajasthan
5	Jakhrana	5	Rajasthan
6	Chegu	2	Himachal Pradesh
7	Non-descript	1	Jammu

Table 4.2: Reported alleles and position of variations at nucleotide level of coding region of *CSN1S2* gene. (Exon number and position of nucleotide in exon was highlighted by green)

Reported Alleles	Position of nucleotide variation in coding region						Accession no. of nucleotide sequences	Reference
	64	235	374	408-418	545	623		
A	G	G	G	No deletion	A	C	X65160; NM_001285585	Boulanger *et al.*, 1984; Bouniol *et al.*, 1993
B		A E9,10					Not available	Boulanger *et al.*, 1984
C					T E16,5		S74171	Bouniol *et al.*, 1994
D				Deletion			Not available	Ramunno *et al.*, 2001(b)
E					T	G E16,83	AJ249995	Lagonigro *et al.*, 2001
F	A E3,13						AJ289716	Ramunno *et al.*, 2001(b)
0			A E11,80				AJ289715	Ramunno *et al.*, 2001(a)

4.1.2 Sirohi

Sirohi breed of goat was local to Sirohi district of Gujarat and Rajsmand of Rajasthan but this breed was also found in Karnal district of Haryana. Milk from two Sirohi goat was included in the present study. Out of twenty clones, two clones (sirs1 & sirs2) of *CSN1S2* gene each belonging to those two unrelated animals were sequenced and multiple aligned with reference sequences. Both the sequences are allele C as there was transversion from A to T at 545th region of *CSN1S2* coding region leading to substitution from Lys_{167} to Ile_{167} in amino acid sequences 16th exon.

4.1.3 Osmanabadi

Osmanabadi breed which are reared for milk and meat were local to different districts of Maharashtra. Milk was collected from Hiraj village in Solapur district of

Maharashtra. RNA was isolated from milk from two unrelated animals and further proceeded the experiments to produce recombinant clones. Ten colonies belonging to each animal were selected for screening for insert. Four clones, S2OSM5 and osms2 from one animal and S2OSM33 and S2OSM88 from another animal were sequenced. The nucleotide and amino acids sequences were multiple aligned with reference sequences. All the clones except one were similar to A allele (Boulanger *et al*., 1984; Bouniol, 1993). In osms2 clone, there was a transition from T to C at 94^{th} position leading to substitution in amino acid sequence from Phe_{17} to Lue_{17} of mature functional protein which found to be a new variation leading to a new allele.

4.1.4 Marwari

Marwari is a well-known goat breed of Rajasthan reared for milk and meat. Total RNA was isolated from milk collected from five unrelated animals. Out of fifty clones (ten clones belonging to each animal), five clones of five different unrelated animals were sequenced and their nucleotide and amino acid sequences were multiple aligned. Out of five clones two (M7B and M5H) were allele C whereas three clones (M5F, M6A & M6F) were found to be new variant. In M6A variant along with transversion of A to T at 545^{th} position there was transition of A to G at 554^{th} position in nucleotide sequence with amino acid sequence substitution from Lys_{167} to Ile_{167} and Gln_{170} to Arg_{170} respectively. In case of M5F, there was transition of G to A at 247^{th} position, transversion of A to T at 545^{th} and C to G at 622^{nd} position and these variations lead to substitutions in the amino acid sequence from Glu_{68} to Lys_{68}, Lys_{167} to Ile_{167} and Pro_{193} to Ala_{193} respectively. The third new allele (M6F), was formed by transition at 134^{th} and 625^{th} position from A to G and T to C respectively and were responsible for substitution from His_{30} to Arg_{30} and Trp_{194} to Arg_{194} of mature functional protein. Skipping of 6^{th} exon in M5F and M7B and 13^{th} exon in M6A clone was also observed.

4.1.5 Jakhrana

Jakhrana breed of goats belong to Alwar district of Rajasthan and milk was collected from five different animals. Five clones of five different animals were sequenced where two clones (J6A and J6E) were not having any sequence difference so they were found to be A allele. One clone (J7F) was having variation at 235^{th} position similar to B allele (Boulanger *et al*., 1984) but there was also presence of silent mutations at 505^{th} and 630^{th} position. There were two new variants (J7A and J8E) where

in one clone (J7A) three transition at 235th from G to A, 505th from C to T and 604th from C to T in nucleotide sequence and there was no change in protein sequence due to transition at 505th position but other two transitions leads to a new allele showing substitutions from Gln_{64} to Lys_{64} and His_{187} to Tyr_{187}. The other new variant (J8E) show four changes, one was silent change, G to A at 111th position and other three non-silent variations at 499th from A to G, 545th from A to T and 600th from T to A position which substitute the amino acids from Thr_{152} to Ala_{152}, Lys_{167} to Ile_{167} and Asp_{185} to Glu_{185} region of protein. There were two exon skipping, 3rd exon in J7A and 6th exon in J6A clone.

4.1.6 Chegu

Chegu is a local breed of Himachal Pradesh. Milk was collected from two goats belonging Lahul district of H.P. Three (CH1S2_3, CH3S2_6 and CH3S2_8) out of total twenty clones of two animals were selected for sequencing and analyzed that the three clones where allelic C variant due to transversion from A to T at 545th nucleotide position also reported by Bouniol and coworkers (1994) and there was skipping of 6th exon in all the three clones.

4.1.7 Non-descript goat

Milk from a non-descript goat from Jammu was also collected. Three clones (S2LJ1, S2LJ3 and S2LJ5) were sequenced, analyzed and no variation was observed in *CSN1S2* gene nucleotide sequence when compared to reference sequence. All clones were allele A.

4.1.8 Nucleotide sequences of complete coding region of *CSN1S2* gene of all clones of six breeds

bkws2

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

S2BK2

TTGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTAAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCTCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAATTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

S2BK4

TTGAAGTTCTTCATTTTTACTTGCCTTTTGGCCGTTGCCCTTGCAACGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTAAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAATTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

S2BK5

CTGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAG
TATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTC
TCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAATTGACTGAAGAAGAAAAGAATCGCCTAAAT
TTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCT
AAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

sirs1

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

sirs2
ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

osms2

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCCTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

S2OSM5

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

S2OSM33

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

S2OSM88

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

M5F

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTGAAGTTGCCCCAAAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGGCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

M5H

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

M6F

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCGTCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCACGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

M6A

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAATAATCAGCCGGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

M7B

ATGAAGTTCTTCATTTTTACCTGCCTCTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

J6A

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

J6E

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

J7A

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATGAACCTATCAATATCTTCCAGGAAATATATAAGCAGGAAAAGAATATGGCC
ATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTAAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAATTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGTATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

J7F

ATTAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTAAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAATTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAAAAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAA
GCTATGAAGCCATGGACCCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

J8E

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAACAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAGAAACTTTGCACCACATCCTGTGAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATAT
TCTATCAGATCATCTAGTGAGGAATCTGCTGAAGTTGCCCCAGAGGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAAT
CAGTTTTATCAGAAGTTCCCCCAGTATCTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCTTTACT
CCCACCGTGAACAGAGAGCAGCTCTCCACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAGCTAAACTGACT
GAAGAAGAAAAGAATCGCCTAAATTTTCTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGAACAGCATCAGAAA
GCTATGAAGCCATGGACTCAACCTAAGACAAATGCTATTCCCTATGTGAGGTACCTTTAA

CH1S2_3

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTGAAGTTGCCCCAGAAGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

