Casein Loci SNPs and Haplotypes Association with Milk Production Traits of Norwegian Goats and their French Alpine Crosses

Takele Taye Desta

Norwegian University of Life Sciences

Department of Animal and Aquaculture

Master Thesis 30 Credits 2010

European Master in Animal Breeding and Genetics

Casein loci SNPs and haplotypes association with milk production traits of Norwegian goats and their French Alpine crosses

Takele Taye Desta

THESIS ANIMAL BREEDING AND GENETICS (M30-IHA)

May 2010

Norwegian University of Life Sciences (UMB)

SUPERVISORS

Tormod Ådnøy (Associate Prof., UMB)

Binyam Sime Dagnachew (PhD Student, UMB)

Education and Culture

Erasmus Mundus

Table of contents

List of Tables

List of Figures

Acknowledgements

First of all I deeply and humbly acknowledge THE ALMIGHTY GOD, Who makes all possible and Who provides me the courage and the stamina to complete my thesis project. My profound appreciation and sincere thanks is for my major supervisor Tormod Ådnøy for his painstaking effort and unreserved support that has brought my research project to this end. My co-supervisor Binyam Sime Dagnachew deserves my due appreciation for his consistent support throughout my research work.

The support of Jørgen Ødegård in ASReml analysis is duly acknowledged. I would like to acknowledge Inger Anne Boman for providing SAS command that used for pedigree construction and Megumi Ohta Fog for facilitating my study at UMB. Most of the dataset were obtained from TINE. The comparative study was based on the buck circle and crossbreeding program supported by NSG. Nine goat farms anonymous owners participated in data collection of test day milk yield together with TINE agents. Helga Kvamsås is acknowledged for giving information on farm location.

European Union Erasmus Mundus program is acknowledged for its generous scholarship award and EM-ABG consortium for providing this study opportunity. One year ASReml software license used for this analysis was purchased from the quality goat milk for cheese project fund of the Norwegian Research Council.

I would like to appreciate Dagim Jirata, Tsehay Mekonnen and Aster Abebe for creating a family hood environment during my stay in Ås. I would like to acknowledge all my family members in Ethiopia for their encouragement and moral support.

Acronyms and abbreviations

αS1-CN	αS_1-casein
αS2-CN	αS_2-casein
β-CN	β-casein
κ-CN	κ-casein
bp	base pair
Ca	Calcium
CN	Casein
CSN1S1	αS_1-casein coding gene
CSN1S2	αS_2-casein coding gene
CSN2	β-casein coding gene
CSN3	κ-casein coding gene
FAO	Food and Agriculture Organization of the United Nations
hp	haplotype
kb	Kilo base pair
kg	Kilo gram
LD	Linkage Disequilibrium
MALDI–TOF–MS	Matrix Assisted Laser Desorption Ionization Time–of–Flight Mass Spectrometry
NCBI	The National Center for Biotechnology Information
SAS	Statistical Analysis System
SNP	Single Nucleotide Polymorphism
SCC	Somatic Cell Count

Abstract

Due to their quantitative and qualitative implications on milk yield and its component traits, casein loci have been and would remain the subject of intense research. Our work was also intended to resolve haplotype structures and diversity, and to study the additive and dominance effects of casein SNPs and SNP haplotypes on milk production traits in the Norwegian goats and their French Alpine crosses (crosses). Our analysis used 376 does (216 Norwegian goats and 160 crosses) that have phenotypic records on test day milk yield and genotype data. Linkage disequilibrium (LD) resolution and haplotype construction were done using 38 SNP markers detected in the entire casein loci block of the two goat populations. Our result showed that, there was extensive LD, especially for αS1- and κ-casein loci; however, the LD is relatively weak in the crosses. The extent of LD varies across the casein loci segment from nearly zero to almost complete LD. The intra-locus LD was stronger than the LD found for inter-loci. Due to extensive LD, the numbers of plausible haplotypes constructed were by far less than what were expected. The diversity of plausible haplotypes is high for the crosses especially for β- and κ-casein. Therefore, more tagger SNPs were detected for the crosses. Our results from SNP halpotypes additive and dominance fixed effect analysis of the entire casein loci and of the individual casein genes showed significant effect on studied milk production traits except lactose percentage. Therefore, this genetic variation observed among haplotypes can be used in the genetic improvement program of Norwegian goats through haplotype assisted selection. We also found significant effect of some casein SNPs on milk production traits, especially for αS2- and κ-casein SNPs. The SNP effect is, however, localized within a segment of adjacent SNPs even at locus level. SNP 14 deletion (D) of exon12 in αS1-casein has significant additive genetic effect on FFA and urea content of the milk and a significant dominance effect of AD genotype on SCC. Our analysis showed that the frequency of D allele for Norwegian goats has reduced to 0.66, which showed that the selection program is against this allele. Among studied traits, the Norwegian goats and the crosses significantly differ for lactose percentage only. However, the crosses were as good as the Norwegian goats in the rest of the traits, which indicates the importance of this crossbreeding program beyond creating genetic variation. Therefore, comparative study at pure breed level would be a plausible option to substantiate whether this difference is due to heterotic effect or real genetic difference between breeds.

Norsk sammendrag

På grunn av deres kvantitative og kvalitative effekter på melkeproduksjonen og melkekomponenter, har kaseinloci vært og vil fortsatt være gjenstand for intens forskning. Vårt arbeid var også ment å studere haplotypestrukturer og –diversitet og additive og dominante effekter av kasein-SNPer og haplotyper på melkeproduksjonstrekk i norske geiter og deres franske alpinkryssinger (kryssinger). Vår analyse brukte 376 geiter (216 norske geiter og 160 kryssinger) som har fenotypiske registreringer på test dagen av melkeproduksjonen og genotypedata slik at en kunne beregne faste additive- og dominans-effekter av SNPer og haplotyper på studerte trekk. Analyse av koplingsulikevekt (LD) og haplotypekonstruksjon ble gjort med 38 SNP markører i alle kaseinloci for de to geitepopulasjoner. Våre resultater viste at det er omfattende koplingsulikevekt, spesielt for αS1- og κ-kasein, men LD er relativt svak for kryssingene. Omfanget av LD varierer fra null til fullstendig mellom loci. Kopling innen locus var sterkere enn på tvers av loci. På grunn av omfattende LD var antallet plausible konstruerte haplotyper langt mindre enn forventet antall. Mangfoldet av plausible haplotyper er høyt for kryssingene særlig for β- og κ-kasein. Derfor ble mer tagger-SNPer observert i kryssingene. Våre resultater fra analyse av SNP-haplotypenes additive- og dominans- fast effekt og av effektene av hele kaseinloci og individuelle kaseingener viste signifikante effekt på melkeproduksjonen og dens komponenter bortsett fra laktoseprosent. Derfor kan denne genetiske variasjonen blant haplotypene brukes i det genetiske forbedringsprogrammet for norske geiter gjennom haplotypeassistert seleksjon. Vi fant også signifikant effekt av noen kasein-SNPer på melkeproduksjonsegenskaper, spesielt for αS2- og κ-kasein. Denne SNP-effekten er imidlertid lokalisert innenfor et segment av tilstøtende SNPer innen ett locus. SNP 14 delesjon (D) i exon12 i αS1-kasein har betydelig additiv genetisk effekt på FFA og urea-innholdet i melk og det er en betydelig dominans effekt av AD-genotypen på SCC. Analysen av vårt materiale viste at frekvensen av D-allelet er redusert til 0,66, noe som viser at seleksjonsprogrammet virker mot dette allelet. Blant de studerte egenskaper hadde de norske geitene og kryssingene betydelige forskjeller bare for laktoseprosent. Men kryssingene er like gode som de norske geitene for resten av egenskapene, - noe som viser betydningen av dette kryssingsprogrammet utover det å skape genetisk variasjon. Derfor ville sammenlignende studier på renrasenivå være et alternativ for å dokumentere hvorvidt denne forskjellen skyldes heteroseeffekt eller skyldes reelle genetiske forskjeller mellon rasane.

1. Introduction

Goat milk is increasingly a valuable source of protein in many African, Asian and European countries (Hayes *et al* 2006). As a result; the world goat population has increased by 66%, and parallelly goat milk and cheese production have increased from 1.7 to 2.5 million tons and 132 to 180 thousand tons from 1985 to 2005, respectively (Dubeuf and Boyazoglu 2009). Europe is a home for 4% (26, 092, 000) of the world's 710 million goat population (Scherf 2000) and produces 18% of the total world goat milk (Dubeuf 2005). Dairy goat farming is an ancient practice in Norway (Dubeuf *et al*. 2004) and accounts for 99% of goats' use and 1% of the country's milk production (Sæther 2002).

Milk is probably the best known food with respect to its biosynthesis and composition, and the chemical structure of its individual components (Martin and Grosclaude 1993). Milk proteins are classified into two major groups caseins (*αS1-* , *αS2-*, *β-* , and *κ*-CN) and whey proteins (Trujillo *et al* 2000). Caseins are the major milk proteins and in ruminants, the four caseins represent about 80% of the milk proteins (Ramunno *et al* 2004; Sulimova *et al* 2007; Sztankóová *et al* 2009) and they account for 95% of the milk protein together with whey proteins (Martin and Grosclaude 1993; Martin *et al* 2002). The four caseins: *CSN1S1, CSN2, CSN1S2,* and *CSN3*, respectively coding for proteins *αS1*-CN, *β*-CN, *αS2*-CN, and *κ*-CN and are encoded by 4 closely linked genes (Threadgill and Womack 1990; Martin and Grosclaude 1993; Martin *et al* 2002; Caroli *et al* 2006; Caroli *et al* 2007; Gigli *et al* 2008).

These evolutionarily related casein genes; the so-called "Calcium-sensitive" (Ca-sensitive) caseins *α*-s1 (*CSN1S1* or *Csna*), *β*- (*CSN2* or *Csnb*), and *α*-s2 (*CSN2*S2, A and B, or *Csng* and *Csnd*)) and the physically and functionally linked *κ*-casein (*CSN3* or *Cnsk*) gene (Rijnkels 2002) in that order, located within 250kb genomic DNA region of caprine chromosome 6 (Martin and Grosclaude 1993; Martin *et al* 2002; Marletta *et al* 2007; Vacca *et al* 2009). Figure 1 showed the graphical representation of the casein genes cluster. As a result casein genes are usually inherited from parents to progeny as haplotype (Caroli *et al* 2007).

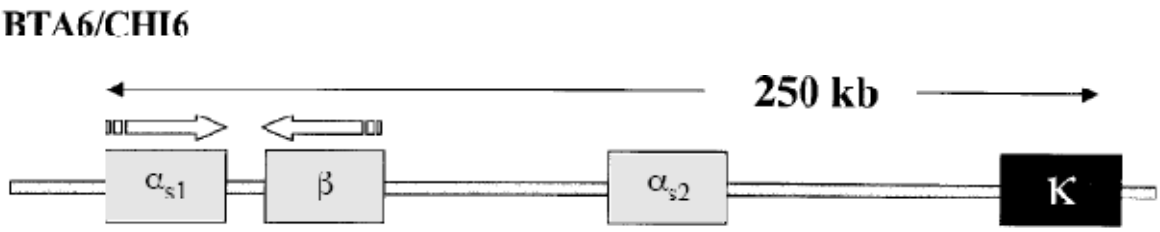

Figure 1. Genomic organization of the bovine/goat casein loci.
(Source: Martin *et al* 2002)

The analysis of caseins in goat species is quite complex due to large number of mutations involving the four coding genes (Caroli *et al* 2006; Caroli *et al* 2007). Therefore, extensive genetic polymorphisms in caprine casein have been the focus of research due to their effects on milk production taits, milk quality and milk composition, dairy performance, and technological properties of milk (Lundén *et al* 1997; Martin *et al* 1999; Marletta *et al* 2007; Gigli *et al* 2008; Kevorkian *et al* 2009). Many of the processing properties of milk are also a function of its structure and relative concentrations of its protein components (Martin and Grosclaude 1993).

The proportion of milk protein components show individual variation due to environmental and genetic factors (Lundén *et al* 1997; Marletta *et al* 2005). Genetic variation is caused by alleles associated with differences in the level of expression (Marletta *et al* 2005). Therefore, genetic variation in milk yield and its component traits can be studied at molecular level using SNPs and/or haplotype information. However, the tendency for haplotype analysis increased power over SNP analysis because SNPs most likely affect phenotypes through their joint effects (Templeton *et al* 2005), even though, the effect of individual SNP on milk production traits was also reported (Hayes *et al* 2006).

Moreover, due to tight linkage among casein genes (Ådnøy *et al* 2006; Caroli *et al* 2006), the variability of the whole haplotype has to be considered when analyzing the goat caseins (Caroli *et al* 2006). For example; the caprine αS1 and beta-casein genes are 12-kb apart (Leroux and Martin 1996) and they might be convergently transcribed and the alleles of these genes can be inherited as haplotype (Leroux and Martin 1996; Marletta *et al* 2005). Therefore, the study of casein haplotypes could provide more information than the study of individual

casein genes (Hayes *et al* 2006; Gigli *et al* 2008). This shows that haplotypes information can be used for genetic selection (Marletta *et al* 2005; Gigli *et al* 2008) and for optimal estimation of genetic relationship among breeds (Moioli *et al* 2007). Therefore, casein haplotypes could be used in the selection program instead of a single locus (Sacchi *et al* 2005). There are a number of papers dealing with the wide heterogeneity of the individual casein genes; however, studies on the combined effect of the casein loci are relatively limited (Marletta *et al* 2007).

Comparative analysis of the casein gene cluster region shows the unusual high divergence of the casein genes coding regions (Rijnkels 2002). Therefore, having the knowledge of genetic polymorphism at goat casein loci allows the set up of breeding plans targeted in the improvement of milk production traits (Feligini *et al* 2005). Compared to sheep, much has been done on casein genes polymorphism in goats. However, it remains a focus of further research due to its high polymorphic nature. The genetic variant of milk protein is a heritable trait and they differ from breed to breed in their occurrence and frequency (Garg *et al* 2009). For example, the relative frequencies of the αS1-casein alleles show marked differences between breeds (Moatsou *et al* 2004).

Moreover, exon12 deletion at αS1-casein creates the unique multi-allelic SNP in Norwegian goats (Hayes *et al* 2006). This deletion is segregating in the population with high frequency (Ådnøy *et al* 2003; Hayes *et al* 2006) inspite of its negative effect on dry matter content of the milk. Surprisingly, this is against the objective of the national goat breeding program (Hayes *et al* 2006). This might indicate the special importance of this unique deletion, which requires further studies.

The national goat and sheep breeding program carried out a crossbreeding of Norwegian goats by using French Alpine semen in 2007. This crossbreeding program is aimed at testing the influence of crossbreeding on milk production traits of the Norwegian goats in selected flocks. Therefore; the effect of this crossbreeding program on milk production traits requires comparative studies between Norwegian goats and the crosses. This is because crossbreeding is considered to be one of the practical ways of improving economically important traits in goats (Shrestha and Fahmy 2007) and it can be used for the exploitation of breed differences through attainable heterotic advantages (Rincon *et al* 1982). In response to this concern, our study has tried to address the following objectives:

1) To study the extent of linkage disequilibrium (LD) among casein genes SNPs in the Norwegian goats and their French Alpine crosses;
2) To study haplotypes diversity in the Norwegian goats and their French Alpine crosses;
3) To study casein genes' SNPs additive and dominance effect on milk production traits of Norwegian goats and their French Alpine crosses;
4) To study casein genes' haplotypes additive and dominance effect on milk production traits of Norwegian goats and their French Alpine crosses;
5) To study the effect of αS1-casein exon 12 deletion on milk yield and its component traits and
6) To study the effect of crossbreeding on milk production traits.

2. Review of Literatures

Caseins are a group of acidic, proline-rich, phosphoproteins aggregated to form large, spherical, micellar structures in colloidal suspension with calcium phosphate in milk (Rijnkels 2002; Marletta *et al* 2007). Caseins are the main source of amino acids, calcium, and phosphate and provide several bioactive peptides (Rijnkels 2002). For example, the concentration of calcium and phosphate in milk is highly correlated with milk casein content (Chanat *et al* 1999).

2.1. Casein polymorphism

A number of genetic variants of the casein genes that affect milk production traits have been described (Hayes *et al* 2006). High polymorphism has been found at the 4 genes *CSN1S1, CSN2, CSN1S2,* and *CSN3* within the goat CN cluster (Caroli *et al* 2007) and research has continued to reveal the extensive casein polymorphism (Marletta *et al* 2007). As a result, the genetic polymorphism of the casein fraction and the chemical structure of its individual components are well documented (Martin *et al* 1999) especially for alpha s1 casein (Raynal–Ljutovac *et al* 2005).

The high molecular divergence in the caseins appears to have resulted from both the variation in the splicing patterns of exons (Chessa *et al* 2007; Marletta *et al* 2007) and due to point mutations involving insertion/deletion (Martin and Grosclaude 1993; Marletta *et al* 2007). It also can be resulted from amino acid substitutions (Martin and Grosclaude 1993), evolutionary divergence (Rijnkels 2002) and by posttranslational modifications (Martin *et al* 2002; Chessa *et al* 2007), in addition to environmental effects (Chessa *et al* 2007). Casein heterogeneity in milk can also be caused by post-translational modifications (Neveu *et al* 2002; Marletta *et al* 2007), such as different levels of phosphorylation and glycosylation (Marletta *et al* 2007). Therefore, the qualitative and quantitative variability of goat caseins originates from the high level of genetic polymorphisms (Chessa *et al* 2007). However, most of the casein alleles differ from each other by a few base substitutions that cause one or two amino acid changes in the protein (Lien *et al* 1995). The structural organization of the casein genes is presented in Figure 2.

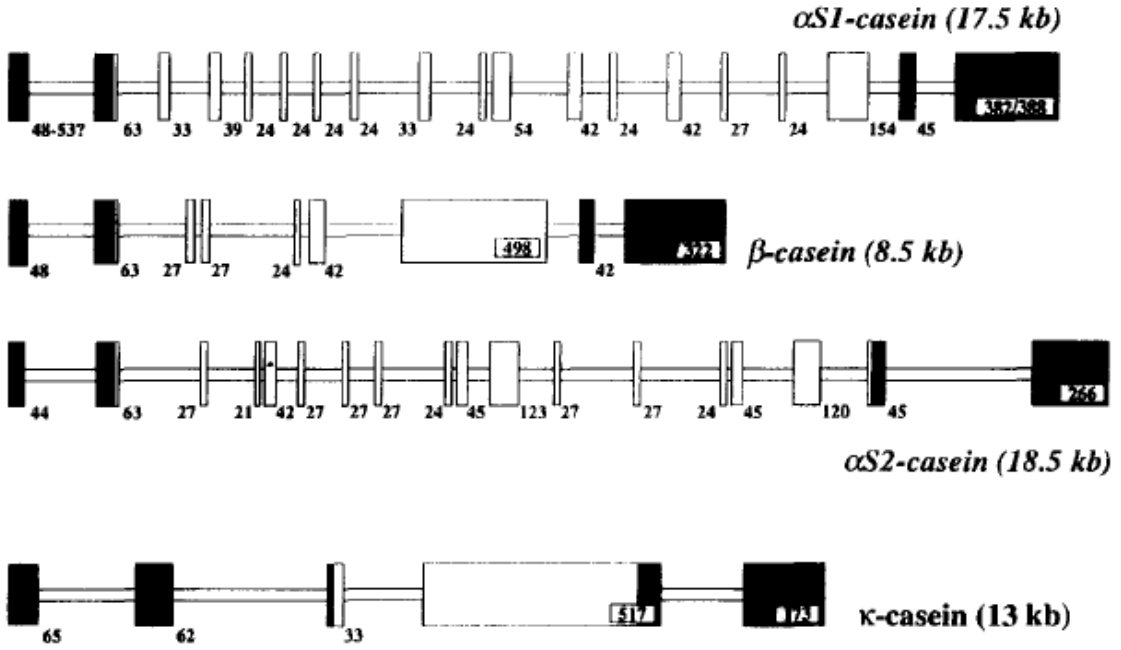

Figure 2. Structural organization of the four bovine casein transcription units.

Open bars in Figure 2 represent introns, and exons are depicted by large, stippled (5' and 3' uncoding regions), black (part of exons encoding the signal peptide) and open (exons and part of exons encoding matured proteins) boxes. Sizes of exons are indicated in base pairs (Source: Martin and Grosclaude 1993).

2.2. Goats' casein

Like thier ruminant counterparts, caprine casein also composed of four genes: αS1- , *β*- , αS2- and κ-casein.