CH3S2_6

ATGAAGTTCTTCATTTTTACCTGCCTTTTGGCCGTTGCCCTTGCAAAGCATAAGATGGAACATGTCTCCTCCAGTGAGGAACCTATCAATATCTTCCAGGAA
ATATATAAGCAGGAAAAGAATATGGCCATTCATCCCAGAAAGGAAGTTGTAAGGAACGCAAATGAAGAGGAATATTCTATCAGATCATCTAGTGAGGAATCT
GCTGAAGTTGCCCCAGAAGAAATTAAGATTACTGTGGACGATAAGCACTACCAGAAAGCCCTGAATGAAATCAATCAGTTTTATCAGAAGTTCCCCCAGTAT
CTCCAGTATCCGTATCAAGGTCCAATTGTTTTGAACCCATGGGATCAGGTTAAGAGAAATGCTGGCCCCCTTTACTCCCACCGTGAACAGAGAGCAGCTCTCC
ACCAGTGAGGAAAATTCAAAGAAGACCATTGATATGGAATCAACAGAAGTATTCACTAAGAAAACTAAACTGACTGAAGAAGAAAAGAATCGCCTAAATTTT
CTGAAAATAATCAGCCAGTATTACCAGAAATTTGCCTGGCCCCAGTATCTCAAGACTGTTGATCAGCATCAGAAAGCTATGAAGCCATGGACTCAACCTAAG
ACAAATGCTATTCCCTATGTGAGGTACCTTTAA

4.2 Comparison of all complete coding sequences of *CSN1S2* gene of Bakarwaki, Marwari, Sirohi, Jakhrana, Osmanabadi and chegu breeds of goat.

All the nucleotide and amino acid sequences of all clones belonging to different breeds mentioned above were analyzed with the known *CSN1S2* sequences from GenBank (Table 4.2). A and B allele were identified by SDS-PAGE in Barbari breed of goat (Garg *et al*., 2009). Jamunapari, Barbari, Jakhrana, Marwari, Sirohi, Beetal and Ganjam, A, B and 0 alleles were present (Sharma *et al*., 2017). In the present study, the sequence analyses of clones revealed that allele C was predominant to all other alleles followed by allele A and B, where C allele is present in Bakarwali, Sirohi, Marwari and Chegu out of six breeds. Allele A was present in Osmanabadi and Jakhrana and B allele was found in Bakarwali and Jakhrana (Table 4.3).

Table 4.3: Result of complete coding *CSN1S2* gene sequences of different breeds analyses. The number of clones that are sequenced and the alleles found.

Breed (no. of animals)	No. of clones	Reported alleles	No. of new alleles
Osmanabadi (2)	**4**	**A**	**1**
Bakarwali(2)	**4**	**C and B**	**1**
Sirohi(2)	**2**	**C**	**-**
Marwari(5)	**5**	**C**	**3**
Jakhrana(5)	**5**	**A and B**	**2**
Chegu(2)	**3**	**C**	**-**
Non-descript goat(Jammu)	**3**	**A**	**-**

Out of 26 sequences, nine belong to allele A, eight belong to allele C, two belong to B allele and the rest seven clones were new alleles. Total 31 types of variations were found at different position considering allele A nucleotide sequence as reference. Out of total variations, seven were silent mutation which were not changing amino acid sequences, nine variations were responsible for formation of B and C allele which are reported earlier Boulanger *et al*., 1984; Bouniol *et al*., 1994) and fifteen variations indicated seven type of new alleles. The new alleles were found in four breeds, one allele in Osmanabadi, one allele in Bakarwali, three alleles in Marwari and two alleles in Jakhrana breed of goat. In the earlier study, variation were found in 3rd, 9th, 11th and 16th exons {Boulanger *et al*., 1984; Bouniol *et al*., 1994; Lagonigro *et al*., 2001; Ramunno *et al*., 2001(a) ; Ramunno et al., 2001(b)} whereas in the present study variations were found at 4th, 5th, 11th, 15th and 16th exons. The percentage of variation was higher in case of 16th exon of *CSN1S2* gene according to present (Table 4.4) and earlier studies, thus this exon is more prone to mutations. Exon skipping was observed in clones of Bakarwali, Jakhrana, Marwari and Chegu breeds (Table 4.5).

Table 4.4: Novel nucleotide variations and its position at nucleotide level of coding region of *CSN1S2* gene in comparison to allele A. (Exon number and position of nucleotide in exon was highlighted by green and red font represent the silent mutations).

NOVEL ALLELES	Position of nucleotide variation in coding region											
	94	111	134	247	326	499	505	554	600	604	622	625
Allele A	T	G	A	G	C	A	C	A	T	C	C	T
osms2	C E4,16											
S2BKW2					T E11,32		T E15,10					
M6A								G E16,14				
M5F				A E9,22							G E16,82	
M6F			G E5,32									C E16,85
J7A							T			T E16,64		
J8E		A E5,9				G E15,4			A E16,60			

Skipping of exon was observed in case of exon 3, exon 6, exon 13 of 27 nt and exon 9 of 24 nt. Due to their small size, these exons are prone to skipping and exon 6 skipping is common in these breeds due to alternative splicing. The same was also reported in sheep and is known as a stabilized mechanism for production of multiple proteins.

Table 4.5: Clones belonging to different breeds and the exon skipped and its size (nucleotide base pairs)

CLONES	BREED	EXON SKIPPED and its size
Bkws2	BAKARWALI	6 (27 nt)
S2BKW2		6 (27 nt)
S2BKW5		9 (24 nt)
M6A	MARWARI	13 (27 nt)
M5F		6 (27 nt)
M7B		6 (27 nt)
J6A	JAKHRANA	6 (27 nt)
J7A		3 (27 nt)
CH1S2_3	CHEGU	6 (27 nt)
CH3S2_6		6 (27 nt)
CH3S2_8		6 (27 nt)

The amino acids sequences of all 26 clones along with that of reference sequences, (Acc. no: NP_001272514.1, AAB32166.1, CAB59920.1, CAB94236.1 and CAA10361 corresponding to A, C, F, E and 0 alleles respectively. No reference sequences for B and D allele were found (Boulanger *et al*., 1984; Rammuno *et al*., 2002) a phylogenetic tree was built (Fig. 4.5). Three main cluster (A, B & C) revealed in the phylogenetic tree. A major cluster (A) was found for allele A which was matching to our J6E, S2LJ3, S2LJ5, S2LJ1, S2OSM5, S2OSM33and S2OSM88 clone sequences J7F and S2BK4 are originated from allele A but present together were these belong to B allele. A new allele

in osms2 clone was present close to allele A and allele F. Another cluster (B) represents the clones (sir1, sir2 and M5H) similar to C allele. The three new alleles J8E, M6F and J7A along with allele E are present in same cluster but originated separately. Third cluster (C) is different mainly due to exon skipping where CH1S2_3, CH3S2_6, CH3S2_8, bkws2 and M7B are similar to C allele but they are grouped together separately as deletion of exon 6. The new alleles represented by S2BK2, M6A and M5F are present in this cluster. The S2BK5 and J6A were similar to A allele but they were present at different nodes as exon 9 and 6 were skipped respectively. One null allele (allele 0) having a premature stop codon is distantly located from all the three clusters representing as out group.

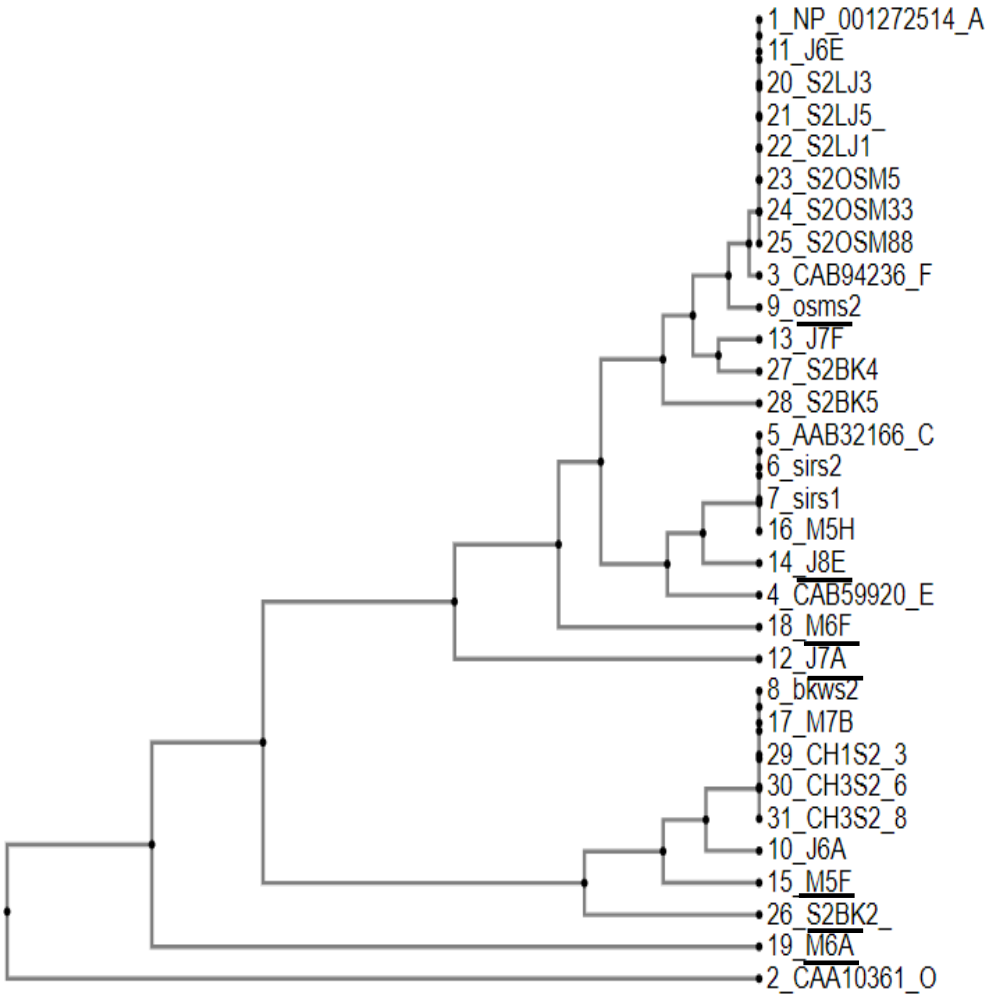


Fig. 4.5: Phylogenetic tree of amino acid sequences of *CSN1S2* gene of all 26 clones of different breeds and reference sequences of five alleles (A, C, E, F and 0) and underlined were seven novel alleles.

In objective 2, exon 11 and exon 16 of *CSN1S2* gene was studied. Genomic DNA was isolated from blood or milk, the exon regions were amplified and analyzed for variation.

4.3 Amplification of large exons and screening for sequence variation:

Variation was screened at genomic DNA level in 11th and 16th exons of 123 and 120 bp region of *CSN1S2* gene of different Indian goat breeds. DNA was isolated from 5ml of blood by phenol chloroform method from different number of animals from different breeds (Table 4.6) namely Jakhrana, Barbari & Jamunapari from CIRG, Makhdoom; Sirohi from Avikanagar; Ganjam from Chiplima, Odisha; Black Bengal from Chiplima and Tripura; Chegu from Himachal Pradesh; Osmanabadi from Maharashtra; Marwari from Rajasthan and Bakarwali from Jammu. DNA was also isolated from 15 ml of milk by phenol chloroform method of Gaddi breed from Mandi district of Himachal Pradesh. Total number of goats was 140 belonging to eleven goat breeds that were screened for mutations. The DNA was isolated, quantified and used for amplification of particular exon region for screening of variation. In India till date there was no study at genomic level for variation or to find out new alleles in the *CSN1S2* gene of goats.

Table 4.6: List of breeds of goat and number of animals per breed screened for variation in *CSN1S2* gene in exon 11 and 16 region

S.no	Breed	No of animals per breed
1	JAKHRANA (CIRG, Makhdoom)	20
2	JAMUNAPARI (CIRG, Makhdoom)	9
3	SIROHI (Avikanagar, Rjasthan)	10
4	GANJAM (Chiplima,Odisha)	10
5	BLACK BENGAL (Chiplima and Tripura)	10
6	BARBARI (CIRG, Makhdoom)	10
7	GADDI (Mandi district, Himachal Pradesh)	21
8	CHEGU (Himachal Pradesh)	10
9	OSMANABADI (Solapur, Maharashtra)	5
10	MARWARI(Rajasthan)	20
11	BAKARWALI (Jammu)	15

The D allele was formed due to deletion of 106 nt, 11 nt from 3' end of exon 11 and 95 nt from 5' end of immediate next intron in genomic DNA leading to deletion of three amino acids in αs2 casein protein (Ramunno *et al*., 2001(b)). In this present study 106 nt deletion in exon 11 region was selected and amplification results were observed in agarose gel. 140 goats belonging to eleven different Indian goat breeds were screened for deletion in this region. The forward F(11) and reverse primers R(11) designed for this region were and respectively (Table 3.1) bind to genomic DNA in such a way that in absence of D allele in genomic DNA, the amplified product band will be observed at 244 nt length position indicating no deletion. In presence of D allele (homozygous), band was observed at 138 nt length position and in heterozygous condition two bands at 244 nt and 138 nt length position can be observed. According to our results, it was found that the no animal was positive for D allele as there was band at 244 nt position (Fig 4.6a, b, c). The D alleles responsible for decrease in αs2 casein protein concentration was absent in Indian breeds.

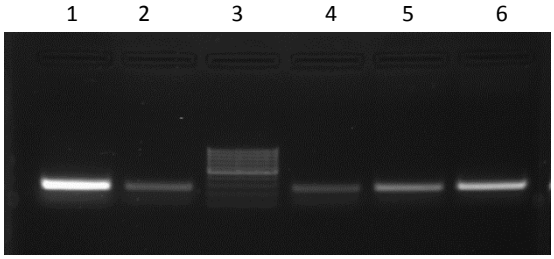


Fig. 4.6a: Agarose gel (2%) electrophoresis of exon 11 PCR amplified product from genomic DNA of Osmanabadi goats for the screening of deletion of 106 nucleotides in *CSN1S2* gene. Lane 3: 100bp linear double stranded DNA marker and in others bands of 244 bp PCR product.