2.2.1. αS1-casein, CSN1S1

The goat CSN1S1 gene extends over 16785bp including 1138bp of exonic regions and 15647bp of intronic regions. It contains 19 exons ranging in size from 24 (exons 5, 6, 7, 8, 10, 13 and 16) to 385 bp (exon 19) and 18 introns from 90 bp (intron 10) to 1685 bp (intron 2) (Ramunno *et al* 2004). *CSN1S1* (casein *alpha-S1*) is the main calcium sensitive casein in ruminant milk (Sztankóová *et al* 2007), which exhibits a high degree of unusual polymorphism (Neveu *et al* 2002; Moatsou *et al* 2004). As many as 18 different alleles have been identified for CSN1S1 (Hayes *et al* 2006). Therefore, *CSN1S1* is distingushed and

characterized by high quantitative and qualitative variation that have qualitative and quantitative implications (Trujillo *et al* 2000; Feligini *et al* 2005; Chiatti *et al* 2007; Sztankóová *et al* 2007).

Based on the milk content of casein alpha-s1, the CSN1S1 variants can be classified into 4 groups: strong alleles (A, B1, B2, B3, B4, C, H, L, and M) producing almost 3.5 g/l of casein alpha-S1 each; intermediate alleles (E and I; 1.1 g/l); weak alleles (D, F and G; 0.45 g/l); and null alleles (01, 02, and N) producing no αS1-casein (Sztankóová *et al* 2007) and the references therein. Therefore, milk produced by goats with different CSN1S1 genotypes shows a variable amount of αS1-casein, ranging from 7 $g \cdot L^{-1}$ in strong allele homozygous goats, to 0.9 $g \cdot L^{-1}$ and 0 $g \cdot L^{-1}$ in weak and null homozygotes, respectively (Marletta *et al* 2007).

2.2.2. β-casein, CSN2

β-casein can represent up to 60% of total casein in goat milk (Neveu *et al* 2002). The CSN2 gene is smaller than the other two Ca-sensitive casein genes, consisting of 9 exons ranging from 24 to 492 bp (Marletta *et al* 2007). The β-casein gene structure has undergone fewer duplications and rearrangements than the other *α*-ike casein genes (Rijnkels 2002) and most likely due to this it has long been considered to be monomorphic (Neveu *et al* 2002; Marletta *et al* 2007). Greater conservation of the β-casein gene might be due to its proposed function in determining certain structural properties of the casein micelle (Rijnkels 2002).

2.2.3. αS2-casein, CSN1S2

Alphas2-casein is the most phosphorylated casein (Trujillo *et al* 2000). In goats CSN1S2 locus at least seven alleles characterized by different levels of expression have been identified (Marletta *et al* 2004; Chessa *et al* 2007; Vacca *et al* 2009). These alleles are associated with three different levels of αS2-casein in milk (Vacca *et al* 2009). The strong alleles A, B, C, E and F are associated with a normal αS2-casein content (about 2.5 g/l for allele), D allele with an intermediate content (1.5 g/l) and the "null" allele 0 in homozygosis is associated with the apparent absence of αS2-casein in milk (Ramunno *et al* 2001, reviewed by Vacca *et al* 2009).

2.2.4. κ-Casein, CSN3

κ-casein constitutes about 15% of the total caseins (Reale *et al* 2005). The caprine *κ*-CN gene comprises five exons with the coding region for mature protein contained in exons 3 (9 amino acids) and 4 (162 amino acids) (Yahyaoui *et al* 2003). The number of goat CSN3 variants has reached 16 (Caroli *et al* 2006; Chiatti *et al* 2007) and are corresponding to 13 *κ*-CN variants, and 3 synonymous mutations (Caroli *et al* 2006). The κ-casein gene includes 5 exons, 4 of them carrying more than 90% of the information encoding for mature protein (Marletta *et al* 2007). *CSN3* is not evolutionarily related to the "calcium-sensitive" casein genes, but is physically linked to this gene family, and is functionally important in stabilizing the Ca-sensitive caseins in the micelle (Rijnkels 2002). *κ*-casein also determines the size and specific function of milk micelles, and its cleavage by chymosin is responsible for milk coagulation (Yahyaoui *et al* 2003). However, the influence of CSN3 on milk production traits still remains to be evaluated (Hayes *et al* 2006).

Compared to the Ca sensitive caseins, CSN3 exhibits distinctive properties: it is the only glycosylated and hydrophilic casein, so it is soluble in a broad range of calcium ions and presents a lower phosphorylation level (Marletta *et al* 2007). *κ*-CN differs from other caseins in its solubility over a broad range of calcium ion concentrations and contains a hydrophilic C–terminal region (Yahyaoui *et al* 2003). Overall the selective pressure on *CSN3* appears to be the strongest compared to other casein genes; this may be due to its role in casein micelle structure organization (Rijnkels 2002).

2.3. Effects of casein genes on milk production traits

Factors affecting milk production generally belong to three categories: zootechnical, environmental and genetical (Feligini *et al* 2005). The protein and fat contents are variable among the different caprine breeds and they are genetically controlled, especially by the αS1-casein locus (Moatsou *et al* 2004). This genetic polymorphism in goat milk is strongly related to the casein content (Trujillo *et al* 2000). For example, goat milk contains at least one null type genetic variant for CSN2 showed poorer coagulation properties (Albenzio *et al* 2009). Moreover, several alleles of the 3 calcium-sensitive CN (*αS1*-CN, *β*-CN, and *αS2*-CN) are associated with a null or reduced expression of the specific protein (Caroli *et al* 2007). For

example, polymorphism in *CSN1S1* is associated with different concentrations of αS1-CN in milk (Gigli *et al* 2008).

In goats, the *CSN1S1* gene has an important effect on the protein content of goat milk and a smaller effect on total protein yield (Sztankóová *et al* 2007). It could be, therefore, useful to select animals with strong alleles at CSN1S1 and CSN2 destined to produce milk for cheese making. Animals with weak or null alleles for CSN1S2 and CSN1S1 should be used in breeding programs aimed at producing milk with hypoallergenic properties (Albenzio *et al* 2009). This is because, the characteristics of the casein fraction are important for the cheese making properties of caprine milk (Moatsou *et al* 2004). Even though there are inconsistencies among different findings, the association between kappa casein and cheese-making is generally accepted (Feligini *et al* 2002).

2.4. Linkage disequlibrium in casein genes

Associations between casein alleles and protein yield could occur if linkage disequilibria exist between mutations in the coding regions and in the regulatory sequences of casein genes (Lien *et al* 1995). Since casein loci are tightly linked (Feligini *et al* 2002), casein genes might be in linkage disequilibrium (Bovenhuis *et al* 1992). This condition was substantiated for example by the study of Caroli *et al* (2007) in west African goat breeds, Caroli *et al* (2006) for Italian goats and Sztankóová *et al* (2009) in Czech goats. Although recombination among the casein genes is essential in explaining the haplotype variability, strong linkage disequilibrium resulting in an unbalanced distribution of the haplotypes among breeds (Caroli *et al* 2006). Therefore, estimates of casein genotype effects obtained using the single-gene model might be affected by effects of linked casein genes (Bovenhuis *et al* 1992). This linkage might be resulted from selection of specific casein haplotypes for their nutritional importance (Caroli *et al* 2007).

3. Materials and Methods

3.1. The study populations

The two goat populations used in our analysis were Norwegrian goats and their 50% French Alpine crosses kept at 9 farms. According to Helga Kvamsås (personal communication 2010), these farms are located in Stranda municipality, Møre og Romsdal County of the northernmost part of western Norway. The Norwegian goats were earlier divided into geographical groups, because mating mainly occurred within a region. However, due to extensive use of AI, most goats are now considered as one breed (Sæther 2002). Therefore, the Norwegian dairy goat breed is a landrace without the requirements for uniform colour (Finocchiaro *et al* 2008). There are 60000 Norwegian dairy goat population (Asheima and Eik 1998), included in the main active breeding dairy goat population (Sæther 2002). The French Alpine (Alpine) goat is the oldest among dairy goat breeds in Europe. The Alpine is noted for its long and productive lactation and for its good flavour and fat rich milk (Bowling 1929).

3.2. Breeding and feeding strategies

In Norway, cooperative genetic improvement of goats through a buck-circle system was introduced in the 1960s and 1970s (Andonov *et al* 2007). In this system, about 15% of the goats are bred to elite bucks to get replacement goats (Bagnicka *et al* 2007). This was supported by progeny testing and selection with the main objectives of improving the dry matter content of the milk and ease of milking (Andonov *et al* 2007). Normally kidding in Norwegian goats takes place from January to March (Asheima and Eik 1998) and the genetic evaluation system classifies the kidding season into December to February, March to May, and June to November (Andonov *et al* 2007). However, Bagnicka *et al* (2007) grouped the kidding season into two classes: October through March and April to September. According to Tormod Ådnøy (personal communication, 2010), the feeding strategy of goats in Norway can be broadly classified into two; the indoor feeding (October to April) and the outdoor feeding (May to September).

3.3. Genotyping

Blood samples were collected from Norwegian goats and their French Alpine crosses by TINE[1] for DNA isolation. SNP candidates in the casein region were found from the literatures and by sequencing the DNA of previously genotyped goats that showed polymorphism (Ådnøy *et al* 2006). Primers for resequencing of casein loci were designed for the promoters, selected exons, and introns of CSN1S1, CSN2, and CSN3, including exon 16 of CSN1S2 and exon 7 of CSN2 (Hayes *et al* 2006). SNPs were genotyped using Matrix Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry, MALDI-TOF-MS (Sequenom, San Diego, CA, USA) at CIGENE (Center for Integrative Genetics, Norwegian University of Life Sciences) laboratory. This sequencing has detected 39 SNPs showing polymorphism for the cluster of the four casein genes (Ådnøy *et al* 2006) and 38 of them were used in our analysis. These 38 SNPs; 14 at CSN1S1, 6 at CSN2, 4 at CSN1S2 and 14 at CSN3 were detected in the promoter, exons and introns regions of the casein loci.

3.4. Phenotype recording

Phenotypic recording on test day milk yield (kg) was done at farm level by the participating farmers and agents of TINE. Moreover, the lab analysis results for fat, lactose and protein percentage, somatic cell count, free fatty acid content, urea content and milk taste score were obtained from TINE. After keeping only those does with both phenotypic records and genotype information on test day milk yield, 376 does with 1670 records were left for analysis.

3.5. The data structure and traits studied

This study was based on the dataset of test day milk yield and milk component traits of 216 lactating does of Norwegian goats and 160 crosses born between September 2007 and September 2008 and that are kept in 9 farms. Descriptive statistics summary for the flock structure is presented in Table 10. Eight milk production traits were considered in our analysis, namely, test day milk yield (kg), fat precentage, protein percentage, lactose

[1] The main company involving in the production, distribution and export of dairy products in Norway.

percentage, somatic cell count (SCC, cell/ml), free fatty acids content (FFA), urea content (urea) and milk taste score (taste score).

3.6. Data analysis

Descriptive statistics of SAS (2002) was used to edit the dataset and to do the preliminary analysis. Outliers were checked and removed using univariate analysis of SAS (2002). Histograms indicating the distribution of observations after removal of outliers and the corresponding Q_Q plot as test of normality are presented in the Appendix Figures 38 through 53. Somatic cell count and FFA were transformed into their logarithmic value to normalize their highly skewed distribution. For fixed effect variables other than SNP and haplotype effect, model effects were tested at $p<0.05$ significant level. However, some biologically important parameters were kept in the model even if they are not significant at $p<0.05$ (see Appendix Table 11 through 18). Type III sum of squrare of SAS (2002) was used to determine the significance level.

3.6.1. Preliminary data analysis

Preliminary data analysis were done using the following fixed effect model other than the fixed effect of SNP and haplotype.

$$y_{ijklmnopqr} = \mu + fm_j + br_k + sn_l + fs_m + ls_n + br * fm_p + kd_q + td_r + e_{ijklmnopqr} \quad (1)$$

where $y_{ijklmnopqr}$ is test day milk yield or somatic cell count or fat% or protein% or lactose% or urea content or free fatty acids content or milk test score for animal i in farm j (9 levels); for breed k (2 levels, Norwegian goat and the crosses); season l (3 levels, December to February, March to May and June to November); feeding strategy m (2 levels, indoor October through April and outdoor May to September); stage of lactation n (defined in 30 days interval into 6 levels, 1 through 5 and ≥6); interaction of breed with farm p; the covariate kidding date q and test day r (16 levels) for animal i, $e_{ijklmnopqr}$ is the random residual term and μ is the overall common mean. By using equation (1) systematic environmental effects were corrected for subsequent least square estimation of SNP or haplotype effect using univariate analysis of

linear mixed model in ASReml (Gilmour *et al* 2009). This model is, however, except test date (this explanatory variable was used for milk taste score analysis only) was fitted for test day milk yield analysis only; whereas different models were fitted for other response variables (see Appendix Tables 11 to 18).

3.6.2. Haplotype construction

Unlike the performance dataset analysis; for haplotype construction a total of 605 genotyped goats were used (including male goats and does having no phenotypic records). These include 144 male crosses; 185 female crosses, 3 Norwegian male goats and 273 Norwegian female goats. PHASE program (Stephens *et al* 2001) was used to construct haplotypes using 38 SNPs obtained from genotypes of 605 goats. SNP haplotypes were predicted for each casein gene and for the complete segment of the casein loci. Best pairs summary of PHASE software (Stephens *et al* 2001) was used to pick up plausable haplotypes (to identify putative haplotype blocks) and these were fitted in our linear mixed model (equation 2) for haplotype effect analysis.

3.6.3. Linkage disequilibrium analysis

The level of linkage disequilibrium between all pairs of loci was estimated using the r^2-statistic (Hudson 1985) as implemeted in HaploView program, and the result was visualized using the HaploView program (Barret *et al* 2005). Positions of SNP markers for CSN1S1 directly taken from the genotype data, whereas for othe casein gene positions of SNP markers were calculated by using bovine genome sequence information of NCBI (http://www.ncbi.nlm.nih.gov/). Observed and expected heterozygosity, minimum allelic frequency, deviations from Hardy-Weinberg and type of alleles at each marker presented in Table 19 and 20 were computed by HaploView software (Barret *et al* 2005). Moreover, to investigate the effect of crossbreeding on the extent of LD, separate analysis were done for the two goat populations using HaploView (Barret *et al* 2005).

3.6.4. SNP tagging

The default of the tagger option (pair tagging only) of the HaploView software (Barret *et al* 2005) was used to tag SNPs. Accordingly, all the 38 SNps were tested for their capacity of inferring about other SNPs.

3.6.5. Mixed model with repeated records analysis

3.6.5.1. The pedigree structure

The additive genetic relationship matrix A includes 2733 animals including the tested ones. The pedigree was seven generations deep for dam line and for does having Norwegian goat buck, whereas the pedigree is two generations deep for crosses descended from Alpine goat sire line.

3.6.5.2. The linear mixed model

The mixed model fitted accounting for repeated records including the fixed effect of SNP and haplotype is:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Xq} + \mathbf{Zu} + \mathbf{Zpe} + \mathbf{e} \qquad (2)$$

where ***y*** is a vector of phenotypic observations on the trait, **β** is a vector of fixed effects comprising a general mean and other fixed effects shown in equation (1); **q** is a vector of fixed SNP or haplotype additive and dominance effects; **u** is a vector of additive polygenic effects other than SNP and haplotype effects, **pe** is a vector of permanent environmental effects and **e** is a vector of residual effects. The matrix **X** is the incidence matrix relating observations to fixed effects and **Z** is the incidence matrix relating observations to random animal effects.

Model definitions:

$$E(y) = X_{\beta} + X_{q} \qquad E(u) = 0 \qquad E(p_e) = 0 \qquad E(e) = 0$$

The random terms are assumed to follow a normal distribution:

$$u \sim (0, A\sigma_u^2) \qquad p_e \sim (0, I\sigma_{p_e}^2) \qquad \mathrm{e} \sim (0, I\sigma_e^2)$$

$$\mathrm{var}(u) = G_1 \qquad \mathrm{var}(p_e) = G_2 \qquad \mathrm{var}(e) = R \qquad \mathrm{var}(y) = ZG_1Z' + ZG_2Z' + R$$

The three random variables have the following distribution:

$$\mathrm{var}\begin{pmatrix} u \\ pe \\ e \end{pmatrix} = \begin{pmatrix} A\sigma_u^2 & 0 & 0 \\ 0 & I\sigma_{pe}^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{pmatrix}$$

where $\sigma^2 u$ is the direct additive genetic variance, $\sigma^2{}_{pe}$ is the variance due to permanent environmental effects and σ_e^2 is the variance of the residual term.

Based on equation 2 the following mixed repeatability model least square equation was fitted in ASReml (Gilmour *et al* 2009) to get the least square estimates of fixed and random effects for traits studied.

$$\begin{pmatrix} \hat{\beta} \\ \hat{q} \\ \hat{u} \\ \hat{p}_e \end{pmatrix} = \begin{pmatrix} X_\beta' R^{-1} X_\beta & X_\beta' R^{-1} X_q & X_\beta' R^{-1} Z & X_\beta' R^{-1} Z \\ X_q' R^{-1} X_\beta & X_q' R^{-1} X_q & X_q' R^{-1} Z & X_q' R^{-1} Z \\ Z' R^{-1} X_\beta & Z' R^{-1} X_q & Z' R^{-1} Z + \alpha A^{-1} & Z' R^{-1} Z \\ Z' R^{-1} X_\beta & Z' R^{-1} X_q & Z' R^{-1} Z & Z' R^{-1} Z + \lambda I \end{pmatrix}^{-1} \begin{pmatrix} X_\beta' R^{-1} Y \\ X_q' R^{-1} Y \\ Z' R^{-1} Y \\ Z' R^{-1} Y \end{pmatrix}$$

Where $\alpha = \sigma_e^2 / \sigma_u^2$ and $\lambda = \sigma_e^2 / \sigma_{p_e}^2$

Variance component estimation using small datasets most likely leads into erranous results. Therefore, to overcome this problem we used the variance components estimated from large dataset for the six traits by the Norwegian association of sheep and goat farmers (Norsk Sau og Geit, NSG). However, because of lack of variance component estimates for milk taste score and urea content of milk from NSG, we used the variance components estimated from our small dataset.

3.6.5.2.1. SNP haplotypes effect analysis

For both 37 bi-allelic markers and 1 multi-allelic marker, the linear model in equation 2 used in our analysis for all fixed effects assumes the following distribution:

$$y \sim (X\beta + Yq, \sigma_e^2)$$

where β is a vector containing an overall mean (μ) and all fixed variables affecting the trait of interest (equation 1). The vector q represents the additive and dominant fixed effect of a given SNP or haplotype. The incidence matrice X connects the phenotypes to other fixed effects and the incidence matrice Y connects the phenotypes to SNP or haplotype additive and dominance effects. For bi-allelic SNPs (for example; AA, AG, GG) these values equal (+1, 0,–1) for additive, and (0, 1, 0) for dominance effects; and allele A was taken as the common wild allele, despite high frequency of allele D. For multi-allelic marker SNP14 in exon12; 6 genotypes were observed (AA, GG, DD, AD, AG and GD, where D is point deletion), with additive effect (1, –1, –1, 0, 0, -1) for allele A, and dominance effect for SNP genotypes were (0, 0, 0, 1, 1, 1), respectively.

For most common SNP haplotypes and individual SNP additive and dominance effect analysis the following model was fitted:

$$y = \mu + m + n_i + n_j + n_{ij} + u + p_e + e \qquad (3)$$

Where y is trait of interest; m the effect of the other fixed factors (other than SNP and haplotype effect, equation 1); n_i and n_j are maternal haplotype and paternal haplotype (either for the whole casein loci or for individual genes); n_{ij} is the SNP or haplotypes dominance interaction, u is the animal's polygenic effect; p_e is permanent environment; e is the error term and μ is the overall mean.

3.6.5.2.2. αS1-casein exon 12 deletion effect analysis

Due to multi-allelic nature of this locus its effect on milk yield and its component traits was treated separately in our analysis. Moreover, genotype effect of this locus was estimated besides its additive and dominance effects. According to Hayes *et al* (2006) sequences of the three polymorphisms detected in exon 12 harboring the deletion on which our analysis was based on are:

Allele 1(D): GAACAGCTTCTCAGACTGAAAAATACAACGTGCCCCAGCTG
Allele 3 (G): GAACAGCTTCTCAGACTGAAGAAATACAACGTGCCCCAGCTG
Allele 6 (A): GAACAGCTTCTCAGACTGAAAAAATACAACGTGCCCCAGCTG

4. Results

4.1. Linkage disequilibrium and haplotype diversity

4.1.1. Linkage disequilibrium among casein SNPs

Seven SNPs were departed from Hardy-Weinberg equilibrium in the crosses; one at αS1-casein and six at κ-casein, whereas only two SNPs at αS1-casein were departed from Hardy-Weinberg equilibrium in the Norwegian goats (data not shown). The SNP markers were not evenly distributed along the casein loci block. Some are found very close to each other whereas there are also some adjacent SNPs distantly located. The LD was less extensive in the crosses, especially for αS1-casein and κ-casein compared to the Norwegian goats (Figures 3 & 4). These figures displayed the level of LD between markers measured by r^2 showed that intra locus LD is stronger than inter loci LD. Moreover, as the distance between the loci is increased the LD get less extensive. However, the haplotypes display showed that except for β-casein and αS2-casein, the inter-loci LD between adjacent loci was stronger in the Norwegian goats than it was in the crosses (see Appendix Figures 54 & 55). Therefore, this showed that the amount of historical recombination between the two adjacent casein genes was weak, even though, the trend between CSN2 and CSN1S2 looks different especially in the Norwegian goats (Figure 54).