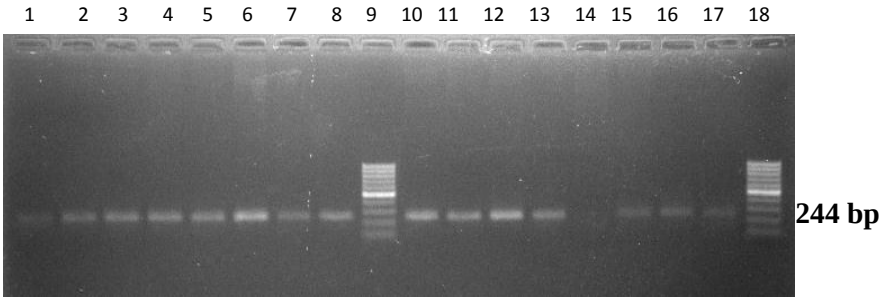


Fig. 4.6b: Agarose gel (2%) electrophoresis of PCR amplified product from genomic DNA of Jakhrana and Jamunapari goats for the screening of deletion of 106 nucleotides in *CSN1S2* gene. Lanes 1 to 8 and 10 to 17: ~244 bp PCR product. Lanes 9 and 18: 100bp DNA marker.

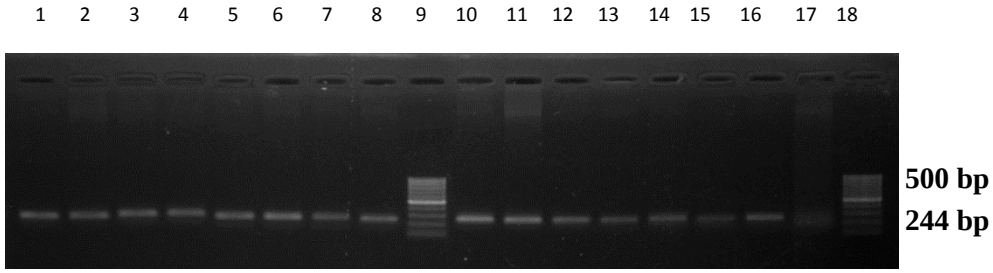


Fig. 4.6c: Agarose gel (2%) electrophoresis of PCR amplified product from genomic DNA of of Barbari and Black Bengal goats for the screening of deletion of 106 nucleotides in *CSN1S2* gene. Lanes 1 to 8 and 10 to 17: ~244 bp PCR product & in lanes 9 and 18, 100bp linear double stranded DNA marker.

In reported alleles there were maximum variations in exon 16 responsible for formation of C and E allele (Table 4.2). CDS analyses of *CSN1S2* gene of six Indian breeds, seven novel alleles were found out of which five new alleles have maximum variations in exon 16 (Table 4.4). Therefore, exon 16 was selected to screen unknown variation in Indian breeds. The full 16th exon (120 nt) was amplified using F(16) and R(16) as forward and reverse primers (Table: 3.1) from genomic DNA and the length amplified product was 166 nt (Fig. 4.7), thereafter, SSCP was done in 10% non-denaturing acrylamide gel with the same PCR product. PCR-SSCP analysis 20 goats each of Marwari and Jakhrana breed, 5 goats of Osmanabadi and 15 goats of Bakarwali breed was done. While analyzing the SSCP pattern of exon 16, it was revealed that

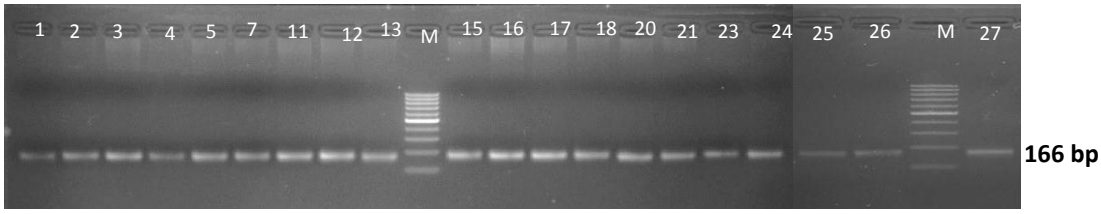


Fig. 4.7: Agarose gel (2%) electrophoresis of PCR amplified product of exon 16 in *CSN1S2* gene from genomic DNA of different goats of Marwari breed. Run in 2.5% agarose gel, 1x TAE buffer for 20 mins at100 volts. Amplified product is of 166 bp against 100bp DNA marker.

almost all animals (belonging to Marwari, Jakhrana, Osmanabadi and bakarwali breeds) showed a similar pattern. This is assumed to be wild type that is allele A (Fig. 4.8). But in Fig. 4.9 in lane of M19 and BK2 different band pattern was observed. This variant one is a heterozygous situation for the exon 16 region because of which a different pattern was found and it is most likely that these animals (M19 & BK2 belong to Marwari and Bakarwali breed) were having one wild type allele along with one new variant or allele for exon 16.

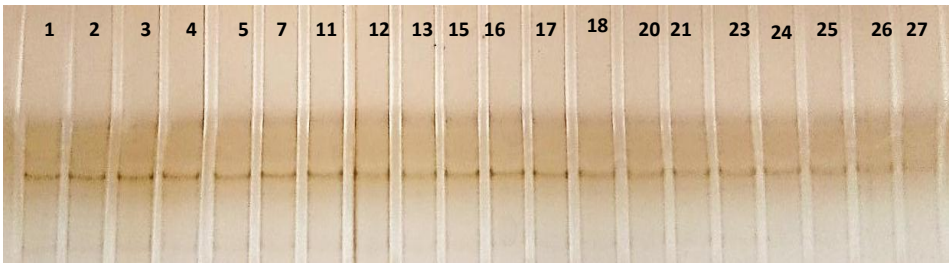


Fig. 4.8: Double stranded PCR-SSCP analysis of exon 16 in *CSN1S2* gene, from genomic DNA of different goats of Marwari breed. For non-denaturing PAGE, 10% polyacrylamide gel with 5% glycerol is used. The electrophoresis was done in 0.5× TBE for 18 h at 4°C at 130 V. A monomorphic band pattern was seen.

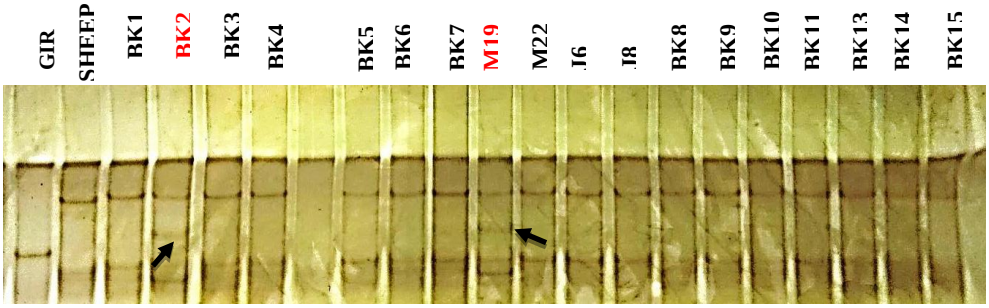


Fig. 4.9: Double stranded PCR-SSCP analysis of goat *CSN1S2* gene, exon 16 (from genomic DNA of different goats of Bakarwali (BK), Marwari (M) and Jakhrana (J) along with Gir (Cattle) & Sheep. For non-denaturing PAGE, 10% polyacrylamide gel with 5% glycerol is used. The electrophoresis was done in 0.5× TBE for 18 h at 4°C at 130 V. Two band patterns were seen. BK2 and M19 are having different pattern of band indicated by arrow mark compared to others but similar to each other).

Further confirmation was done through sequencing for which BK2 and M19 genomic DNA were selected for exon 16 region of genomic DNA amplification. The amplified product of these two goats was present at 166 nt position in 2.5% agarose gel was precisely cut and purified. The purified 166 nt DNA was ligated to linear vector (pTZ57R/T) and then transformed into XL-1 blue MRF' strain of *E.coli*. Ten recombinant colonies of each animal were selected and PCR-SSCP was done. Two types of band pattern were observed, M19.1, M19.2, M19.3, M19.10, BK2.4, BK2.6 and BK2.9 clones have same pattern and M19.4, M19.5, M19.7, M19.8, M19.9, M19.10, BK2.1, BK2.2, BK2.3, BK2.7, BK.8, BK2.10 clones have same pattern (Fig. 4.10). Eight clones (M19.3, M19.5, M19.7, M19.8, BK2.1, BK2.3, BK2.6 and BK2.8) of M19 and BK2 were selected for capillary sequencing showing both the pattern.

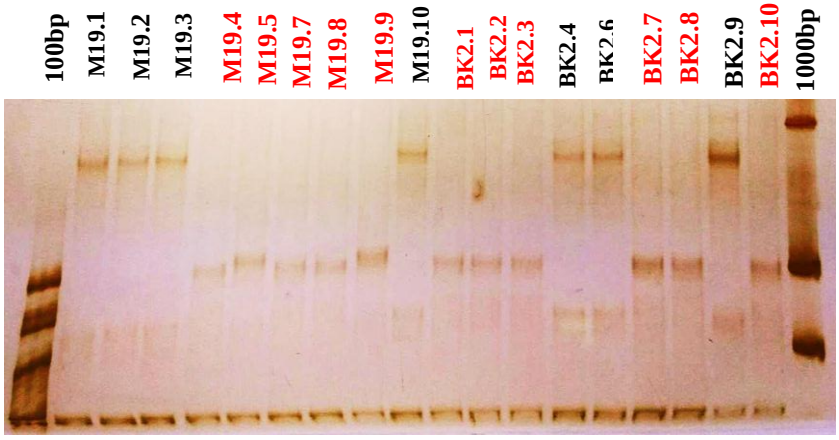


Fig. 4.10: Double stranded PCR-SSCP analysis of goat *CSN1S2* gene, exon 16 from recombinant clone of M19 and BK2 goats after denaturation at 95°C for 5 min, in 10% polyacrylamide gel with 5% glycerol. The electrophoresis was done in 0.5× TBE for 18 h at 4°C at 130 V. Two patterns of band are indicated by the black and red color font.

The nucleotide sequences were edited, translated to amino acid sequences and multiple aligned with the amino acid sequences of exon 16 region of reference alleles that were earlier reported (Accession no: NP_001272514.1, AAB32166.1, CAB59920.1 and CAB94236.1). In two clones (M19.3 and BK2.6), the sequences were similar to that of allele A. In other six clones (M19.5, M19.7, M19.8, BK2.1, BK2.3 and BK2.8) a transversion (C to G) at 623rd nt position of coding region of the exon 16 responsible for substitution of Pro_{193} to Arg_{193} in amino acid sequence which is also reported in allele E (Lagonigro *et al.*, 2001) but these are not similar to E allele because along with the

mentioned transversion another transversion (A to T) should be there at 545th nt position which was not present in these clones.

Therefore, these observation reports a new allele which was found in both Bakarwali and Marwari breeds as these were not matching with any reported alleles. The variation position of these clones was also not found in the seven new alleles (Table 4.4). Therefore, total eight new alleles are found in present study. The alignment of one clone, M19.5 was shown in fig. 4.11 with reference alleles along with the new seven alleles that were found during the study of objective 1.

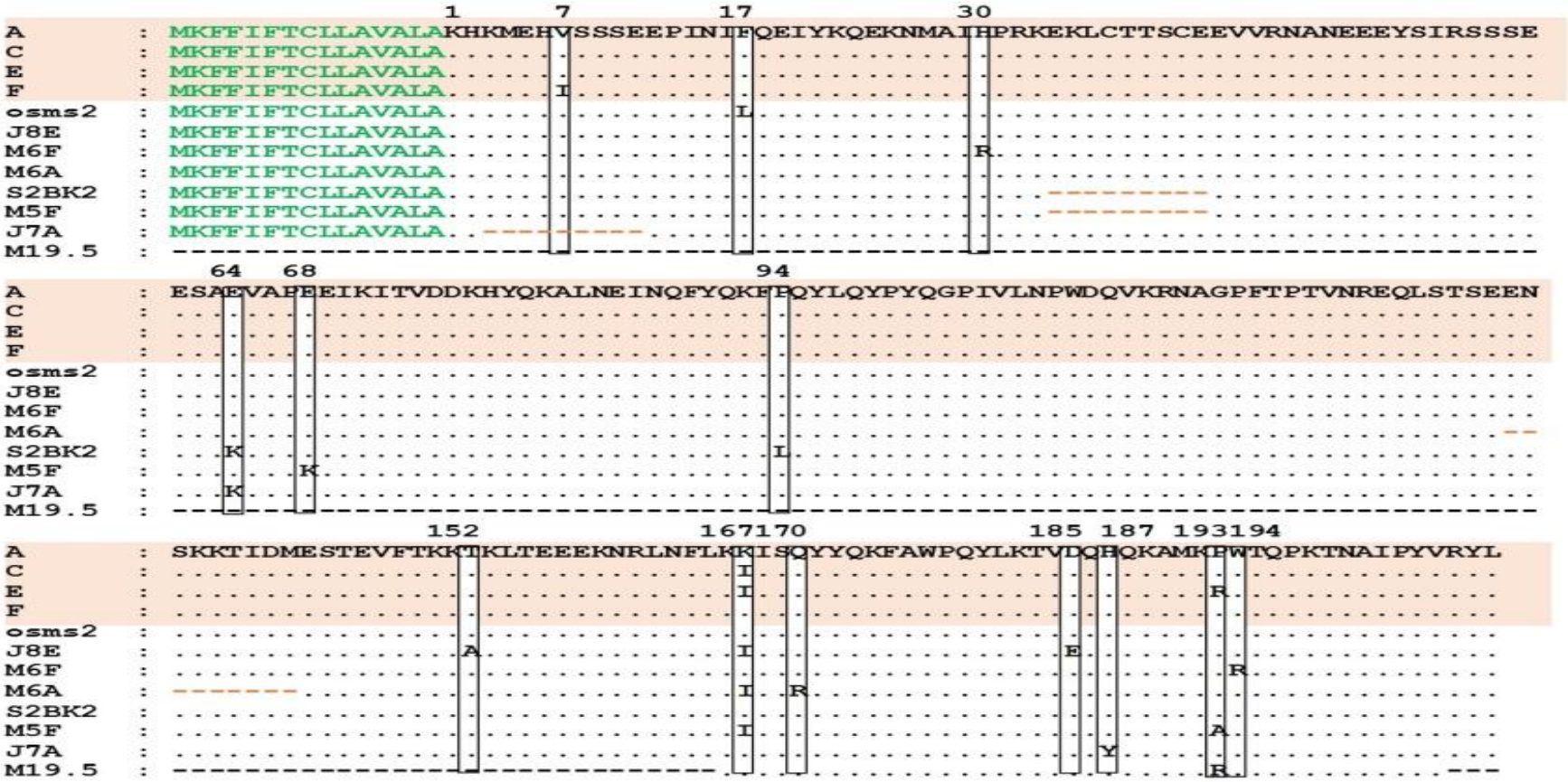


Fig. 4.11: **Alignment analysis of aminoacid sequences of coding region of *CSN1S2* gene of eight novel alleles of Indian goats (**Jakhrana, Marwari, Osmanabadi and Bakarwali Breeds) and exon 16 (M19.5) against the reference sequences NP_001272514.1 (allele A), AAB32166.1 (allele C), CAB59920.1 (allele E) and CAB94236.1 (allele F). The first 15 amino acids in green colour font indicate the signal peptide and the coloured dash (-) lines indicate the exons being skipped. Alignment was performed with ClustalW, and the results are displayed by the GenDoc program. Polymorphisms are shown in the text box and their positions written on top.