.

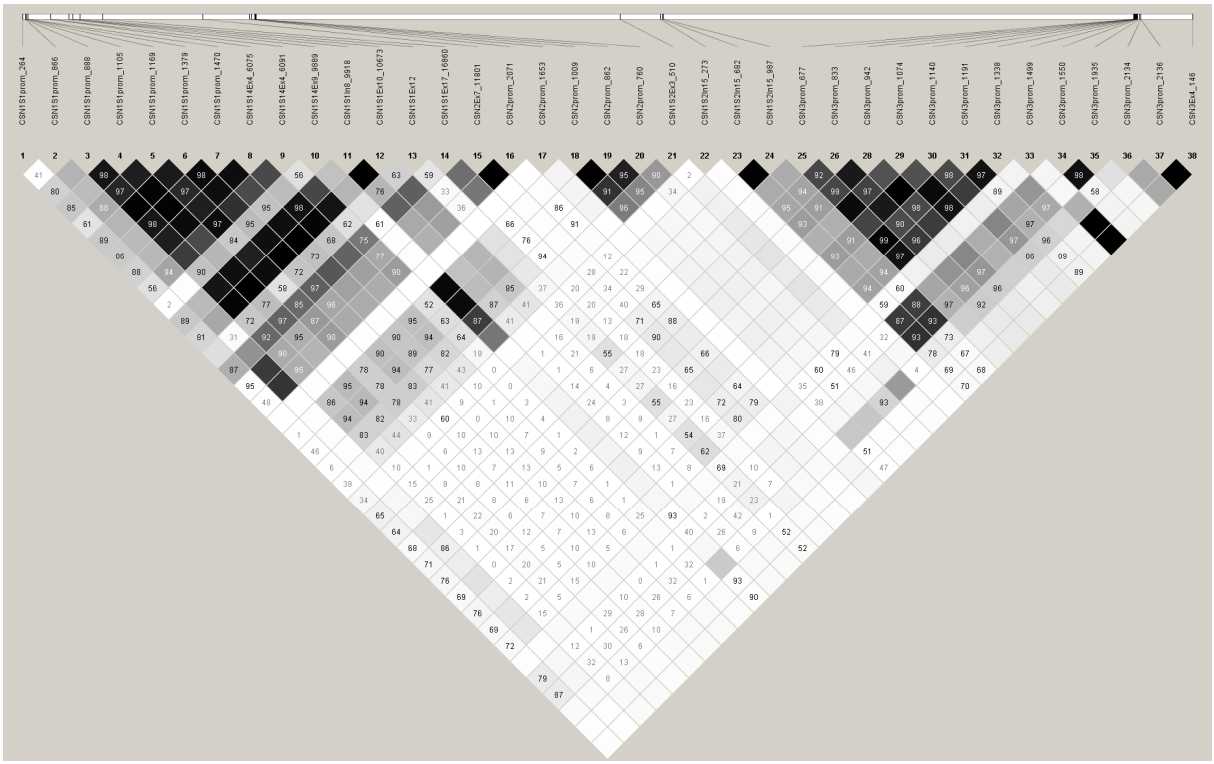

Figure 3. The extent of LD observed across the four casein genes' SNPs in Norwegian goats.

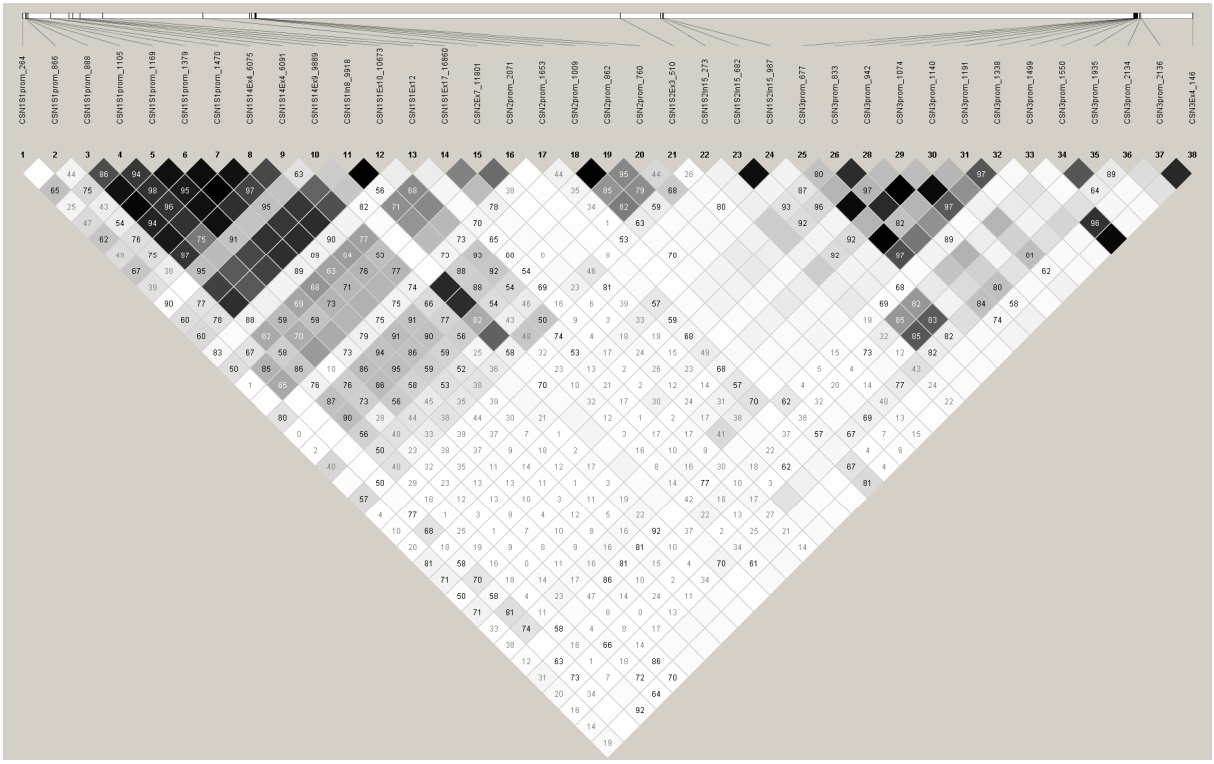

Figure 4. The extent of LD observed across the four casein genes' SNPs in the crosses.

4.1.2. Tagger and tagged SNPs

There were less tagger SNPs in the Norwegian goats than the crosses (Table 1 *vs* 3), which showed the presence of strong LD in the Norwegian goats. For Norwegian goats, there were 16 tagger SNPs (Table 1). These 16 tagger SNPs were tested in 16 tests that can predict all the 38 SNPs at $r^2 \geq 0.8$ with mean maximum r^2 of 0.971. Out of these 16 tagger SNPs, 7 can be used to inferred about one or more other SNP(s) (Table 2).

Table 1. List of tagger SNPs obtained in the Norwegian goats.

Tagger SNP	Tagger SNP	Tagger SNP	Tagger SNP
CSN1S1Ex10_1067	CSN3prom_1499	CSN1S1prom_264	CSN1S1Ex12
CSN2prom_1009	CSN3prom_942	CSN1S1Ex17_16860	CSN3prom_2134
CSN3prom_1191	CSN2Ex7_11801	CSN2prom_1653	CSN1S1prom_866
CSN3prom_1935	CSN1S2In15_273	CSN1S2Ex3_510	CSN3prom_852

Table 2. List of tagger SNPs used to predict other SNPs in the Norwegian goats.

Tagger SNP	Tagged SNP(s)
CSN1S1Ex10_10673	CSN1S1prom_1169,CSN1S1prom_1470,CSN1S1prom_1379,CSN1S14Ex4_6091,CSN1S1prom_1105,CSN1S1In8_9918,CSN1S14Ex4_6075,CSN1S1prom_888
CSN2prom_1009	CSN1S14Ex9_9889,CSN2prom_760,CSN2prom_862
CSN3prom_1191	CSN3prom_1338,CSN3prom_1074,CSN3prom_677
CSN3prom_1935	CSN1S2In15_987,CSN3prom_1550,CSN1S2In15_682
CSN3prom_1499	CSN3prom_2136,CSN3Ex4_146
CSN3prom_942	CSN3prom_1140,CSN3prom_833
CSN2Ex7_11801	CSN2prom_2071

Using HaploView pairwise tagging option; in the crosses 21 tagger SNPs in 21 tests, which can predict all the 38 SNPs at $r^2 \geq 0.8$ with mean maximum r^2 of 0.955 were found (Table 3). Out of these 21 tagger SNPs, 6 can be used to predict about one or more other SNP(s) in the segment (Table 4).

Table 3. List of tagger SNPs obtained in the crosses.

Tagger SNP	Tagger SNP	Tagger SNP	Tagger SNP
CSN1S1prom_1379	CSN1S2In15_273	CSN1S1Ex12	CSN3prom_2134
CSN3prom_677	CSN3prom_852	CSN1S2Ex3_510	CSN3prom_1935
CSN3Ex4_146	CSN1S1prom_866	CSN3prom_1338	CSN2prom_760
CSN2prom_862	CSN3prom_1550	CSN1S1Ex17_16860	CSN2prom_1653
CSN3prom_833	CSN1S1prom_264	CSN2Ex7_11801	CSN2prom_2071
CSN1S2In15_987			

Table 4. List of tagger SNPs used to predict other SNPs in the crosses.

Tagger SNP	Tagged SNP(s)
CSN1S1prom_1379	CSN1S1prom_1105,CSN1S14Ex4_6091,CSN1S1Ex10_10673,CSN1S14Ex4_6075,CSN1S1In8_9918,CSN1S1prom_1470,CSN1S1prom_1169 and CSN1S1prom_888
CSN3prom_677	CSN3prom_1191 and CSN3prom_1074

There were no common tagger SNPs for the two populations, and therefore a SNP that was used as tagger in one population is tagged in the other (Table 2 & 4).

4.1.3. Haplotype diversity

4.1.3.1. Haplotypes diversity output from HaploView

To see the variation in haplotype structure between the two populations separate analysis were made for haplotype construction using HaploView (Barret *et al* 2005). Our results in Table 5 & 6 showed remarkable variation in diversity and frequency of the four individual casein genes haplotypes between the two populations. Accordingly, 8, 3, 4 and 8 SNP haplotypes with frequencies of ≥0.01 were detected for αS1-, β-, αS2- and κ-casein, respectively for the Norwegian goats; whereas these numbers were, respectively 9, 6, 5 and 11 in the crosses. This showed that there is considerable variation in the number of SNP haplotypes detected especially for β- and κ-casein. Variation in the individual SNP haplotypes frequency and haplotype frequency re-ranking were observed in the two populations.

Therefore, our results in Table 5 & 6 showed that, there were 17 haplotypes for αS1-casein in the two populations and 4 of them were shared by both populations. For β-casein's 9 haplotypes were detected in the two populations and 3 of them were shared by both populations. There were 9 haplotypes for the two populations' αS2-casein and 4 of them were detected in both populations. There were 19 haplotypes for the two populations κ-casein and 7 of them were shared by both populations. Therefore, the two population common haplotypes were larger for αS2-casein and followed by κ-, β- and αS1-casein.

Moreover, when the entire casein loci is considered as a block, there were only 12 haplotypes with a frequency of ≥0.01 in the Norwegian goats whereas there were 21 in the crosses (data not shown). Therefore, there were 33 SNP haplotypes for the two populations entire casein loci and none of them were shared between the two populations.

Table 5. SNP haplotypes of individual casein genes frequencies in the Norwegian goats.

Block	Haplotype	Freq.
αS1- casein		
	ACGGGCGTCCACCC	0.635
	ACGGGCGTCCACAC	0.162
	GCAAATACGCGGAT	0.04
	ACAAATACGGGGAC	0.039
	ATAAATACGCGGAT	0.027
	ATAAATACGCGGCT	0.025
	ACAAGTACCGGGAC	0.014
	GCAAGTACCCGGAT	0.01
β- casein		
	TAGATC	0.858
	CGGATC	0.079
	TAGGAT	0.05
αS2-casein		
	GCCA	0.567
	GCTT	0.321
	ACCA	0.098
	GGCA	0.012
K-casein		
	AATACGGATGAGC	0.385
	GGATGTTACTAGC	0.302
	GGATGTTATGGGC	0.102
	GGATGTTGTGGAT	0.08
	GATTCTTATGAGC	0.05
	GGATGTTACGGGC	0.024
	AGTTCGGATGAGC	0.012
	GATTCTTATGGGC	0.011

For αS1-casein SNP 13 position (exon 12 deletion); C represents the deletion and A represents A or G. For SNP 10 position (exon 9 deletion) G represents the deletion and A represents C.

Table 6. Haplotype blocks of individual casein genes and their frequencies in the crosses.

Block	Haplotype	Freq.
αS1- casein		
	ACGGGCGTCCACAC	0.471
	ACGGGCGTCCACCC	0.243
	ACGGGCGTCCGGAT	0.06
	ACAAATACGGGGAC	0.037
	GCGGGCGTCCACAC	0.03
	GCAGATACGCGGAT	0.022
	ACAAATACGCGGAT	0.022
	ATAAATACGCACAT	0.017
	GCAAATACGCGGAT	0.017
β- casein		
	TAGATC	0.803
	CGGATC	0.056
	TAGATT	0.044
	TAGGAT	0.042
	CAGATC	0.027
	CAAATC	0.01
αS2-casein		
	GCCA	0.588
	GGCA	0.165
	GCTT	0.161
	ACCA	0.042
	AGCA	0.031
K-casein		
	GGATGTTATGGGC	0.22
	GGATGTTACTAGC	0.179
	AATACGGATGAGC	0.162
	GATTCTTATGAGC	0.156
	GGATGTTGTGGAT	0.062
	GATTCTGATGAGC	0.06
	GGATGTTACGAGC	0.06
	GGATGTTACGGGC	0.018
	GATTCTTATGGGC	0.014
	AGTACGGATGAGC	0.014
	GGTTCTTATGAGC	0.01

For αS1-casein SNP 13 position (exon 12 deletion); C represents the deletion and A represents A or G. For SNP 10 position (exon 9 deletion) G represents the deletion and A represents C.

4.1.3.2. Haplotypes diversity from PHASE output

Phase program (Stephens *et al* 2001) output showed that for the entire casein loci, 18 haplotypes that have a frequency of ≥1% were obtained in the whole population (Table 7). For individual casein genes haplotypes having frequencies of ≥1% were 12, 5, 4 and 10 for αS1- , β- , αS2- and κ-casein, respectively (Table 8). These haplotypes were used to estimate the additive and dominance effects of SNP or haplotypes on traits studied using linear mixed model that accounts for repeated records.

Table 7. The entire casein loci block SNP haplotypes.

№	Haplotypes	Freq.
1	ACGGGCGTCCAC 1 CTAGATCGCCAAAGTACGGATGAGC	0.19
2	ACGGGCGTCCAC 1 CTAGATCGCCAGGGATGTTGTGGAT	0.04
3	ACGGGCGTCCAC 1 CTAGATCGCTTGGGATGTTACTAGC	0.11
4	ACGGGCGTCCAC 1 CTAGATCACCAGGGATGTTATGGGC	0.03
5	ACGGGCGTCCAC 6 CTAGATCGCCAAAGTACGGATGAGC	0.01
6	ACGGGCGTCCAC 6 CTAGATCGCCAGAGTTCTGATGAGC	0.01
7	ACGGGCGTCCAC 6 CTAGATCGCCAGAGTTCTTATGAGC	0.10
8	ACGGGCGTCCAC 6 CTAGATCGCCAGGGATGTTATGGGC	0.04
9	ACGGGCGTCCAC 6 CTAGATCGCCAGGGATGTTGTGGAT	0.01
10	ACGGGCGTCCAC 6 CTAGATCGCTTGGGATGTTACTAGC	0.06
11	ACGGGCGTCCAC 6 CTAGATCGGCAGGGATGTTATGGGC	0.01
12	ACGGGCGTCCAC 6 TTAGATCGCCAGAGTTCTGATGAGC	0.01
13	ACGGGCGTCCGG 3 TCAGATCGGCAGGGATGTTACGAGC	0.01
14	ACAAATACGDGG 6 CTAGGATACCAGGGATGTTATGGGC	0.02
15	ATAAATACGCGG 3 TCGGATCGCCAAAGTACGGATGAGC	0.02
16	GCGGGCGTCCAC 6 CTAGATTGGCAGGGATGTTATGGGC	0.01
17	GCAGATACGCGG 3 TTAGATTGGCAGGGATGTTACTAGC	0.01
18	GCAAATACGCGG 3 TTAGATCGCTTGGGATGTTACTAGC	0.02

Table 8. Individual casein genes SNP haplotypes.

Block	№	Haplotypes	Freq.
αS1- casein			
	1	ACGGGCGTCCAC 1 C	0.41
	2	ACGGGCGTCCAC 3 C	0.03
	3	ACGGGCGTCCAC 6 C	0.28
	4	ACGGGCGTCCAC 6 T	0.01
	5	ACGGGCGTCCGG 3 T	0.02
	6	ACAAATACGCGG 3 T	0.01
	7	ACAAATACGDGG 6 C	0.04
	8	ATAAATACGCGG 3 T	0.02
	9	GCGGGCGTCCAC 6 C	0.03
	10	GCAGATACGCGG 3 T	0.01
	11	GCAAGTACCCGG 3 T	0.01
	12	GCAAATACGCGG 3 T	0.04
β- casein			
	1	TAGATC	0.84
	2	TAGATT	0.02
	3	TAGGAT	0.05
	4	CAGATC	0.01
	5	CGGATC	0.07
αS2-casein			
	1	GCCA	0.65
	2	GCTT	0.21
	3	GGCA	0.06
	4	ACCA	0.07
K-casein			
	1	AAGTACGGATGAGC	0.27
	2	AGGTACGGATGAGC	0.01
	3	GAGTTCTGATGAGC	0.02
	4	GAGTTCTTATGAGC	0.12
	5	GAGTTCTTATGGGC	0.01
	6	GGGATGTTATGGGC	0.16
	7	GGGATGTTACGAGC	0.02
	8	GGGATGTTACGGGC	0.02
	9	GGGATGTTACTAGC	0.25
	10	GGGATGTTGTGGAT	0.07

4.2. Haplotypes additive and dominance effects on traits studied

Haplotypes additive and dominance effects were estimated using equation 3 for all traits considered.

4.2.1. Haplotypes additive effect

4.2.1.1. The entire casein loci haplotypes additive effect

Least squre estimates results for the additive effect of the entire casein haplotypes on studied traits are presented in Figures 5 through 8. None of the haplotypes showed statistically significant effect on test day milk yield and lactose% at $p \leq 0.1$. However, haplotype 3 for fat%; haplotypes 6, 7, 10, 16 & 18 for FFA content; haplotypes 1 & 15 for protein percentage; haplotype 9 for SCC; haplotypes 1 & 10 for milk taste score, and haplotypes 4 & 18 for urea content showed signficant effect at $p \leq 0.05$.