CHAPTER –5

Summary and Conclusions

SUMMARY AND CONCLUSION

The present investigation entitled, "Cloning and screening for sequence variation in *CSN1S2* gene in selected Indian goat breeds" was undertaken to study to any possible variation in the sequences in coding region of *CSN1S2* gene.

- The complete coding sequences of six Indian goat breeds namely Bakarwali, Sirohi, Osmanabadi, Marwari, Jakhrana, Chegu and one non-descript goat of Jammu were sequenced.
- Total 19 cDNA were from 19 animals of different Indian goat breeds above mentioned were synthesized, PCR amplified, 190 clones (10 from each animal) were screened for insert of *CSN1S2* gene and 26 clones were sequenced by capillary method.
- Allele A, B & C were found in these breeds of goats which were reported earlier. Allele C was found to be predominant in Indian breeds.
- Seven novel alleles were found due to new variations in nucleotide sequences which were responsible for substitution of amino acid sequences. Three alleles were identified in Marwari, two alleles in Jakhrana, one allele in Bakarwali and in Osmananbadi breed of goats.
- In novel alleles the variations in nucleotide sequences were found in exon 4, 5, 11, 15 and 16 where as in reported alleles the variations were found in exon 3, 9, 11 and 16.
- Most mutations were found in exon 16 in case of novel alleles and also in reported alleles.
- Exon skipping due to alternative splicing was observed. Exon 3, 6, 9 and 13 were deleted in different clones of different breeds namely Bakarwali, Marwari, Jakhran and Chegu. Among these 6^{th} exon found to be most common to be skipped.
- Exon 11 region was screened for presence of deletion (Allele D). Total 140 goats belonging to eleven different breeds (Jakhran, Jamunapari, Sirohi, Ganjam, Black

Bengal, Barbari, Gaddi, Chegu, Osmanabadi, Marwari and Bakarwali) were screened and in no animal deletion were found.

- PCR-SSCP was performed for exon 16 in 60 goats belonging to four different breeds.
- A new variation was found in exon 16 which is responsible for formation of a new allele in Bakarwali and Marwari breeds of goat.

Conclusion:

The present study showed polymorphism in *CSN1S2* gene of Indian goat breeds on sequencing the complete coding sequence of *CSN1S2* gene six breeds namely Bakarwali, Sirohi, Osmanabadi, Marwari, Jakhrana, Chegu and one non-descript goat of Jammu. A, B and C alleles were found in these breeds of goats which were reported earlier and along with them seven novel alleles were identified from Marwari, Jakhrana, Bakarwali and Osmanabadi breeds of goat. Exon skipping due to alternative splicing was also observed in exon 3, 6, 9 and 13 in different clones of different breeds namely Bakarwali, Marwari, Jakhrana and Chegu and 6th exon found to be most common to be skipped. Most mutations were found in exon 16 in case of novel and also in reported alleles. PCR-SSCP was performed for exon 16 in 60 goats belonging to four different breeds. A new variation was found in exon 16 which is responsible for formation of a new allele in Bakarwali and Marwari breeds of goat. Thus in Indian goat breeds, eight new alleles were found for *CSN1S2* gene.

Bibliography

BIBLIOGRAPHY

Acharya, R.M., Misra, R.K. & Patil, V.K. (1982). Breeding Strategy for Goats in India: A Discussion Paper. *Indian Council of Agricultural Research*.

Ahmed, A.S., El-Bassiony, T., Elmalt, L.M. & Ibrahim, H.R. (2015). Identification of potent antioxidant bioactive peptides from goat milk proteins. *Food Research International*, 74, 80-88.

Aurousseau, B., Vermorel, M., Theriez, M., Vezinhet, A., Duboisset, F., Donnat, J.P., Garambois, Y., Gasnet, S., Martinaud, M., Moins, G. & Souchet, R. (1989). Effects of substitution of tricaprylin or coconut oil for tallow in milk replacers offered to preruminant lambs. *In Annales De Zootechnie*. 38, 49-59.

Aziz, M.A. (2010). Present status of the world goat populations and their productivity. *World*, 861(1078.2), 1.

Ballabio, C., Chessa, S., Rignanese, D., Gigliotti, C., Pagnacco, G., Terracciano, L., Fiocchi, A., Restani, P. & Caroli, A.M. (2011). Goat milk allergenicity as a function of αS1-casein genetic polymorphism. *Journal of Dairy Science*, 94(2), 998-1004.

Barłowska, J., Szwajkowska, M., Litwińczuk, Z. & Król, J. (2011). Nutritional value and technological suitability of milk from various animal species used for dairy production. *Comprehensive Reviews In Food Science and Food Safety*, 10(6), 291-302.

Bevilacqua, C., Ferranti, P., Garro, G., Veltri, C., Lagonigro, R., Leroux, C., Pietrola, E., Addeo, F., Pilla, F., Chianese, L. & Martin, P. (2002). Interallelic recombination is probably responsible for the occurrence of a new αs1- casein variant found in the goat species. *European Journal of Biochemistry*, 269(4), 1293-1303.

Bia, R.R., Virgirinia, R.P., Setiawan, B., Soewondo, A. & Fatchiyah, F. (2015). Goat milk CSN1S2 is able to decrease the severity scoring, TNF-a, and RAGE expression in complete Freund's adjuvant-induced rheumatoid arthritis model of rats. *Biomarkers and Genomic Medicine*, 7(2), 64-71.

Boisnard, M., Hue, D., Bouniol, C., Mercier, J.C. & Gaye, P. (1991). Multiple mRNA species code for two non- allelic forms of ovine αs2- casein. *European Journal of Biochemistry*, 201(3), 633-641.

Boulanger, A., Grosclaude, F. & Mahé, M.F. (1984). Polymorphisme des caséines α s1 et α s2 de la chèvre (Capra hircus). *Génétique, Sélection, Evolution*, 16(2), 157.

Bouniol, C., (1993). Sequence of the goat αs2-casein-encoding cDNA. *Gene*, 125(2), 235-236.

Bouniol, C., Brignon, G., Mahé, M.F. & Printz, C. (1994). Biochemical and genetic analysis of variant C of caprine αs2- casein (Capra hircus). *Animal Genetics*, 25(3), 173-177.

Caravaca, F., Carrizosa, J., Urrutia, B., Baena, F., Jordana, J., Amills, M., Badaoui, B., Sánchez, A., Angiolillo, A. & Serradilla, J.M. (2009). Effect of αS1-casein (CSN1S1) and κ-casein (CSN3) genotypes on milk composition in Murciano-Granadina goats. *Journal of Dairy Science*, 92(6), 2960-2964.

Caroli, A., Chiatti, F., Chessa, S., Rignanese, D., Bolla, P. & Pagnacco, G. (2006). Focusing on the goat casein complex. *Journal of Dairy Science*, 89(8), 3178-3187.