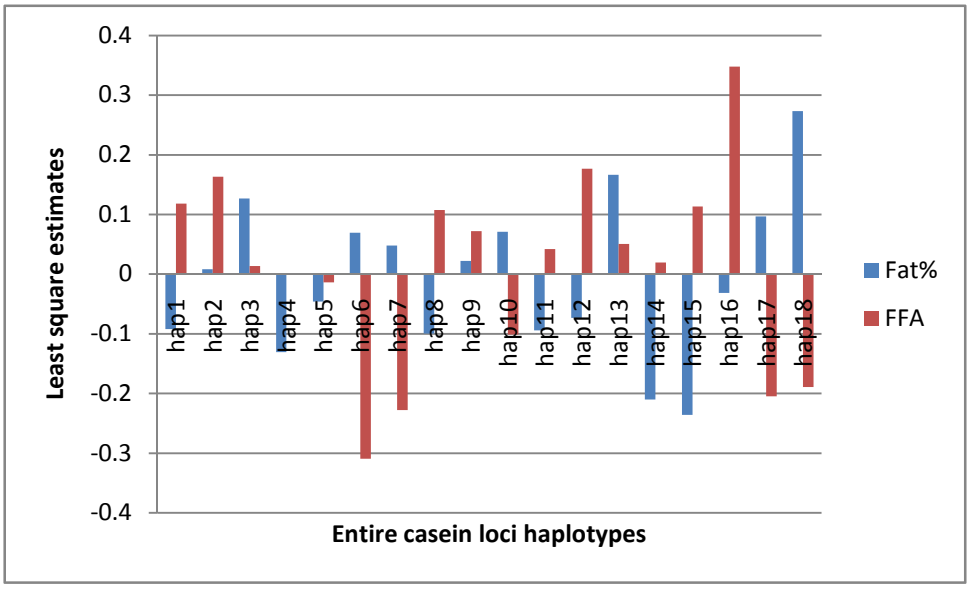

Figure 5. Additive fixed effects of the entire casein loci haplotypes on fat% and FFA content

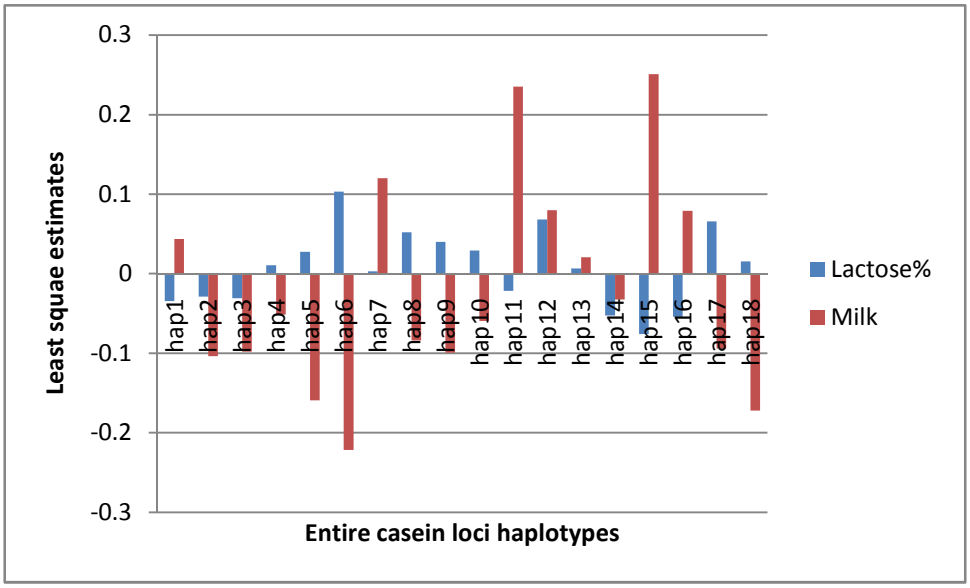

Figure 6. Additive fixed effects of the entire casein loci haplotypes on lactose% and test day milk yield (kg).

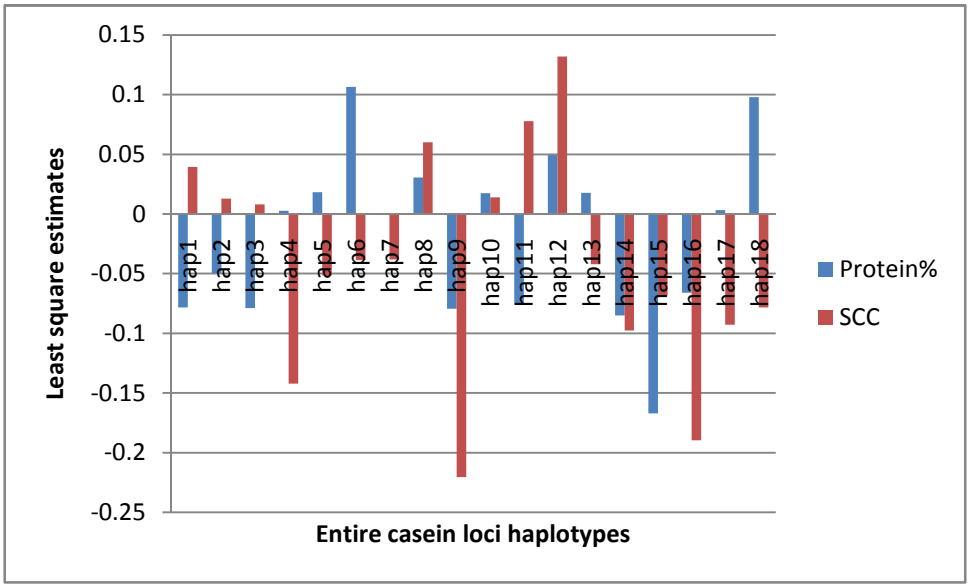

Figure 7. Additive fixed effects of the entire casein loci haplotypes on protein% and somatic cell count content (cell/ml).

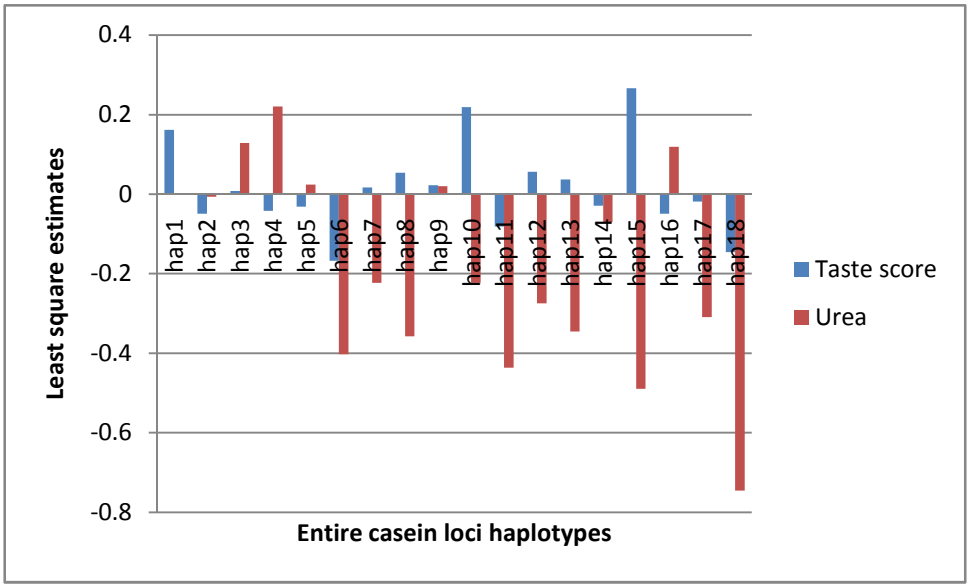

Figure 8. Additive fixed effects of the entire casein loci haplotypes on taste score and urea content.

4.2.1.2. Individual casein locus haplotypes additive effect

4.2.1.2.1. αS1-casein gene haplotypes additive effect

Our results for the effect of αS1-casein gene haplotypes presented in Figures 9 to 12. Our findings showed that for αS1-casein gene haplotype 7 has significant effect on fat%. Similarly, haplotypes 1, 10 and 12 on FFA ($p \leq 0.05$) and haplotypes 2 and 3 on test day milk yield have significant effect ($p \leq 0.1$). Moreover, haplotype 1 on protein%, haplotypes 3 and 8 on taste score, and haplotype 1 on urea content showed significant effect ($p \leq 0.05$). However, none of the αS1-casein SNP haplotypes showed significant effect on lactose percentage.

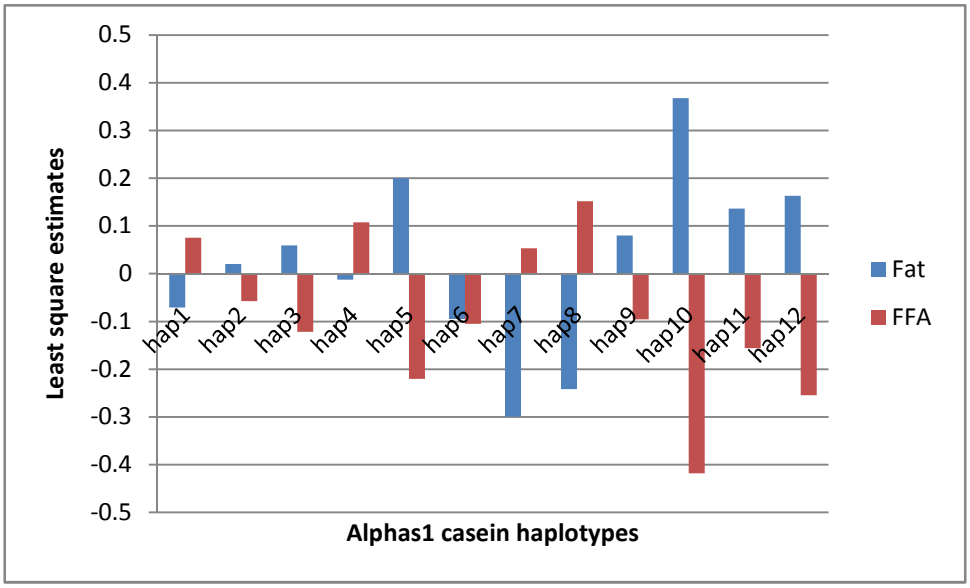

Figure 9. Additive fixed effects of the alphas1 casein loci haplotypes on fat% and FFA content.

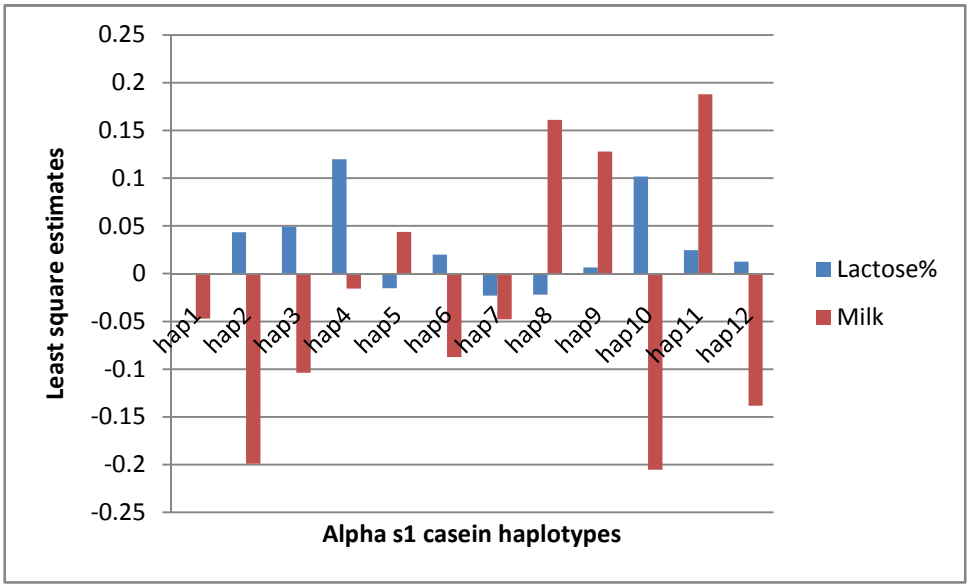

Figure 10. Additive fixed effects of the alphas1 casein loci haplotypes on lactose% and test day milk yield (kg).

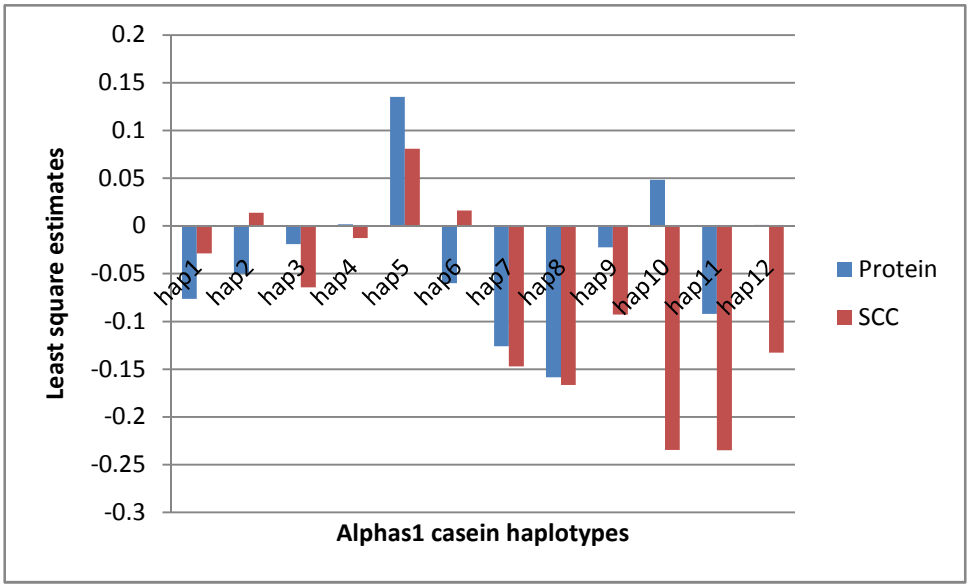

Figure 11. Additive fixed effects of the alphas1 casein loci haplotypes on protein% and somatic cell count.

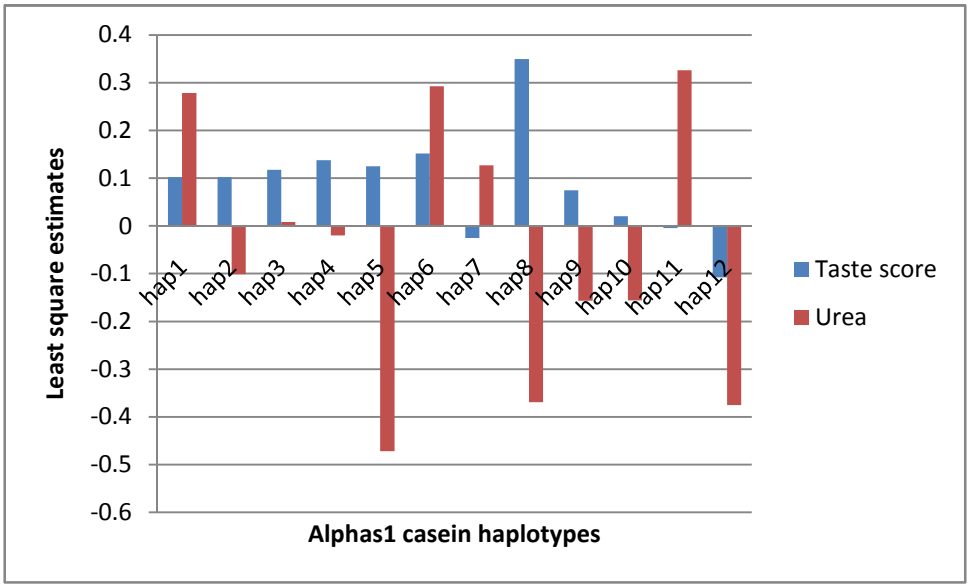

Figure 12. Additive fixed effects of the alphas1 casein loci haplotypes on taste score and urea content.

4.2.1.2.2. β-casein gene haplotypes additive effect

Our results for the effect β-casein gene haplotypes presented in Figures 13 through 15. Accordingly, haplotype 2 on fat%, haplotype 4 on FFA, and haplotype 3 on test day milk yield showed significant effect at $p \leq 0.05$.

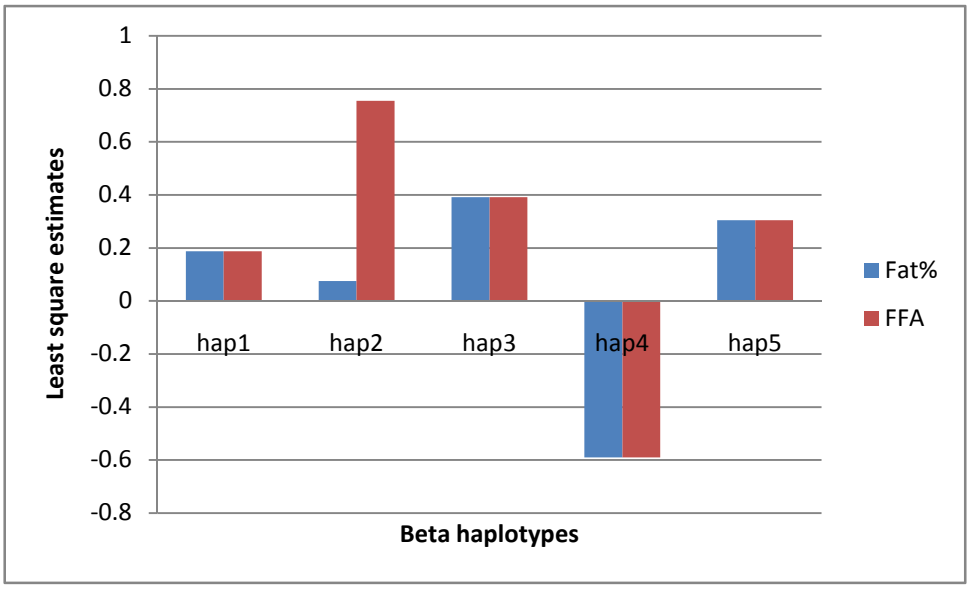

Figure 13. Additive fixed effects of the beta casein loci haplotypes on fat% and FFA content.

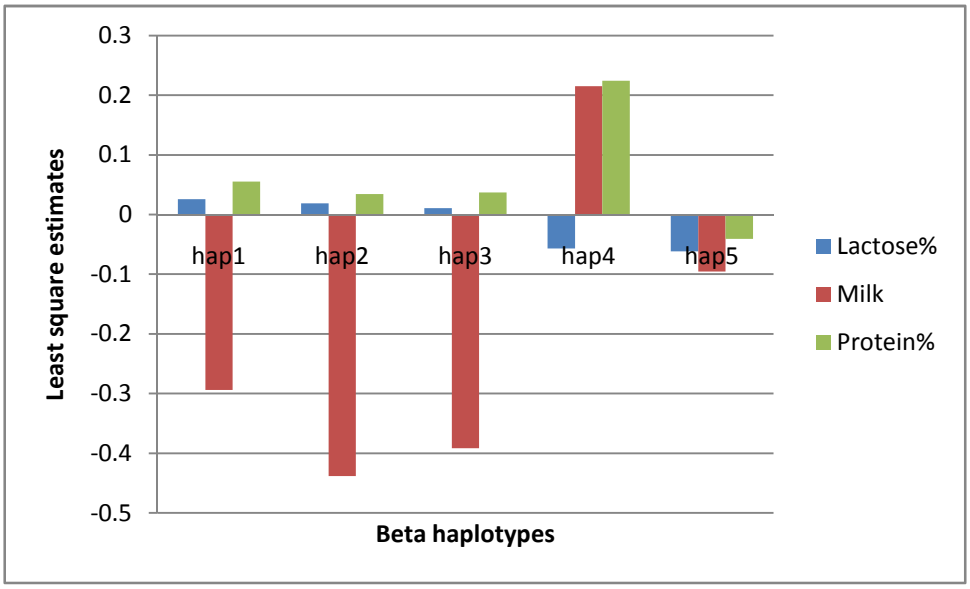

Figure 14. Additive fixed effects of the beta casein loci haplotypes on lactose%, test day milk yield (kg) and protein%.

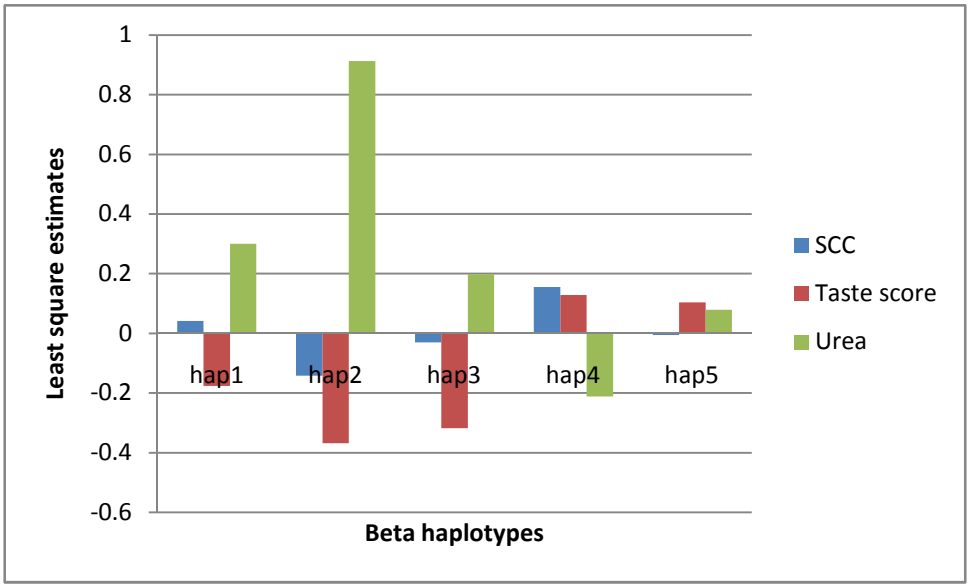

Figure 15. Additive fixed effects of the beta casein loci haplotypes on somatic cell count, taste score and urea content.

4.2.1.2.3. αS2-casein gene haplotypes additive effect

Results from αS2-casein gene haplotypes additive effect on traits studied are presented in Figures 16 and 17. Haplotype 1 on fat percentage and haplotypes 1 and 3 on FFA showed significant effect at $p \leq 0.05$. None of the αS2-casein gene haplotypes showed significant additive effect for the rest of studied traits.

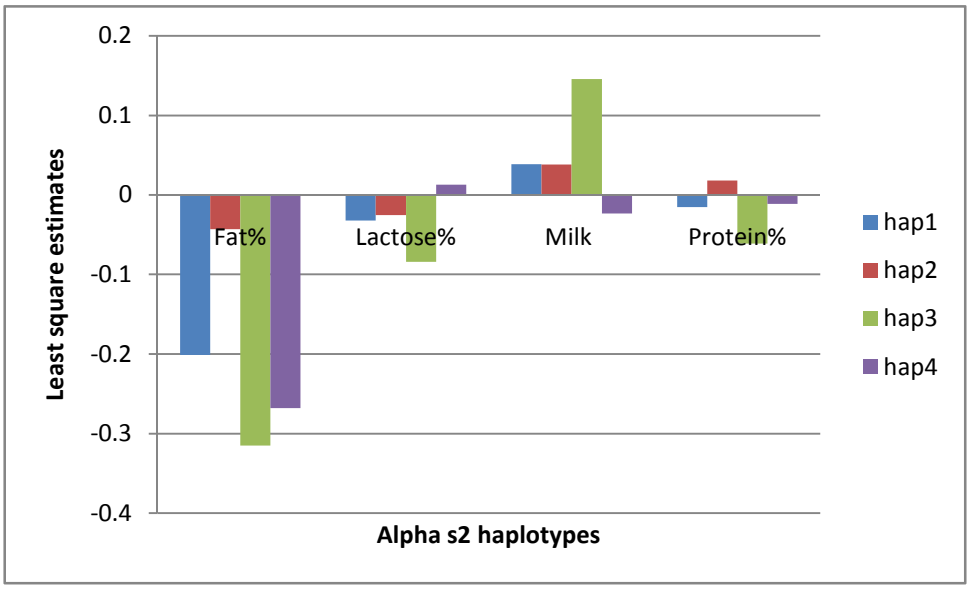

Figure 16. Additive fixed effects of the alphas2 casein loci haplotypes on fat%, lactose%, test day milk yield (kg) and protein%.

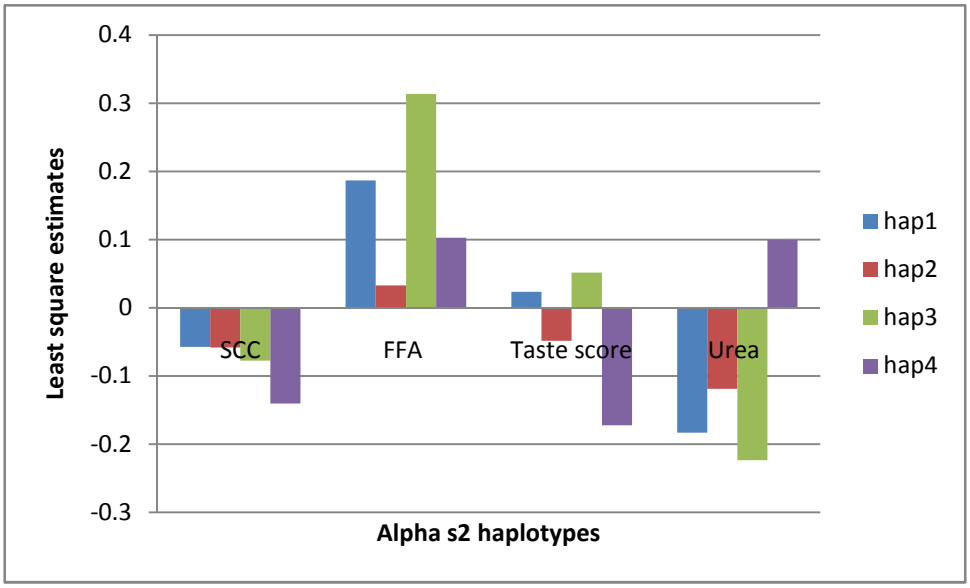

Figure 17. Additive fixed effects of the alphas2 casein loci haplotypes on somatic cell count, FFA content, taste score and urea content.

4.2.1.2.4. κ-casein gene haplotypes additive effect

Outputs from κ-casein gene haplotypes additive effect on milk production traits presented in Figures 18 to 21. Haplotypes 1, 6 and 9 on fat percentage, and haplotypes 1 and 9 on FFA content showed significant effect at $p \leq 0.05$. Similarily, haplotype 8 on test day milk yield, haplotype 6 on SCC, haplotype 1 on taste score and haplotypes 8 and 9 on urea content showed significant effect at $p \leq 0.05$. However, none of the κ-casein gene haplotypes showed signficant effect on protein and lactose percentage.

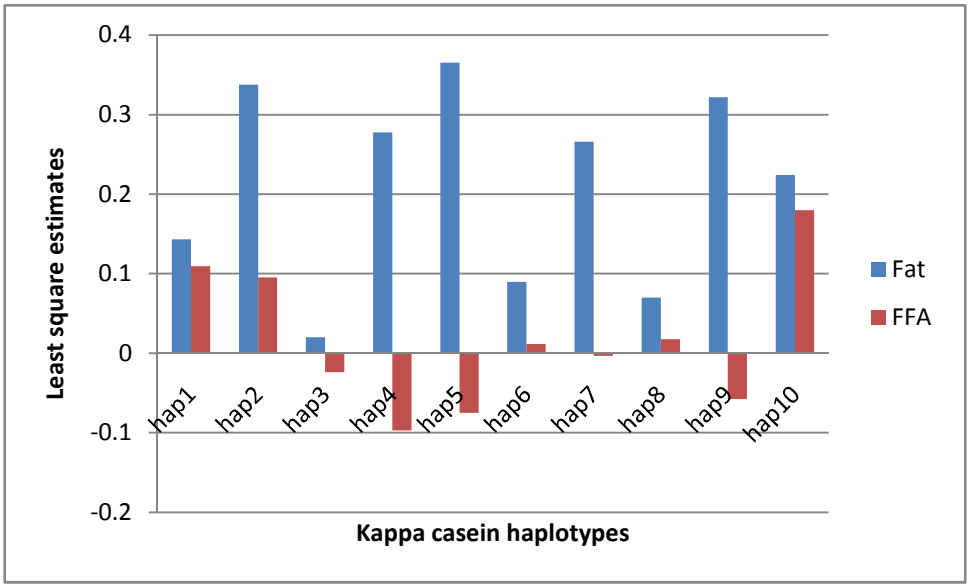

Figure 18. Additive fixed effects of the kappa casein locus haplotypes on fat% and FFA content.

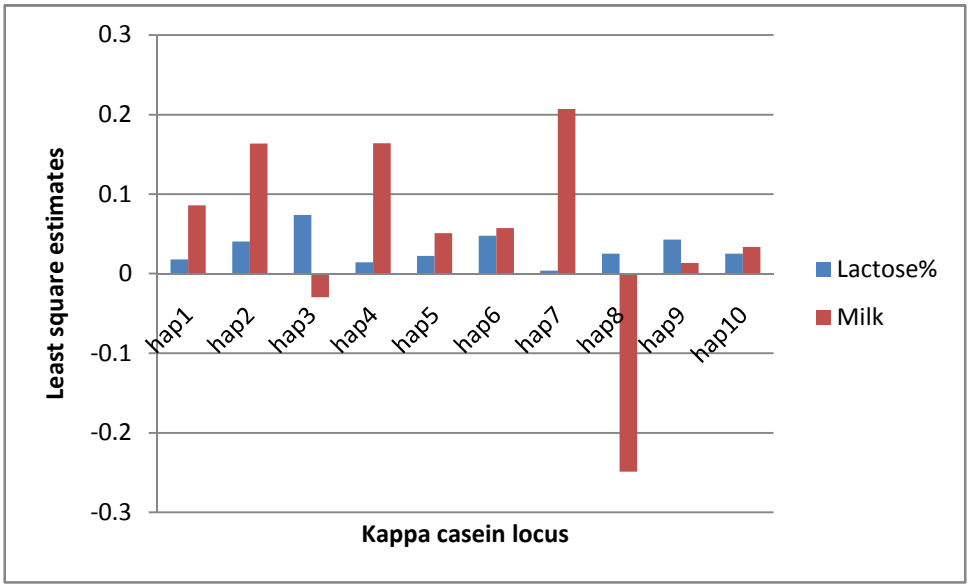

Figure 19. Additive fixed effects of the kappa casein locus haplotypes on lactose% and test day milk yield (kg).

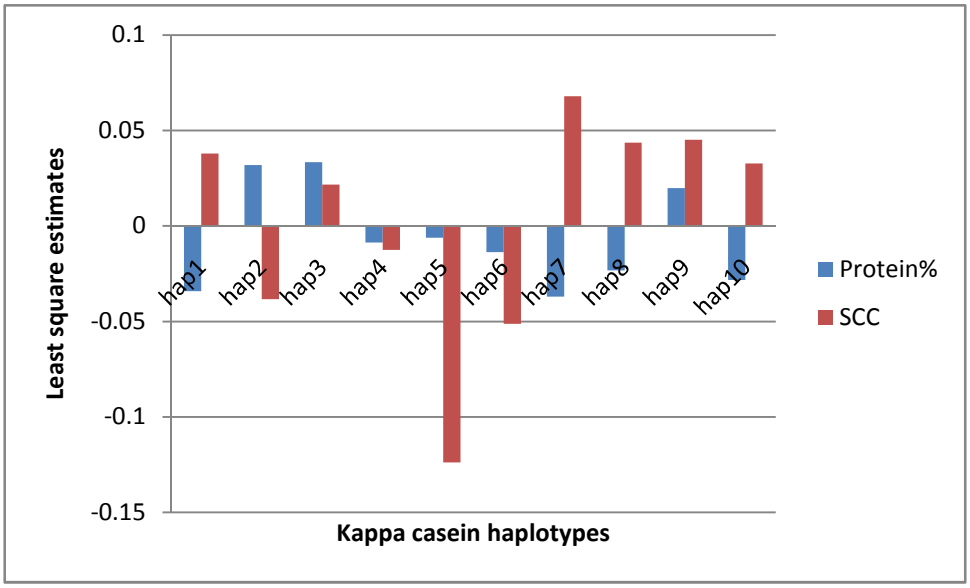

Figure 20. Additive fixed effects of the kappa casein locus haplotypes on protein% and somatic cell count.

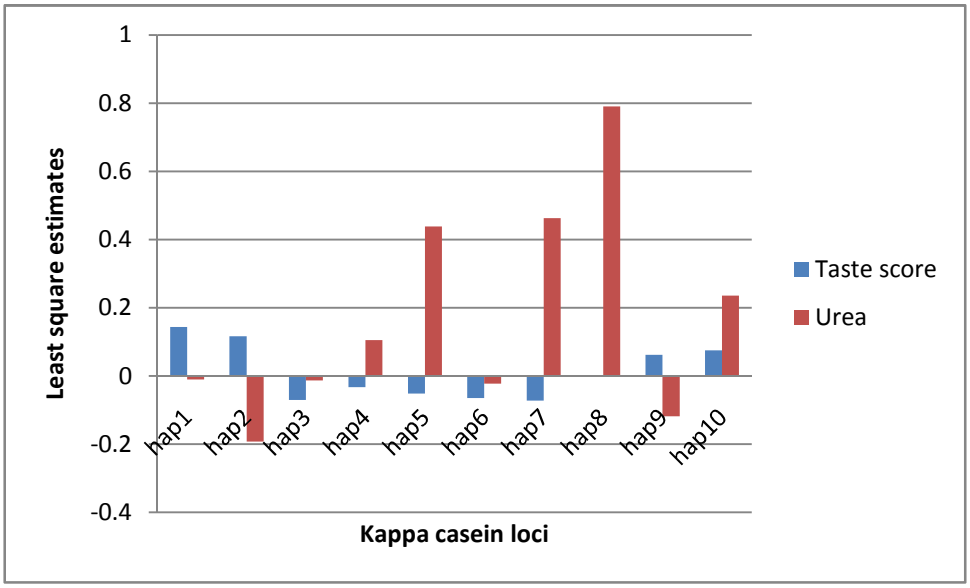

Figure 21. Additive fixed effects of the kappa casein locus haplotypes on taste score and urea content.

4.2.2. Haplotypes dominance effect

4.2.2.1. Entire casein loci haplotypes dominance effect

The entire casein block dominance effect on traits studied is presented in Figure 22. The dominance interaction between haplotypes 7 and 10 has significant effect on fat percentage. Moreover, the dominance interaction between haplotypes 1 and 2, and 1 and 3 on FFA content, and haplotypes 1 and 3, and 1 and 7 on urea content of the milk showed significant effect. However, none of the haplotypes dominance interaction showed significant effect at $p \leq 0.05$ on test day milk yield, protein%, lactose% and milk taste score.

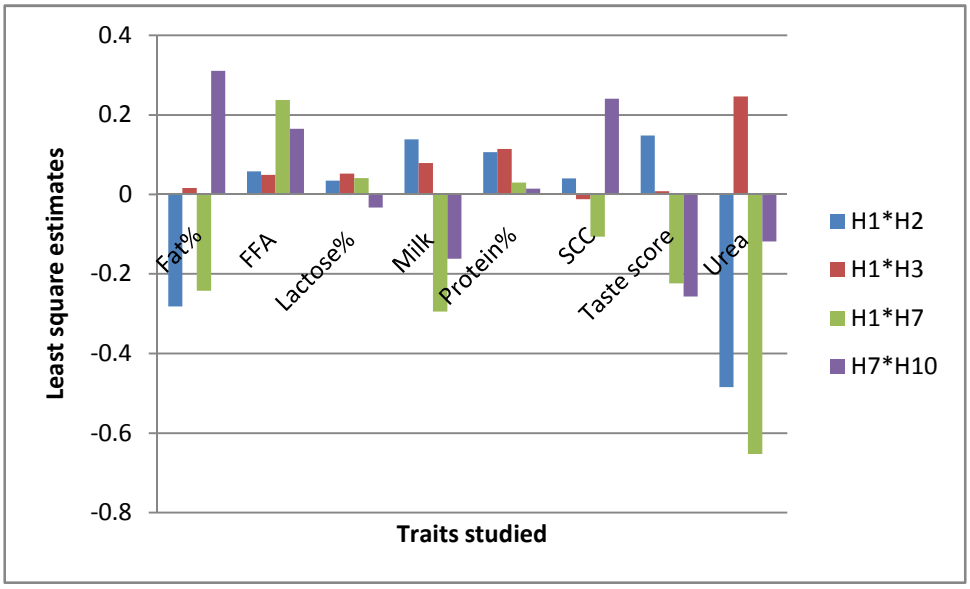

Figure 22. Least sqaure estimates for entire casein loci haplotypes dominance effect.

4.2.2.2. Individual casein genes haplotypes dominance effect

The casein loci haplotypes dominance effect on milk production traits were also estimated at individual genes level.

4.2.2.2.1. αS1-casein gene haplotypes dominance effect

Significant dominance effect of αS1-casein gene were found for haplotypes 1 and 8 on taste score, and haplotypes 1 and 3 on urea content only (Figure 23).

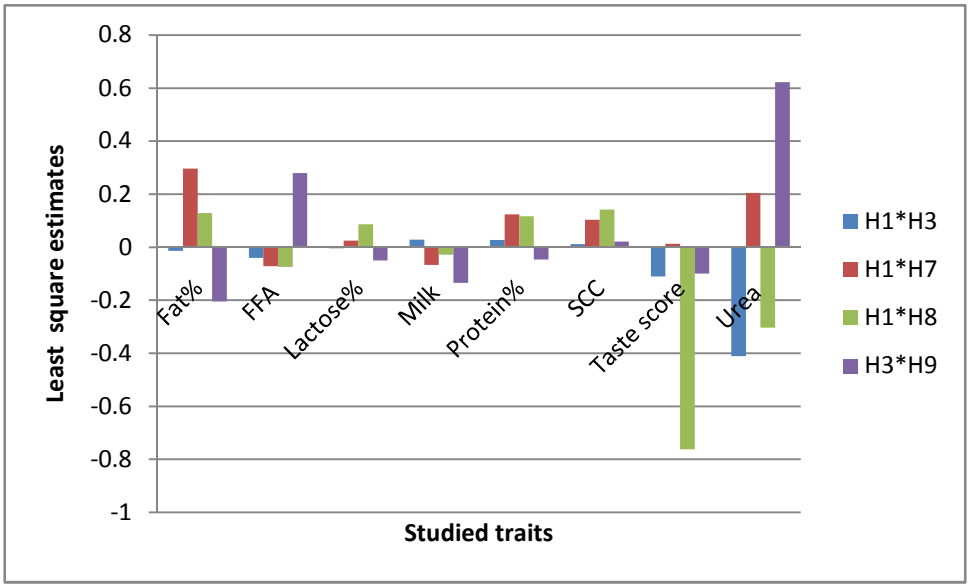

Figure 23. Least sqaure estimates for αS1-casein haplotypes dominance interaction effect.

4.2.2.2.2. β- casein gene haplotypes dominance effect

For β-casein haplotype 1 and 2 dominant interaction has significant effect on SCC (Figure 24).

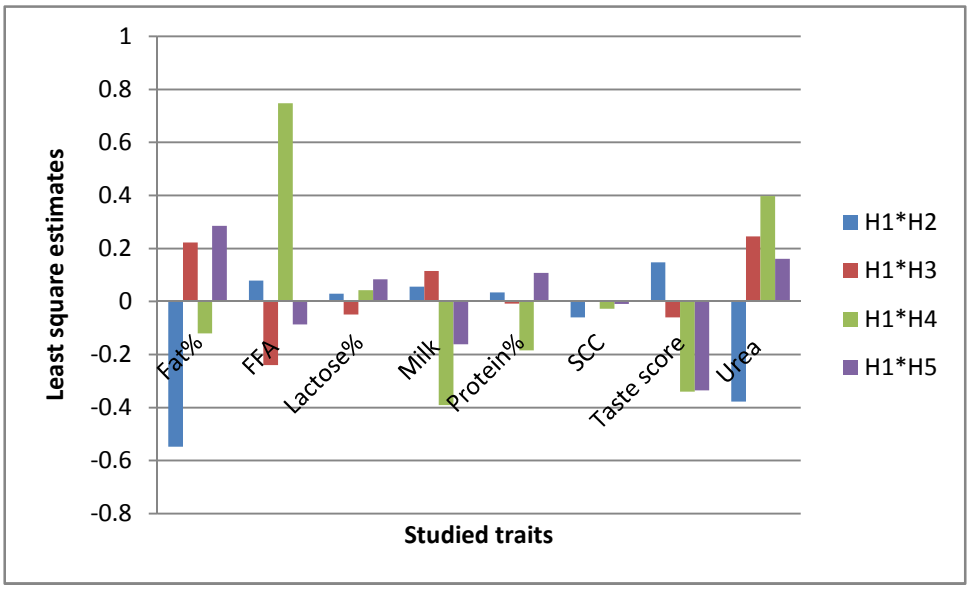

Figure 24. Least sqaure estimates for β-casein haplotypes dominance interaction effect.

4.2.2.2.3. αS2-casein gene haplotypes dominance effect

Our results showed that αS2-casein gene haplotypes 1 and 2 on fat percentage; and haplotypes 1 and 3 both on protein percentage and urea content showed significant effect (Figure 25).

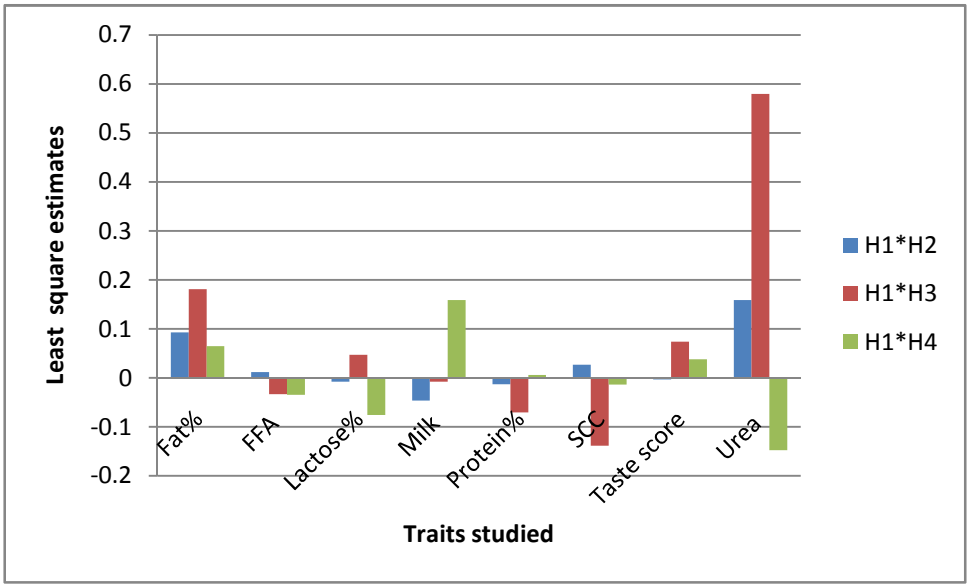

Figure 25. Least sqaure estimates for αS2-casein haplotypes dominance interaction effect.