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 Sc*ience, 92(11)

Ceballos, L.S., Morales, E.R., de la Torre Adarve, G., Castro, J.D., Martínez, L.P. & Sampelayo, M.R.S. (2009). Composition of goat and cow milk produced under similar conditions and analyzed by identical methodology. *Journal of Food Composition and Analysis*, 22(4), 322-329.

Chessa, S., Budelli, E., Gutscher, K., Caroli, A. & Erhardt, G. (2003). Simultaneous identification of five κ-casein (CSN3) alleles in domestic goat by polymerase chain reaction-single strand conformation polymorphism. *Journal of Dairy Science*, 86(11), 3726-3729.

Chessa, S., Rignanese, D., Chiatti, F., Radeghieri, A., Gigliotti, C. & Caroli, A. (2008). Simultaneous identification of CSN1S2 A, B, C, and E alleles in goats by

polymerase chain reaction-single strand conformation polymorphism. *Journal of Dairy Science*, 91(3), 1214-1217.

Chiatti, F., Chessa, S., Bolla, P., Caroli, A. & Pagnacco, G. (2005). Casein genetic polymorphisms in goat breeds of Lombardy. *Italian Journal of Animal Science*, 4(sup2), 46-48.

Cieslak, J., Pawlak, P., Wodas, L., Borowska, A., Stachowiak, A., Puppel, K., Kuczynska, B., Luczak, M., Marczak, L. & Mackowski, M. (2016). Characterization of equine CSN1S2 variants considering genetics, transcriptomics, and proteomics. *Journal of Dairy Science*, 99(2), 1277-1285.

Claeys, W.L., Verraes, C., Cardoen, S., De Block, J., Huyghebaert, A., Raes, K., Dewettinck, K. & Herman, L. (2014). Consumption of raw or heated milk from different species: An evaluation of the nutritional and potential health benefits. *Food Control*, 42, 188-201.

Coll, A., Folch, J.M. & Sanchez, A. (1995). Structural features of the 5'flanking region of the caprine κ-casein gene. *Journal of Dairy Science*, 78(5), 973-977.

Dixit, S.P., Verma, N.K., Aggarwal, R.A.K., Vyas, M.K., Rana, J. & Sharma, A. (2012). Genetic diversity and relationship among Indian goat breeds based on microsatellite markers. *Small Ruminant Research*, 105(1-3), 38-45.

Erhardt, G., Jäger, S., Budelli, E. & Caroli, A. (2002). Genetic polymorphism of goat αs2-casein (*CSN1S2*) and evidence for a further allele. *Milchwissenschaft*, 57(3), 137-140.

FAOSTAT, F., 2017. Available online: http://www. fao. org/faostat/en/# data. QC (accessed on January 2018).

Ford, E. B. (1940). Polymorphism and Taxonomy. *The New Systematics*. 493–513.

Garg, N., Singh, S.K., Rout, P.K. & Mandal, A. (2009). Genetic polymorphism of milk proteins in Barbari goats. *Tropical and Subtropical Agroecosystems*, 11(1).

Gasser, R.B., Hu, M., Chilton, N.B., Campbell, B.E., Jex, A.J., Otranto, D., Cafarchia, C., Beveridge, I. & Zhu, X. (2006). Single-strand conformation polymorphism (SSCP) for the analysis of genetic variation. *Nature Protocols*, 1(6), 3121.

Gautam, D., Vats, A., Verma, M., Rout, P.K., Meena, A.S., Ali, M., Deepika, S. & De, S. (2019). Genetic variation in CSN3 exon 4 region of Indian goats and a new nomenclature of CSN3 variants. *Animal Genetics*, 50(2), 191-192.

Grobler, R., Visser, C., Chessa, S. & van Marle-Köster, E. (2017). Genetic polymorphism of CSN1S2 in South African dairy goat populations. *South African Journal of Animal Science*, 47(1), 72-78.

Groenen, M.A. & van der Poel, J.J. (1994). Regulation of expression of milk protein genes: a review. *Livestock Production Science*, 38(2), 61-78.

Guan, D., Mármol-Sánchez, E., Cardoso, T.F., Such, X., Landi, V., Tawari, N.R. & Amills, M. (2019). Genomic analysis of the origins of extant casein variation in goats. *Journal of Dairy Science*.

Gupta, S.C., Kumar, D., Pandey, A., Malik, G. & Gupta, N. (2009). New κ-casein alleles in Jakhrana goat affecting milk processing properties. *Food Biotechnology*, 23(1), 83-96.pp.5335-5352.

Haenlein, G.F.W. (2004). Goat milk in human nutrition. *Small Ruminant Research*, 51(2), 155-163.

Hatziminaoglou, Y. & Boyazoglu, J. (2004). The goat in ancient civilisations: from the Fertile Crescent to the Aegean Sea. *Small Ruminant Research*, 51(2), 123-129.

Hayashi, K. (1991). PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods and Applications*. 1, 34-38.

Ibrahim, H.R., Ahmed, A.S. & Miyata, T. (2017). Novel angiotensin-converting enzyme inhibitory peptides from caseins and whey proteins of goat milk. *Journal of Advanced Research*, 8(1), 63-71.

India, M., 2012. 19th Livestock Census—2012: All India Report.

Jenness, R. (1980). Composition and characteristics of goat milk: review 1968− 1979. *Journal of Dairy Science*, 63(10), 1605-1630.

Joshi, M.B., Rout, P.K., Mandal, A.K., Tyler-Smith, C., Singh, L. & Thangaraj, K. (2004). Phylogeography and origin of Indian domestic goats. *Molecular Biology and Evolution*, 21(3), 454-462.

Jung, T.H., Yun, S.S., Lee, W.J., Kim, J.W., Ha, H.K., Yoo, M., Hwang, H.J., Jeon, W.M. & Han, K.S. (2016). Hydrolysis by alcalase improves hypoallergenic properties of goat milk protein. *Korean Journal for Food Science of Animal Resources*, 36(4), 516.

Kawasaki, K., Lafont, A. & Sire, J. (2011). The evolution of milk casein genes from tooth genes before the origin of mammals. *Molecular Biology and Evolution*, 28(7), 2053-2061.

Keen, T.J. & Inglehearn, C.F. (1996). Mutations and polymorphisms in the human peripherin- RDS gene and their involvement in inherited retinal degeneration. *Human Mutation*, 8(4), 297-303.

Lagonigro, R., Pietrola, E., D'Andrea, M., Veltri, C. & Pilla, F. (2001). Molecular genetic characterization of the goat s2- casein E allele. *Animal Genetics*, 32(6), 391-393.

Lahiri, D. K., & Nurnberger Jr, J. I. (1991). A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Research*, 19(19), 5444.

Leroux, C. & Martin, P. (1996). The caprine αs1-and β-casein genes are 12-kb apart and convergently transcribed. *Animal Genetics-Supplement*, 27.

Leroux, C., Mazure, N. & Martin, P. (1992). Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat alpha s1-casein transcripts. Structural organization of the relevant gene. *Journal of Biological Chemistry*, 267(9), 6147-6157.

Leyton, J., Drury, P.J. & Crawford, M.A. (1987). Differential oxidation of saturated and unsaturated fatty acids in vivo in the rat. *British Journal of Nutrition*, 57(3), 383-393.

Luikart, G., Gielly, L., Excoffier, L., Vigne, J.D., Bouvet, J. & Taberlet, P. (2001). Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proceedings of the National Academy of Sciences*, 98(10), 5927-5932.