4.2.2.2.4. κ-casein gene haplotypes dominance effect

For κ-casein haplotype dominance effects were noted for haplotypes 1 and 6 on FFA, and 1 and 4 on urea content (Figure 26).

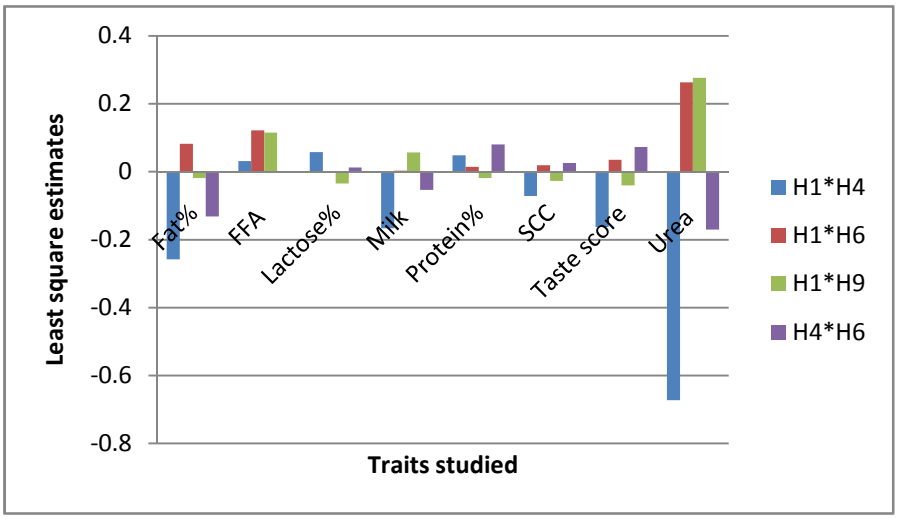

Figure 26. Least sqaure estimates for κ-casein haplotypes dominance interaction effect.

4.3. SNPs additive and dominance effects

Casein SNPs additive and dominance effects were estimated using equation 3 for all studied traits.

4.3.1. Bi-allelic SNPs additive and dominance effects

The least square estimates for additive and dominance effects of casein genes SNPs on milk yield and its component traits shown in Figure 27 to 30 for the additive effect and from Figure 31 through 34 for the dominance effect.

4.3.1.1. Bi-allelic SNPs additive effect

From our analysis; SNPs 1, 24, 25, 26, 29, 30, 31, 32,33, 35 and 36 showed significant additive effect on fat%. Similarily, SNPs 24, 25, 26, 30, 32, 33,34, 35, 36 showed significant additve effect on FFA content. Moreover, SNP18 on test day milk and SNP 8, 24, 25 and 35 on SCC showed significant additive effect. However, for other traits the effect of most of the SNPs is not statistically significant at $p \leq 0.05$. Moreover, for some of the least square estimates the standard errors were large that makes the estimates of the SNP additive effect less reliable.

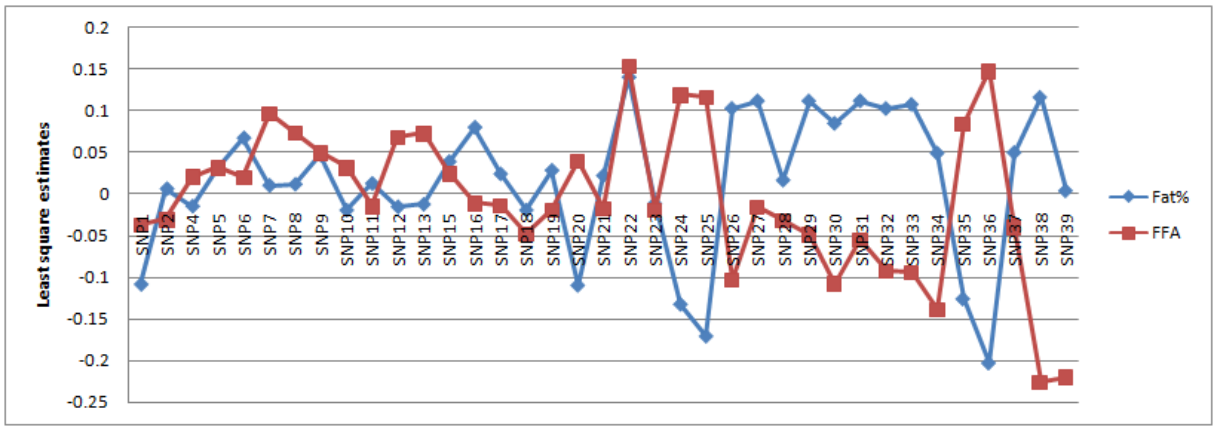

Figure 27. Bi-allelic casein SNPs additive effect on fat% and FFA.

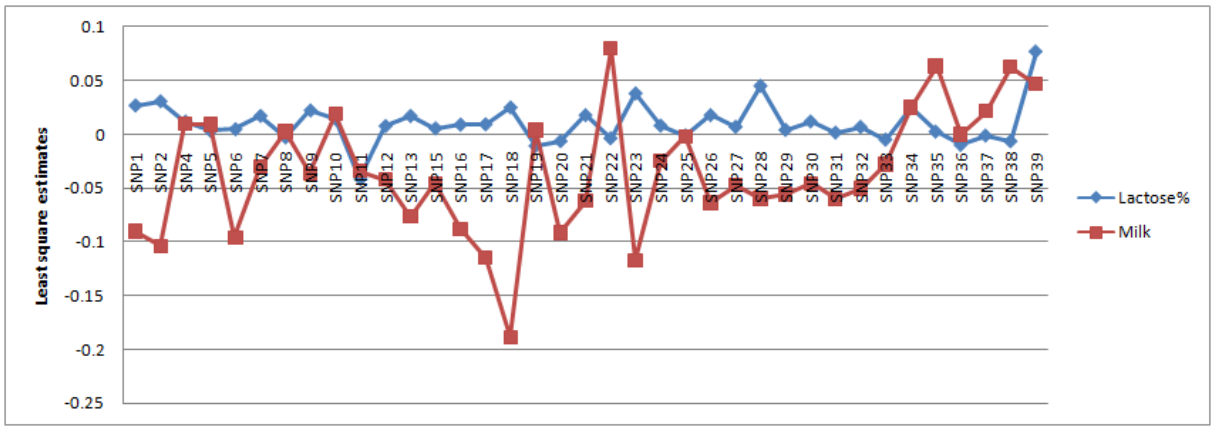

Figure 28. Bi-allelic casein SNPs additive effect on lactose% and test day milk yield (kg).

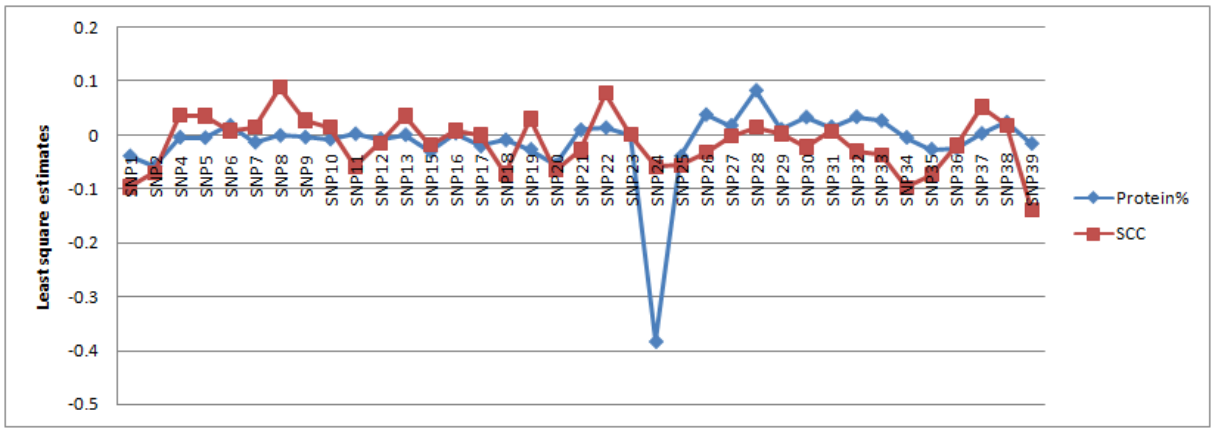

Figure 29. Bi-allelic casein SNPs additive effect on protein% and somatic cell count.

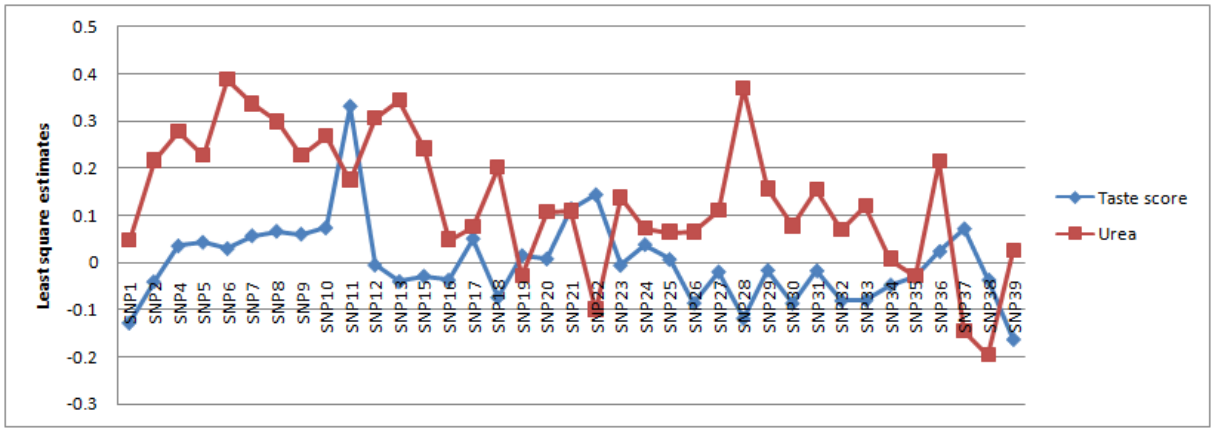

Figure 30. Bi-allelic casein SNPs additive effect on taste score and urea content.

4.3.1.2. Bi-allelic SNPs dominance effect

The dominance effect of casein loci SNPs on milk production taraits is presented in Figure 31 to 34. By and large, the dominace effect of the SNPs on the traits studied is not significant. However, SNPs 20, 27 and 35 showed significant dominance effect on fat%. The dominance effect of SNP 39 on lactose% was also significant. Moreover, the effect of SNP 39 on lactose percentage, SNP 24 on test day milk yield, and SNP 1, 21 and 39 on SCC were significant.

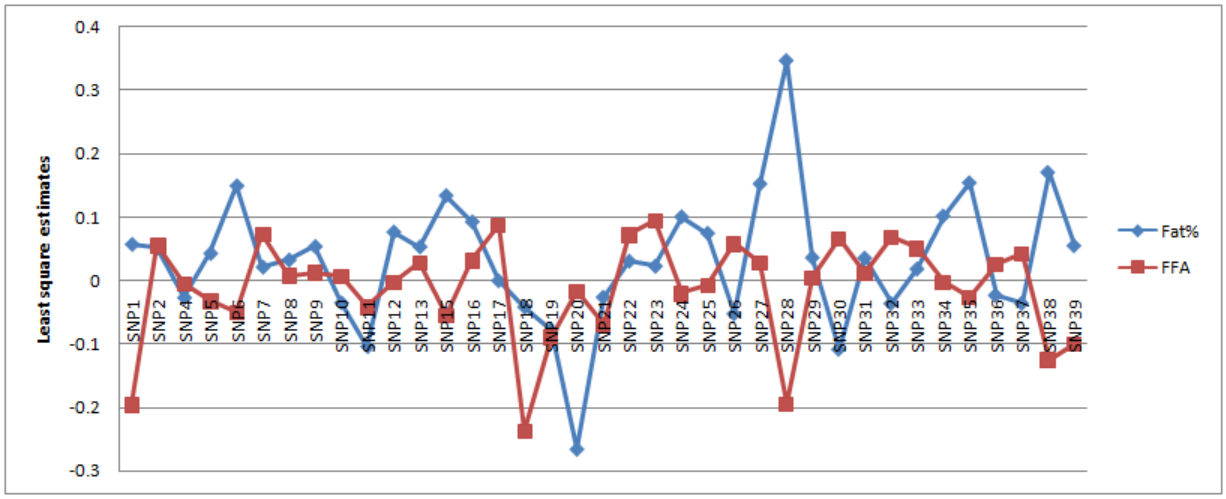

Figure 31. Bi-allelic casein SNPs dominance effect on fat% and FFA.

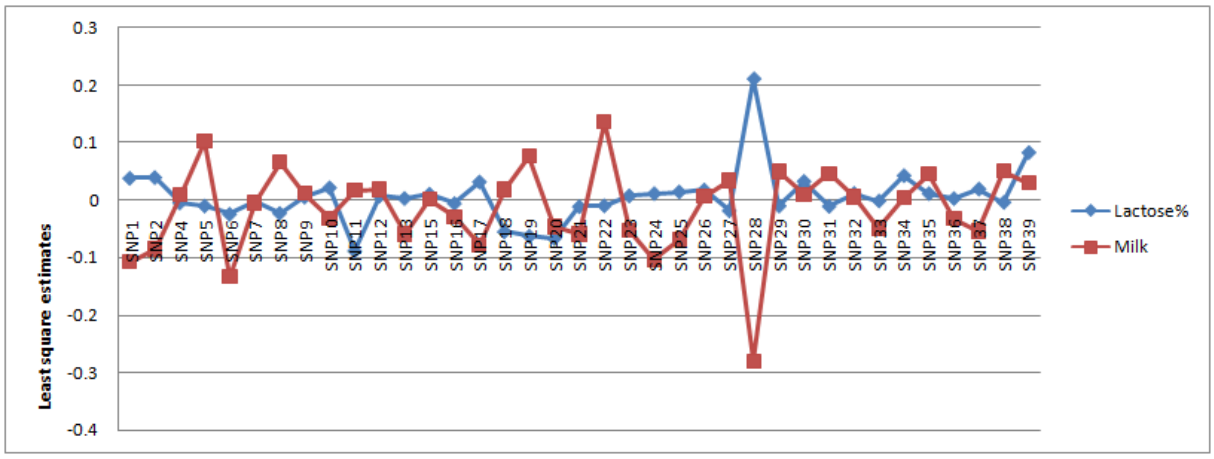

Figure 32. Bi-allelic casein SNPs dominance effect on lactose% and test day milk yield (kg).

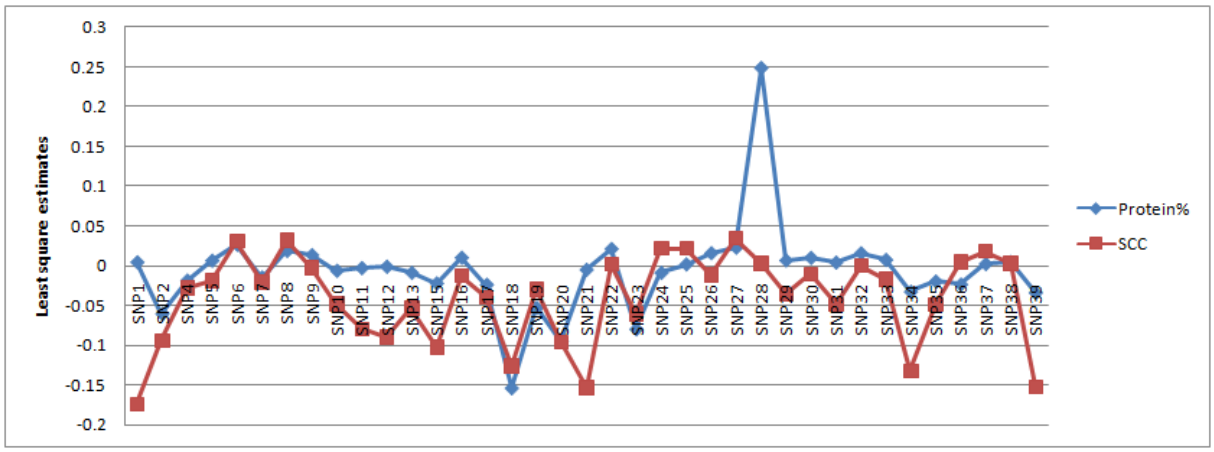

Figure 33. Bi-allelic casein SNPs dominance effect on protein% and somatic cell count.

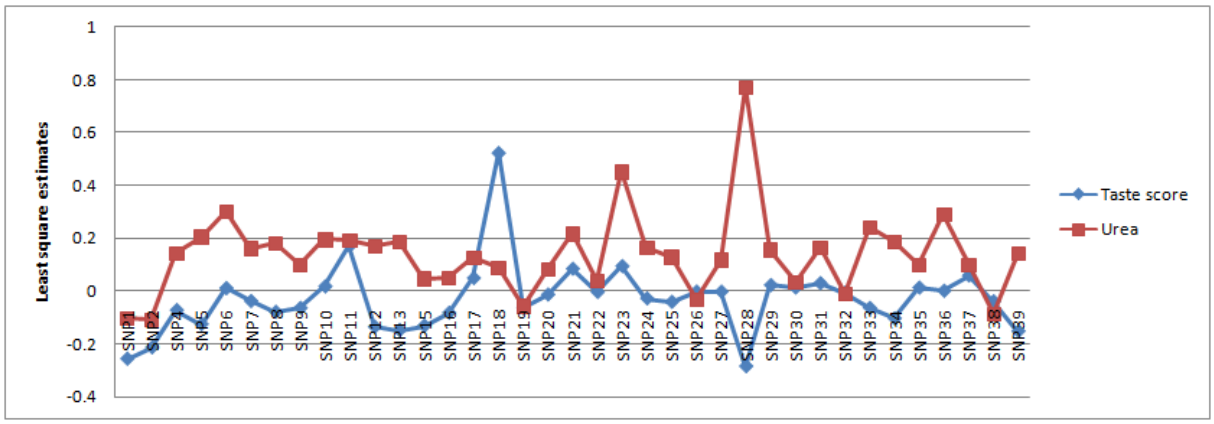

Figure 34. Bi-allelic casein SNPs dominance effect on taste score and urea content.

4.3.2. αS1-casein exon 12 SNP 14 additive and dominance effects

4.3.2.1. αS1-casein exon 12 SNP 14 genotypic effect

Genotypic effect least square estimates for exon 12 deletion in αS1-casein shown in Figure 35. Our findings showed that the least square estimates for genotypic effect of SNP 14 were significant for FFA and urea content only at $p \leq 0.05$.

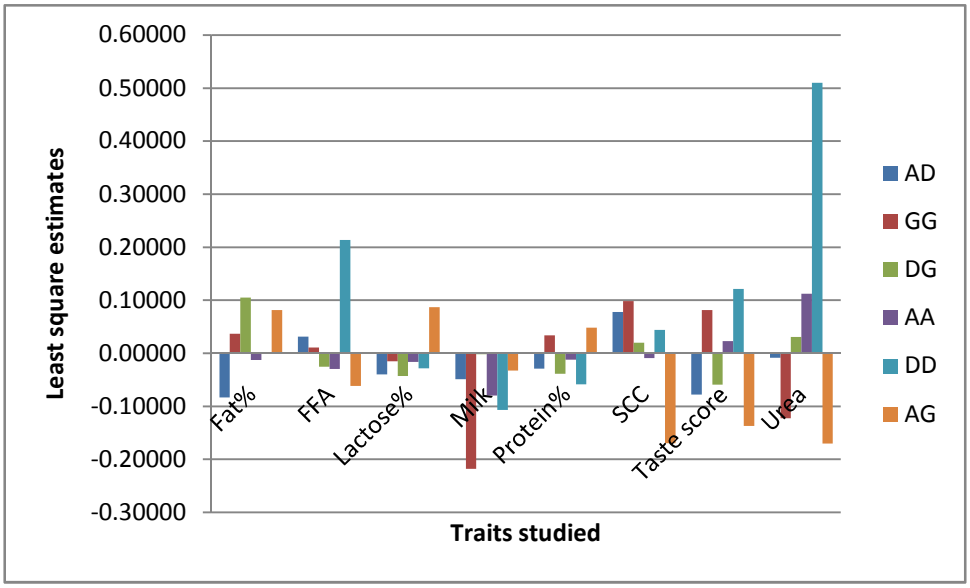

Figure 35. αS1-casein exon 12 SNP14 deletion genotype effects.

4.3.2.2. αS1-casein exon 12 deletion additive and dominance effects

Additive effect of SNP14 is presented in Figure 36. Exon 12 SNP 14 has significant additive genetic effect of allele A on FFA and allele D on FFA and urea content of the milk ($p \leq 0.05$). The AD on SCC, and AG on urea content showed significant dominance effect at $p \leq 0.05$ (Figure 37).

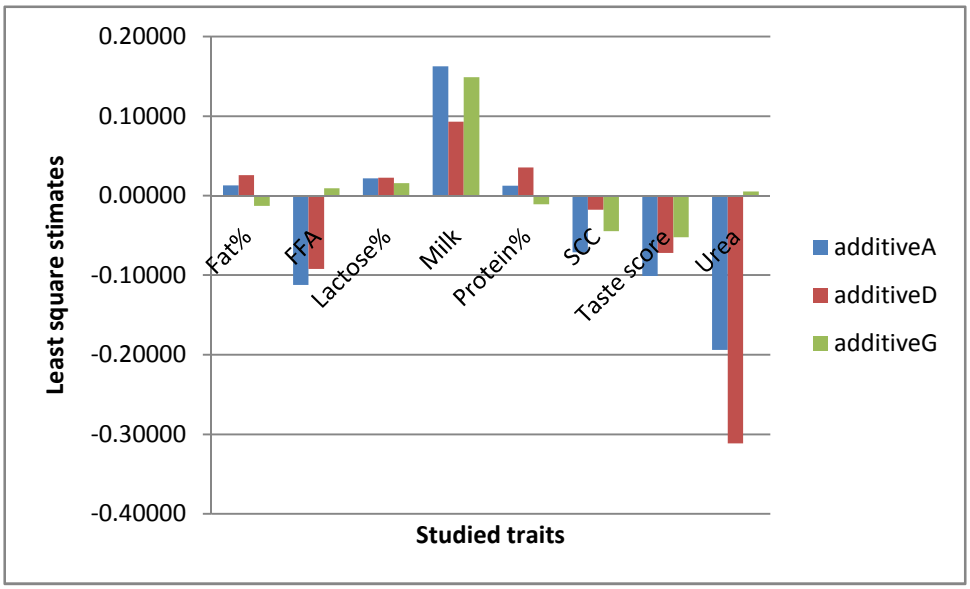

Figure 36. αS1-casein exon12 SNP14 additive effect on studied traits.

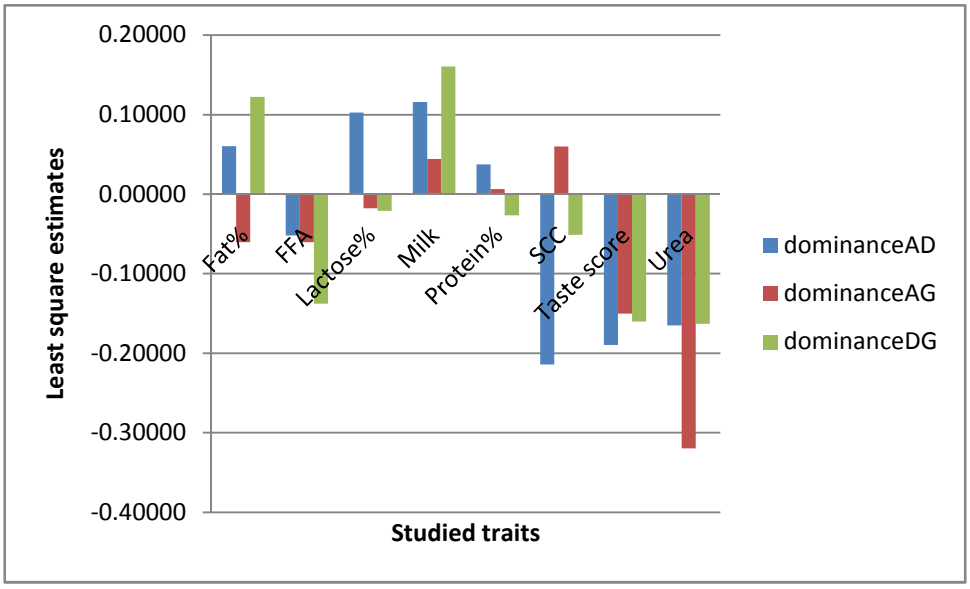

Figure 37. αS1-casein exon12 SNP14 dominance effect on traits studied.

4.4. Allelic frequencies for exon12 deletion

Allele frequencies for exon12 tri-allelic SNP: D (deletion), A and G were calculated for Norwegian goats and the crosses. These frequencies were 0.66, 0.22 and 0.13, and 0.24, 0.55 and 0.22, for D, A and G alleles, in the Norwegian goats and the crosses, respectively. This frequeny was 0.48, 0.36 and 0.16 for D, A and G alleles, respectively for the entire population. Exon12 deletion frequency obtained in our analysis in the Norwegian goats is less than the previous reports (Ådnøy *et al* 2003; Hayes *et al* 2006). This is probably resulted due to the absence of αS1-casein exon12 deletion in the Alpine breed and due to selection against the homozygous deletion.

4.5. Least square estimates of the studied traits for the two study populations

Least square estimates computed by fitting the model in equation 2 without SNP and haplotype effect are presented in Table 9. The only least square estimate significantly differ at $p \leq 0.05$ between the Norwegian goats and the crosses was lactose percentage, for which the crosses were superior. For other traits, even though, the numerical values were larger for the crosses, the two poulations were not significantly different, because most of the estimates have large standard errors.

Table 9. Least square estimates for studied traits in the Norwegian goats and the crosses.

Breed	Fat%	logFFA	Lactose%	Milk	Protein%	logSCC	Taste score	Urea
Norwegian goats	4.023	–0.159	4.524	2.995	2.903	9.692	0.988	10.38
Crosses	4.424	–0.627	4.620	3.278	3.015	8.430	0.919	9.034

5. Discussions

The extensive linkage disequilibrium (LD) observed in our finding especially for αS1-casein is consistent with previous work on Norwegian goats (Hayes *et al* 2006). However, the extent of LD between pairs of SNP markers showed remarkable variation; from complete to nearly zero LD and similar findings were reported (Hayes *et al* 2006; Nilsen *et al* 2009). Moreover, consistent to Hayes *et al* (2006), regions of high LD were not evenly distributed across the chromosome segment (Figures 3 & 4). We found that the LD was much higher between CSN1S1 and CSN2 SNPs and between SNPs in CSN1S2 and CSN3 than between SNPs in CSN2 and CSN1S2 in Norwegian goats (Figure 54), which is similar to the findings of Hayes *et al* (2006). However, the trend observed in the crosses is slightly different. The LD between CSN2 and CSN1S2 is higher in the crosses than it was for the Norwegian goats (Figure 55). The LD extent is relatively stronger within locus than across loci and is in agreement with previous findings (Liu *et al* 2004; Hayes *et al* 2006; Finocchiaro *et al* 2008). Therefore, even if casein genes are found in cluster, they might not be in a strong LD (Finocchiaro *et al* 2008). Crossbreeding is a source of LD (Goddard 1991); however, our findings showed that the level of overall LD is less in the crosses. This might be partly resulted from the effect of within breed artificial selection in the Norwegian goats, which probably increases the LD level than the LD introduced by crossbreeding in the crosses.

Generally, the number of haplotypes found from our analysis is much lower than the expected number and this indicates the presence of extensive LD (Hayes *et al* 2006; Finocchiaro *et al* 2008). More haplotypes were found for the crosses, this is indeed true; because the patrilineally inherited Alpine gene in the crosses has introduced additional Alpine specific haplotypes in the crossbred population. On the other hand, the high level of LD observed in the Norwegian goats (Figure 3) has resulted in low haplotype diversity. This is due to the inverse relationship between the extent of LD and the amount of haplotype diversity. The number of haplotypes obtained in our analysis for individual casein genes is comparable with the previous report for the Norwegian goats (Hayes *et al* 2006).

The two populations shared low number of haplotypes in common for the high polymorphic locus αS1-casein (Table 5 & 6). This is due to high polymorphic nature of this locus (Neveu *et al* 2002; Moatsou *et al* 2004). However, despite the extensive polymorphism in κ-casein

(Sacchi *et al* 2005), the common haplotypes in the two populations for the κ-casein showed the second highest percentage (37%) next to αS2-casein (44%). Despite β-casein's characterstic feature of sequence conservation (Rijnkels 2002), few number of common haplotypes were found for the two populations.

Haplotype diversity within each block can be well explained by a finite of SNPs, called tag SNPs (Liu *et al* 2004). Our results showed much variation in the number of tagging SNPs between the two populations. Fewer tagger SNPs can be used for prediction of other SNPs that are in LD within the casein locus haplotype block for the Norwegian goats than for the crosses. This is due to the high level of LD in the Norwegian goats. The number of tag SNPs obtained in our analysis for Norwegian goats casein SNPs is higher than the report of Hayes *et al* (2006). Therefore, the level of LD detected in our analysis for the Norwegian goats population is less extensive than the findings of Hayes *et al* (2006).

The frequency of SNP haplotype 1 was higher than other haplotypes frequencies obsereved for individual casein genes and the entire casein block (Table 7 & 8). Therefore, it represents the most commonly segregating haplotype in the population. This is most likely happened due to the influence of continious artificial selection in favor of this haplotype. This is indeed true; because haplotypes that provide better protein and casein contents favored by selection (Sztankóová *et al* 2009). However, both in our analysis (see Figures for additive effect of SNP haplotypes in the Result section) and in the work of Hayes *et al* (2006), the commonly found haplotype has rather negative effect on dry matter constituents of milk, which is quite surprising and is against the objectives of the breeding program (Hayes *et al* 2006; Andonov *et al* 2007). Haplotype diversity from our study was higher for *CSN1S1* and *CSN3* genes and low for *CSN1S2* and *CSN2* genes (Table 8). This is due to high polymorphic nature of *CSN1S1* and *CSN3* genes (Martin *et al*., 2002; Moioli *et al*., 2007). This was also proved upon SNP detection, because more SNP markers were found for *CSN1S1* and *CSN3* genes in the genotype data used in our analysis.

Significant entire casein block SNP haplotypes additive effect was observed to a varying degree in our study for the traits studied. This includes the effect of the entire casein block some haplotypes on fat%, and a similar finding was reported by Hayes *et al* (2006) for individual casein genes. For FFA, a reasonable number of haplotypes constructed from the entire casein loci SNP showed significant effect. Protein% is also affected by the casein block

haplotype and a similar finding was reported by Hayes *et al* (2006) for individual casein genes haplotypes and Nilsen *et al* (2009) for Ca-sensitive casein loci block haplotypes in the Norwegian Red cattle. Similarily, SCC and milk taste score were affected by casein block haplotypes. Signficant effect of the entire casein SNP haplotypes dominance interaction was found for fat%, FFA and urea content of the milk. However, none of the haplotypes constructed from the casein loci block were showed significant additive and dominance effect on test day milk yield and lactose percentage.

Haplotypes from individual casein genes also showed significant effect on some milk production traits and a similar findings were reported by Hayes *et al* (2006) for the Norwegian goats. Among the casein genes, αS1-casein is the most studied locus for its effect on milk production traits. Therefore, there are a number of poven evidences for notable effect of this locus on milk production traits (Neveu *et al*. 2002; Trujillo *et al* 2000; Martin *et al* 2002; Feligini *et al* 2005; Chiatti *et al* 2007; Marletta *et al* 2007; Sztankóová *et al* 2007), which substantiates our results for additive effect of αS1-casein haplotypes (Figures 9 to 12). However, Caravaca *et al* (2009) reported absence of significant associations between αS1-casein genotype and total protein, fat, and casein contents. This partly agreed with our findings, for example, the effect of αS1-casein SNP haplotypes found from our analysis was less compared to κ-casein haplotypes.

Most of the studied traits (except protein and lactose percentage) were also significantly affected especially by the additive effect of β-casein gene SNP haplotypes. However, the dominance interaction effect of the most commonly segregating haplotypes was found significant for SCC only. This makes sense, because the function of this locus is to determine certain structural properties of the casein micelle (Rijnkels 2002). Moreover, notable addditive effects of αS2-casein haplotypes were found in our analysis on milk production traits and similar findings were reported (Ramunno *et al* 2001; Marletta *et al* 2004; Chessa *et al* 2007; Vacca *et al* 2009). Rijnkels (2002) reported the function of κ-casein gene haplotypes in casein genes organization, which supports the significant effect found in our analysis for κ-casein gene haplotypes on mik production traits. Similarliy, Hayes *et al* (2006) also reported the significant effect of κ-casein polymorphisms on goat milk quality. Therefore, the variation observed in the effect of casein genes genetic polymorphism on milk production traits can be used in selection and genetic improvement programs. As a result, future selection programs

should exploit both the genetic variations that exist among the entire casein cluster (Sacchi *et al* 2005) and within individual casein genes.

The dominance effect of SNP haplotypes constructed from the entire casein loci and individual casein genes on milk production traits was estimated. Significant effect of dominance interaction casein haplotypes was found for some traits. We also noted that, the additive effect of SNP haplotypes on studied traits is more evident than the dominance effect. Therefore, the variation observed in the additive effect of SNP haplotypes on milk production traits can be utilized through haplotype assisted selection (Hayes *et al* 2006).

Our finding for entire casein block versus individual casein gene SNP haplotypes effect on milk production traits is not in line with literature recommendations (Caroli *et al* 2006; Hayes *et al* 2006; Gigli *et al* 2008). The individual casein genes SNP haplotypes effect was observed for most of studied traits than the effect found by fitting the entire casein segment SNP haplotypes. Therefore, our finding showed that the study of individual casein genes effect on milk production traits for which much has been done at protein variant level has to be a focus of study besides entire casein SNP haplotyes. This less effect obtained by fitting the entire casein loci haplotypes might be due to simultaneous action of up-regulation and down-regulation of different casein genes. If this is the case, when expression of one casein gene is down-regulated, the others can be up-regulated to compensate and this will result into insignificant total sum effect (Hayes *et al* 2006).

Individual SNP additive and dominance effect were estimated for traits studied. The additive effect of one or more individual SNP was found for some of the studied traits. Similar trend was observed for dominance effect, however, the number of SNPs that showed dominance effect was less. The effect of individual SNP found in our analysis varies across casein loci. For example, most of the SNPs that showed significant effect on milk production traits localized in αS2-casein and κ-casein. Even within these loci the effect was further localized among adjacent SNPs. For example, signifcant effect of κ-casein SNP 24 and 25 was found for most of the traits in question. A similar trend about localization of SNP effect was also reported by Hayes *et al* (2006). Nilsen *et al* (2009) also reported a number of SNPs that showed significant effect on milk production traits from αS2-casein in the Norwegian red cattle. Therefore, relatively less effect found from αS1-casein SNPs in our analysis is in agreement with the findings of Nilsen *et al* (2009). Moreover, the effect of few individual

SNP on milk production traits was reported by Hayes *et al* (2006) in the Norwegian goats. Additional polymorphism introduced in the crosses that are included in our analysis most likely resulted in large number of SNPs showing significant effect compared to the findings of Hayes *et al* (2006) that based on the study of Norwegian goats only.

The αS1-casein exon 12 deletion is unique for the Norwegian goat population (Hayes *et al* 2006) and hence it becomes interesting area of research in this population. The αS1-casein exon 12 deletion effect on the dry matter content of milk was reported in the Norwegian goats (Hayes *et al* 2006). Our results obtained by fitting SNP 14 genotypes that harbors exon 12 deletion as class explanatory variable showed significant effect on FFA and urea content of milk. As it was noted by Hayes *et al* (2006) the genotypic effect that harbour the deletion showed negative effect on dry matter content of the milk. Similarly, the additive genetic effect of this deletion was significant for FFA content and urea content at $p \leq 0.05$, and protein percentage at $p \leq 0.1$ (Figure 36), which agrees partly with previous report (Hayes *et al* 2006). However, unlike the report of Hayes *et al* (2006) positive trend (it is not statistically siginificant) was obsereved (Figure 36) for the additive effect of D allele on milk dry matter contents and this disagrees from our result obtained by fitting genotypic effect. However, the dominance interaction of the two mutant alleles DG showed negative but insignificant effect on protein percentage, lactose percentage and FFA content of the milk. This inconsistency in SNP14 additive effect might be arosed from low proportion of D allele (0.48) for the entire goat population used in our analysis compared to the Hayes *et al* (2006) finding that was based on relatively high frequency of D allele (0.745) in the Norwegian goats population. Significant dominant interaction effect of this unique deletion for AD allelic combination was detected on SCC at $p \leq 0.05$ and milk taste score at $p \leq 0.1$, and DG allelic combination showed significant effect on FFA at $p \leq 0.1$ (Figure 37).

For all traits studied, the LSD-test was used to compare the least square estimates for Norwegian goats and the crosses. Our results showed that the estimates were not signficantly different except for the lactose percentage of milk. This would be interesting for institutions that involve in the genetic improvement of Norwegian goats, because increasing the lactose percentage of the milk is one of the objectives of the breeding program (Andonov *et al* 2007). Therefore, the effect of this crossbreeding program in creating superior crossbred population for other traits considered is not signficantly high. However, the crosses are at least as good as the Norwegian goats in other traits. This crossbreeding program can be also used as a source

of additional genetic variation which inturn would create a wide genetic pool on which selection and genetic improvement programs can rely. Since the national breeding program is also aimed at increasing lactose content of the milk (Andonov *et al* 2007), the superiority of the crosses in lactose percentage can be exploited through crossbreeding program. However, this might be resulted due to heterotic effect in the F1 population. Therefore, this requires comparative study of Norwegian goats and French Alpine under similar production environment to reach at safe conclusion.

6. Conclusion and Recommendation

6.1. Conclusion

From our results, the level of LD was less extensive in the crosses, which in turn resulted in high haplotype diversity in the crosses. This showed the advantage of this crossbreeding program in creating additional genetic variation. There is noticable additive and dominance effects of some casein loci SNPs and SNP haplotypes for most of the traits studied. Therefore, these variations in haplotypes effect can be used in genetic improvement programs through haplotype assisted selection. However, very stringent results were obtained in our analysis showing signficant negative effect of the commonly segregating SNP haplotypes on the dry matter contents of the milk, which are against the objectives of the breeding program. Significant additive and dominance effect of exon 12 deletion was observed on FFA, urea content and SCC. Therefore, compared to previous reports (Hayes *et al* 2006), the effect of exon 12 deletion was less pronounced in our findings. This is due to the absence of this unique deletion in the haplotypes inherited from Alpine goats in the crosses. There is significant difference in the lactose percentage between the two study populations, which shows the importance of this crossbreeding program. However, further comparative studies of the two breeds under comparable environment is required to substantiate whether this difference is due to heterotic effect or due to real genetic difference between two breeds.

6.2. Recommendations

This piece of work tried to address research questions that limited to its very nature, however, we would like to recommend researchable problems for future works:

- Studying the effect of tightly linked loci haplotypes effect in CSN1S1 and CSN2 block *vs* in CSN1S2 and CSN3 block on milk production traits.
- Due to their evolutional variation, comparative study of Ca-sensitive casein loci block *vs* CSN3 effect on milk production traits is an interesting area of research.
- The negative effect of commonly segregating SNP haplotypes on dry matter contents of milk needs to be further investigated.

- Having variance components estimated from relatively large dataset for milk taste score and milk urea content would help to reach at more reliable results.
- There might be variation among the judges in milk taste scoring. This can be seen from small R^2 value obtained from SAS output in Table 18, which could be partly due absence of judges effect in the fitted model.
- The crosses showed significantly higher estimate for lactose percentage. Therefore, further research is needed to differentiate hetrosis from true genetic difference between two breeds.

7. References

Ådnøy T., Vegarud G., Devold T. G., Nordbø R., Colbjørnsen I. *et al*. 2003. Effects of the 0- and F-alleles of alpha S1 casein in two farms of northern Norway. Proceedings of the International Workshop on Major Genes and QTL in Sheep and Goat, Toulouse, France.

Ådnøy T. , Hagesæther N., Hayes B., Lien S. 2006. Use of information about single genes in breeding–Exemplified by casein genes and milk in goats. 8th World Congress on Genetics Applied to Livestock Production, August 13–18, 2006, Belo Horizonte, MG, Brazil.

Albenzio M., Santillo A., d'Angelo F. Sevi A. 2009. Focusing on casein gene cluster and protein profile in Garganica goat milk. Journal of Dairy Research, 76: 83–89.