Malacarne, M., Martuzzi, F., Summer, A. & Mariani, P. (2002). Protein and fat composition of mare's milk: some nutritional remarks with reference to human and cow's milk. *International Dairy Journal*, 12(11), 869-877.

Mandal, A., Karunakaran, M., Rout, P.K. & Roy, R. (2014). Conservation of threatened goat breeds in India. *Animal Genetic Resources*, 55, 47-55.

Manthani, V., Keerthi, S., & Rao, P. (2018). Comparative Study on Chemical Characteristics of Cow Milk and Goat Milk, Kefir. *International Journal of Pure And Applied Bioscience*, 6 (1), 910-914

Marletta, D., Criscione, A., Bordonaro, S., Guastella, A.M. & D'Urso, G. (2007). Casein polymorphism in goat's milk. *Le Lait*, 87(6), 491-504.

Martin, P., Szymanowska, M., Zwierzchowski, L. and Leroux, C. (2002). The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reproduction Nutrition Development*, 42(5), 433-459.

Matéos, A., Miclo, L., Mollé, D., Dary, A., Girardet, J.M. & Gaillard, J.L. (2009). Equine αS1-casein: Characterization of alternative splicing isoforms and determination of phosphorylation levels. *Journal of Dairy Science*, 92(8), 3604-3615.

Matsuo, T. & Takeuchi, H. (2004). Effects of structured medium-and long-chain triacylglycerols in diets with various levels of fat on body fat accumulation in rats. *British Journal of Nutrition*, 91(2), 219-225.

Michaud, J., Brody, L.C., Steel, G., Fontaine, G., Martin, L.S., Valle, D. & Mitchell, G., (1992). Strand-separating conformational polymorphism analysis: efficacy of detection of point mutations in the human ornithine δ-aminotransferase gene. *Genomics*, 13(2), 389-394.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Scie*nces, 86(8), 2766-2770.

Park, Y.W. (2017). Goat milk–chemistry and nutrition. *Handbook of Milk of Non- Bovine Mammals*, 42-83.

Potočnik, K., Gantner, V., Kuterovac, K. & Cividini, A. (2011). Mare's milk: composition and protein fraction in comparison with different milk species. *Mljekarstvo: Casopis za Unaprjeđenje Proizvodnje i Prerade Mlijeka*, 61(2), 107-113.

Prosser, C.G., McLaren, R.D., Frost, D., Agnew, M. & Lowry, D.J. (2008). Composition of the non-protein nitrogen fraction of goat whole milk powder and goat milk-based infant and follow-on formulae. *International Journal of Food Sciences and Nutrition*, 59(2), 123-133.

Rachagani, S. & Gupta, I.D. (2008). Bovine kappa-casein gene polymorphism and its association with milk production traits. *Genetics and Molecular Biology*, 31(4), 893-897.

Ramunno, L., Cosenza, G., Pappalardo, M., Longobardi, E., Gallo, D., Pastore, N., Di Gregorio, P. & Rando, A. (2001(b)). Characterization of two new alleles at the goat CSN1S2 locus. *Animal Genetics*, 32(5), 264-268.

Ramunno, L., Cosenza, G., Rando, A., Pauciullo, A., Illario, R., Gallo, D., Di Berardino, D. & Masina, P. (2005). Comparative analysis of gene sequence of goat CSN1S1 F and N alleles and characterization of CSN1S1 transcript variants in mammary gland. *Gene*, 345(2), 289-299.

Ramunno, L., Longobardi, E., Pappalardo, M., Rando, A., Di Gregorio, P., Cosenza, G., Mariani, P., Pastore, N. & Masina, P. (2001(a)). An allele associated with a non- detectable amount of αs2 casein in goat milk. *Animal Genetics*, 32(1), 19-26.

Rijnkels, M. (2002). Multispecies comparison of the casein gene loci and evolution of casein gene family. *Journal of Mammary Gland Biology and Neoplasia*, 7(3), 327-345.

Rout, P.K., Kumar, A., Mandal, A., Laloe, D., Singh, S.K. & Roy, R. (2010). Characterization of casein gene complex and genetic diversity analysis in Indian goats. *Animal Biotechnology*, 21(2), 122-134.

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. *Journal of Dairy Science*, 88(4), 1561-1568.

Scott, R. (1986). Cheese making Practice. Elsevier, Amsterdam, *The Netherlands*. 529.

Sharma, G., Rout, P. K., & Singh, G. (2017). Characterization of goat milk protein and comparison of milk proteins using polyacrylamide gel electrophoresis. *International Journal of Scientific Engineering and Applied Science*. 3(4), 13-18.

Simos, Y., Metsios, A., Verginadis, I., D'Alessandro, A.G., Loiudice, P., Jirillo, E., Charalampidis, P., Kouimanis, V., Boulaka, A., Martemucci, G. & Karkabounas, S. (2011). Antioxidant and anti-platelet properties of milk from goat, donkey and cow: An in vitro, ex vivo and in vivo study. *International Dairy Journal*, 21(11), 901-906.

Smith, C.W., Chu, T.T. & Nadal-Ginard, B. (1993). Scanning and competition between AGs are involved in 3'splice site selection in mammalian introns. *Molecular and Cellular Biology*, 13(8), 4939-4952.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. & Soreq, H. (2005). Function of alternative splicing. *Gene*, 344, 1-20.

Stewart, A.F., Bonsing, J., Beattie, C.W., Shah, F., Willis, I.M. & Mackinlay, A.G. (1987). Complete nucleotide sequences of bovine alpha S2-and beta-casein cDNAs: comparisons with related sequences in other species. *Molecular Biology and Evolution*, 4(3), 231-241.

Strzałkowska, N., Jóźwik, A., Bagnicka, E., Krzyżewski, J., Horbańczuk, K., Pyzel, B., Słoniewska, D. & Horbańczuk, J.O. (2010). The concentration of free fatty acids in goat milk as related to the stage of lactation, age and somatic cell count. *Animal Science Papers and Reports*, 28(4), 389-395.

Vacca, G.M., Ali, H.O.A.B., Carcangiu, V., Pazzola, M. & Dettori, M.L., (2009). Genetic structure of the casein gene cluster in the Tunisian native goat breed. *Small Ruminant Research*, 87(1-3), 33-38.

Velázquez, O.C., Seto, R.W. & Rombeau, J.L. (1996). The scientific rationale and clinical application of short-chain fatty acids and medium-chain triacylglycerols. *Proceedings of the Nutrition Society*, 55(1B), 49-78.

Verma, N.K., (2018). Goat genetic diversity of India-An overview. *Journal of Livestock Biodiversity*

Yue, X.P., Fang, Q., Zhang, X., Mao, C.C., Lan, X.Y., Chen, H. & Lei, C.Z. (2013). Effects of *CSN1S2* genotypes on economic traits in Chinese dairy goats. *Asian-Australasian Journal of Animal Sciences*, 26(7), 911.

Appendix

APPENDIX-1

Nucleotide	**Symbols**
Adenine	A
Guanine	G
Thymine	T
Cytosine	C
Amino acid	**Three letter abbreviation**
Alanine	Ala
Arginine	Arg
Asparagine	Asn
Aspartic acid	Asp
Cysteine	Cys
Glutamic acid	Glu
Glutamine	Gln
Glycine	Gly
Histidine	His
Isoleucine	Ile
Leucine	Leu
Lysine	Lys
Methionine	Met

Phenylalanine	Phe
Proline	Pro
Serine	Ser
Threonine	Thr
Tryptophan	Trp
Tyrosine	Tyr
Valine	Val