Andonov S., Ødegård J., Boman I. A., Svendsen M., Holme I. J., Ådnøy T., Vukovic V., Klemetsdal G. 2007. Validation of Test-Day Models for Genetic Evaluation of Dairy Goats in Norway. J. Dairy Sci., 90: 4863–4871.

Asheima L.J. and Eik L.O. 1998. The economics of fibre and meat on Norwegian dairy goats. Small Ruminant Research, 30: 185–190.

Bagnicka E., Wallin E., Łukaszewicz M., Ådnøy T. 2007. Heritability for reproduction traits in Polish and Norwegian populations of dairy goat. Small Ruminant Research, 68: 256–262.

Barrett JC, Fry B, Maller J, Daly MJ. 2005. HaploView: analysis and visualization of LD and haplotype maps. Bioinformatics, 21: 263–265.

Bovenhuis H., Van Arendonk J. A. M., Korver S. 1992. Associations between milk protein polymorphisms and milk production traits. J Dairy Sci, 75: 2549–2559.

Bowling G. A. 1929. Dairy Goats. Extension Bulletin No. 85. Accessed on 1 March 2010 (http://web1.msue.msu.edu/imp/modp1/morefile/E0085.pdf).

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

Caroli A., Chiatti F., Chessa S., Rignanese D., Bolla P., Pagnacco G. 2006. Focusing on the Goat Casein Complex. J. Dairy Sci., 89: 3178–3187.

Caroli A., Chiatti F., Chessa S., Rignanese D., Ibeagha–Awemu E. M., Erhardt G. 2007. Characterization of the Casein Gene Complex in West African Goats and Description of a New αS1-Casein Polymorphism. J. Dairy Sci., 90: 2989–2996.

Chanat E., Martin P. and Ollivier–Bousquet M. 1999. αS1-casein is required for the efficient transport of β- and κ-casein from the endoplasmic reticulum to the Golgi apparatus of mammary epithelial cells. Journal of Cell Science, 112: 3399–3412 .

Chessa S., Rignanese D., Chiatti F., Radeghieri A., Gigliotti C., Caroli A. 2007. Technical Note: Simultaneous Identification of CSN1S2 A, B, C, and E Alleles in Goats by Polymerase Chain Reaction–Single Strand Conformation Polymorphism. J. Dairy Sci., 91: 1214–1217.

Chiatti F., Chessa S., Bolla P., Cigalino G., Caroli A., Pagnacco G. 2007. Effect of κ-Casein Polymorphism on Milk Composition in the Orobica Goat. J. Dairy Sci., 90: 1962–1966.

Dubeuf J.–P. 2005. Structural, market and organisational conditions for developing goat dairy production systems. Small Ruminant Research, 60: 67–74.

Dubeuf J.–P. and Boyazoglu J. 2009. An international panorama of goat selection and breeds. Livestock Science, 120: 225–231.

Dubeuf J.–P., Morand-Fehr P., and Rubino R. 2004. Situation, changes and future of goat industry around the world. Small Ruminant Research, 51: 165–173.

Feligini M., Cubric-Curik V., Parma P., Curik I., Greppi G. F., Enne G. 2002. Polymorphism of κ-Casein in Goat Breeds: A New ACRS–PCR Designed DNA Test for Discrimination of A and B Alleles. Food Technol. Biotechnol., 40 (4): 293–298 .

Feligini M., Frati S., Curik V. C., Brambilla A., Parma P., Curik I., Greppi G. F. and Enne G. 2005. Caprine αS1-Casein Polymorphism: Characterisation of A, B, E and F Variants by Means of Various Biochemical and Molecular Techniques. Food Technol. Biotechnol., 43 (2): 123–132.

Garg N., Singh S.K., Rout P.K. and Mandal A. 2009. Genetic polymorphism of milk proteins in Barbari goats: Short Note. Tropical and Subtropical Agroecosystems, 11: 181–183.

Gigli I., Maizon D. O., Riggio V., Sardina M. T., and Portolano B. 2008. Casein Haplotype Variability in Sicilian Dairy Goat Breeds: Short Communication. J. Dairy Sci., 91: 3687–3692.

Gilmour A.R., Gogel B.J., Cullis B.R., Thompson R. 2009. ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK www.vsni.co.uk

Goddard M.E. 1991. Mapping genes for quantitative traits using linkage disequilibrium. Genet Sel Evol Suppl, 23 (1): 131–134.

Hayes B., Hagesæther N., Ådnøy T., Pellerud G., Berg P., Lien S. 2006. Effects on production traits of haplotypes among casein genes in Norwegian goats and evidence for a site of preferential recombination. Genetics, 174: 455–464.

Hudson R. R. 1985. The sampling distribution of linkage disequilibrium under an infinite allele model without selection. Genetics, 109: 611–631.

Kevorkian S. E. M., Manea , M. A. Gavrila M. G., Rebedea M., Georgescu S. E., Costache M. 2009. Sequencing of exon three of αS1-casein sheep and goat gene. Archiva Zootechnica, 12 (3): 87–92.

Leroux C., and Martin P. 1996. The caprine αS1 and beta-casein genes are 12-kb apart and convergently transcribed. Anim. Genet., 27: 93.

Lien S., Gomu-Raya L.. Steine T., Fimland E., Rogne S. 1995. Associations between Casein haplotypes and milk yield traits. J Dairy Sci , 78: 2047–2056.

Liu T., Johnson J. A., Casella G., Wu R. 2004. Sequencing Complex Diseases With HapMap. Genetics, 168: 503–511.

Lundén A., Nilsson M., Janson L. 1997. Marked effect of b-lactoglobulin polymorphism on the ratio of casein to total protein in milk. J. Dairy Sci., 80: 2996–3005.

Marletta D., Bordonaro S., Guastella A.M., Falagiani P., Crimi N., D'Urso G. 2004. Goat milk with different αS2-casein content: analysis of allergenic potency by REAST-inhibition assay. Small Ruminant Research, 52: 19–24.

Marletta D., Bordonaro S., Guastella A.M., Criscione A., D'Urso. G. 2005. Genetic polymorphism of the calcium sensitive caseins in sicilian Girgentana and Argentata dell'Etna goat breeds. Small Ruminant Research, 57: 133–139.

Marletta D., Criscione A., Bordonaro S., Guastella A. M., D'Urso G. 2007. Casein polymorphism in goat's milk: Review. Lait, 87: 491–504.

Martin P. and Grosclaude F. 1993. Improvement of milk protein quality by gene Technology. Livestock Production Science, 35: 95–115.

Martin P., Ollivier-Bousquet M., Grosclaude F. 1999. Genetic polymorphism of caseins: a tool to investigate casein micelle organization. International Dairy Journal, 9: 163–171.

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: 433–459.

Moatsou G., Samolada M., Panagiotou P., Anifantakis E. 2004. Casein fraction of bulk milks from different caprine breeds. Food Chemistry, 87: 75–81.

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

Neveu, C., Mollé, D., Moreno, J., Martin, P., & L_eonil, J. 2002. Heterogeneity of caprine beta-casein elucidated by RP–HPLC/MS: Genetic variants and phosphorylations. Journal of Protein Chemistry, 21 (8): 557–567.

Nilsen H., Olsen H. G., Hayes B., Sehested E., Svendsen M., Nome T., Meuwissen T., Lien S. 2009. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. Genetics Selection Evolution, 41: 24.

Ramunno L., Cosenza G., Pappalardo M., Longobardi E., Gallo D., Pastore N., Di Gregorio P., Rando A., 2001. Characterization of two new alleles at the goat CSN1S2 locus. Anim. Genet., 32: 264–268.

Ramunno L., Cosenza G., Rando A., Illario R., Gallo D., Berardino D. D., Masina P. 2004. The goat αS1-casein gene: gene structure and promoter analysis. Gene, 334: 105–111.

Raynal-Ljutovac K., Gaborit P. and Lauret A. 2005. The relationship between quality criteria of goat milk, its technological properties and the quality of the final products. Small Ruminant Research, 60: 167–177.

Reale S., Yahyaoui M. H., Folch J. M., Sanchez A., Pilla F., Angiolillo A. 2005. Genetic polymorphism of the κ-casein (CSN3) gene in goats reared in Southern Italy: Short Communication. Ital. J. Anim. Sci., 4: 97–101.

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.

Rincon E. J., Schermerhorn E. C., Dowell R. E. MC and Daniel B. T. MC. 1982. Estimation of Genetic Effects on Milk Yield and Constituent Traits in Crossbred Dairy Cattle. J Dairy Sci, 65: 848–856.

Sacchi P., Chessa S., Budelli E., Bolla P., Ceriotti G., Soglia D., Rasero R., Cauvin E., and Caroli A. 2005. Casein Haplotype Structure in Five Italian Goat Breeds. J. Dairy Sci., 88: 1561–1568.

Sæther N.H. (Ed.). 2002. Norway Country Report on Farm Animal Genetic Resources, 2002. Committee on Farm Animal Genetic Resources. Accessed on 17 Feb. 2010. ftp://ftp.fao.org/docrep/fao/011/a1250f/annexes/CountryReports/Norway.pdf)

SAS. 2002. Statistical Analysis System version 9.00. SAS Institute Inc., Cary, NC, USA.

Scherf B.D. (Ed.). 2000. World Watch List of Domestic Animal Diversity, third ed. Food and Agriculture Organization of the United Nations, Rome, Italy, p. 726.

Shrestha J.N.B., Fahmy M.H. 2007. Breeding goats for meat production: 2. Crossbreeding and formation of composite population–Review. Small Ruminant Research, 67: 93–112.

Stephens M., Smith N. J., Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics, 68: 978–989.

Sulimova G. E., Ahani Azari M., Rostamzadeh J., Mohammad Abadi M. R. and Lazebny O. E. 2007. κ-Casein Gene (CSN3) Allelic Polymorphism in Russian Cattle Breeds and Its Information Value as a Genetic Marker. Russian Journal of Genetics, 43 (1): 73–79.

Sztankóová Z., Mátlová V., Malá G. 2007. Genetic polymorphism at the CSN1S1 gene in two Czech goat breeds. Czech J. Anim. Sci., 52 (7): 199–202.

Sztankóová Z., Mátlová V., Kyseľová J., Jandurová O. M., Říha J., and Senese C. 2009. Polymorphism of casein cluster genes in Czech local goat breeds–Short communication. J. Dairy Sci., 92: 6197–6201.

Templeton A. R., Maxwell T., Posada D., Stengård J.H., Boerwinkle E. and Sing C. F. 2005. Tree Scanning: A Method for Using Haplotype Trees in Phenotype/Genotype Association Studies. Genetics, 169: 441–453.

Threadgill D. W. and Womack J. E. 1990. Genomic analysis of the major bovine milk protein genes. Nucleic Acids Research, 18 (23): 6935–6942.

Trujillo A. J., Casals I., Guamis B. 2000. Analysis of Major Caprine Milk Proteins by Reverse-Phase High-Performance Liquid Chromatography and Electrospray Ionization-Mass Spectrometry. J Dairy Sci, 83:11–19.

Vacca G.M., Ouled Ahmed Ben Ali H., Pazzola M., Sanna M., Dettori M.L. and Carcangiu V. 2009. An investigation on allele frequency at the *CSN1S2* locus and its relationship with milk parameters in the Sarda goat. Journal of Animal and Feed Sciences, 18: 628–637.

Yahyaoui M. H., Angiolillo A., Pilla F., Sanchez A., and Folch J. M. 2003. Characterization and Genotyping of the Caprine κ-Casein Variants. J. Dairy Sci., 86: 2715–2720.

8. Appendices

Table 10. Summary of flock structure of the two goat populations.

Parent	Breed	N	Number of offsprings			
			Mean	Std.dev	Min	Max
Sire	Norwagian	29	7.45	5.94	1	18
	Crosses	15	10.67	3.37	6	16
Dam	Norwagian	318	1.27	0.45	1	2

Table 11. GLM output for fixed effects on test day milk yield.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	32	276.0603228	8.6268851	35.12	<.0001
Error	1617	397.2068530	0.2456443		
Corrected Total	1649	673.2671758			

R-Square	Coeff Var	Root MSE	milk Mean
0.410031	20.75377	0.495625	2.388121

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	69.2620523	8.6577565	35.25	<.0001
breed	1	18.6092925	18.6092925	75.76	<.0001
lactstg	9	129.9205386	14.4356154	58.77	<.0001
season	2	3.9302630	1.9651315	8.00	0.0003
kids number	2	3.0234253	1.5117127	6.15	0.0022
fdstr	1	13.1672587	13.1672587	53.60	<.0001
farm*breed	8	11.3420704	1.4177588	5.77	<.0001
kidding date	1	8.7201229	8.7201229	35.50	<.0001

Table 12. GLM output for fixed effects on somatic cell count.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	55	98011555.4	1782028.3	3.19	<.0001
Error	1478	826319209.9	559079.3		
Corrected Total	1533	924330765.3			

R-Square	Coeff Var	Root MSE	somtic cell count Mean
0.106035	117.1815	747.7161	638.0834

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	7918615.29	1979653.82	3.54	0.0070
breed	1	2779457.67	2779457.67	4.97	0.0259
lactstg	9	5627465.27	625273.92	1.12	0.3460
season	2	197163.58	98581.79	0.18	0.8384
dkontroll	33	35949833.99	1089388.91	1.95	0.0011
kids number	2	3587467.60	1793733.80	3.21	0.0407

Table 13. GLM output for fixed effects on fat percentage.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	22	535.850717	24.356851	41.99	<.0001
Error	1546	896.839070	0.580103		
Corrected Total	1568	1432.689787			

R-Square	Coeff Var	Root MSE	fat Mean
0.374017	17.84300	0.761645	4.268591

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	225.8865499	28.2358187	48.67	<.0001
breed	1	21.2535119	21.2535119	36.64	<.0001
lactstg	9	74.8151885	8.3127987	14.33	<.0001
season	2	1.1943591	0.5971796	1.03	0.3575
fdstr	1	17.2717364	17.2717364	29.77	<.0001
kidding date	1	2.9584700	2.9584700	5.10	0.0241

Table 14. GLM output for fixed effects on lactose percentage.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	29	63.0373016	2.1737001	57.13	<.0001
Error	1547	58.8557281	0.0380451		
Corrected Total	1576	121.8930297			

R-Square	Coeff Var	Root MSE	lact Mean
0.517153	4.430249	0.195051	4.402720

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	8.60802588	1.07600323	28.28	<.0001
breed	1	1.34142734	1.34142734	35.26	<.0001
lactstg	9	37.30507195	4.14500799	108.95	<.0001
season	2	1.64248365	0.82124182	21.59	<.0001
fdstr	1	0.70461549	0.70461549	18.52	<.0001
farm*breed	8	0.61915790	0.07739474	2.03	0.0393

Table 15. GLM output for fixed effects on protein percentage.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	30	57.7371381	1.9245713	43.09	<.0001
Error	1529	68.2930196	0.0446652		
Corrected Total	1559	126.0301577			

R-Square	Coeff Var	Root MSE	prot Mean
0.458122	6.808592	0.211341	3.104038

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	8.31894171	1.03986771	23.28	<.0001
breed	1	2.76824435	2.76824435	61.98	<.0001
lactstg	9	31.00174443	3.44463827	77.12	<.0001
season	2	0.57514554	0.28757277	6.44	0.0016
fdstr	1	10.74545404	10.74545404	240.58	<.0001
farm*breed	8	1.53341310	0.19167664	4.29	<.0001
kidding date	1	2.23191059	2.23191059	49.97	<.0001

Table 16. GLM output for fixed effects on urea content.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	21	3739.288576	178.061361	61.44	<.0001
Error	1507	4367.541980	2.898170		
Corrected Total	1528	8106.830556			

R-Square	Coeff Var	Root MSE	urea Mean
0.461252	17.72876	1.702401	9.602485

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	1012.352770	126.544096	43.66	<.0001
breed	1	144.845893	144.845893	49.98	<.0001
lactstg	9	635.362140	70.595793	24.36	<.0001
season	2	6.645652	3.322826	1.15	0.3180
fdstr	1	1156.075947	1156.075947	398.90	<.0001

Table 17. GLM output for fixed effects on free fatty acids content.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	30	119.6234168	3.9874472	12.54	<.0001
Error	1498	476.3756283	0.3180078		
Corrected Total	1528	595.9990451			

R-Square	Coeff Var	Root MSE	ffa Mean
0.200711	59.76547	0.563922	0.943558

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	8	35.23943257	4.40492907	13.85	<.0001
breed	1	34.13333245	34.13333245	107.33	<.0001
lactstg	9	7.68136157	0.85348462	2.68	0.0043
season	2	1.78896583	0.89448291	2.81	0.0604
fdstr	1	5.31350389	5.31350389	16.71	<.0001
farm*breed	8	9.93331569	1.24166446	3.90	0.0001
kidding date	1	4.26534011	4.26534011	13.41	0.0003

Table 18. GLM output for fixed effects on milk taste score.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	41	33.1203865	0.8078143	3.46	<.0001
Error	1345	314.0259726	0.2334766		
Corrected Total	1386	347.1463590			

R-Square	Coeff Var	Root MSE	testscor Mean
0.095408	41.62672	0.483194	1.160779

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farm	6	1.63283250	0.27213875	1.17	0.3222
breed	1	5.75937317	5.75937317	24.67	<.0001
lactstg	9	2.08532478	0.23170275	0.99	0.4442
season	2	0.55899119	0.27949560	1.20	0.3024
fdstr	1	0.00031776	0.00031776	0.00	0.9706
testdate	14	8.45693040	0.60406646	2.59	0.0011
farm*breed	7	5.05606934	0.72229562	3.09	0.0031

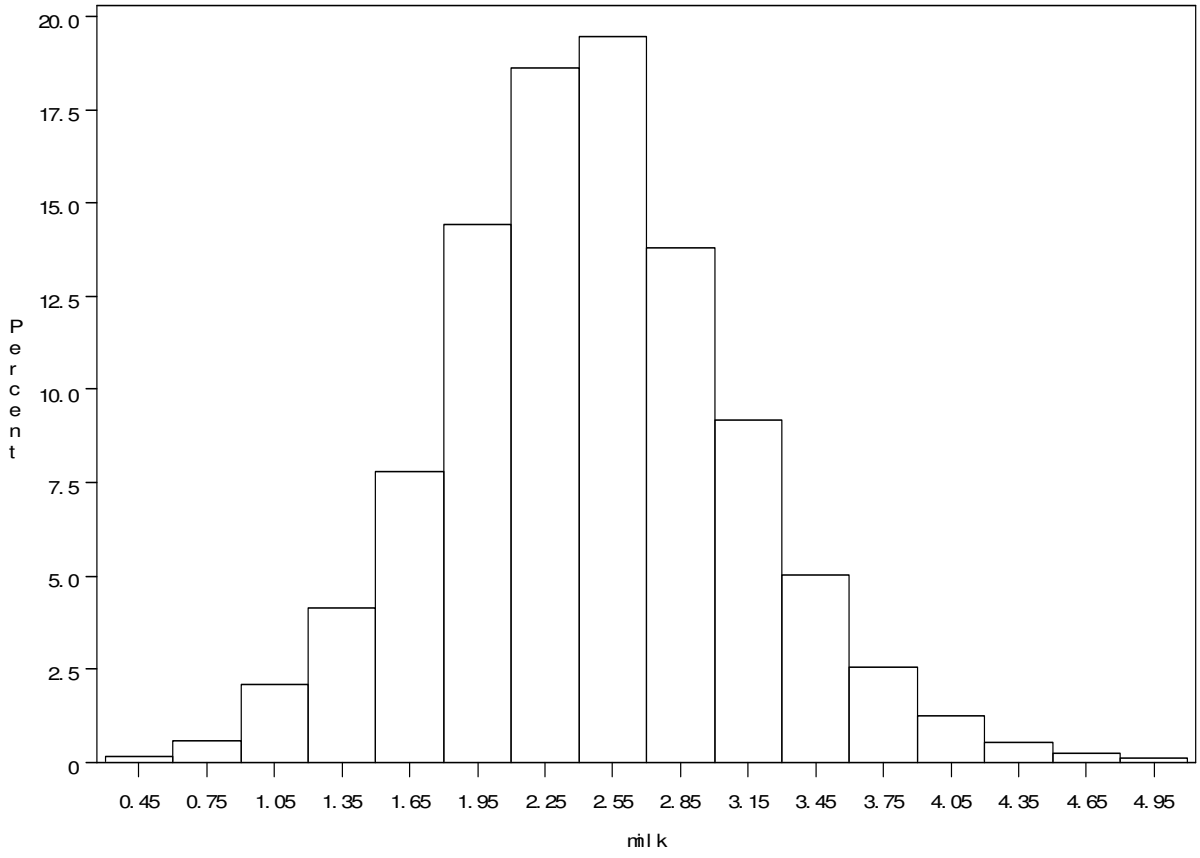

Figure 38. Histogram showing the distribution of test day milk yield records.

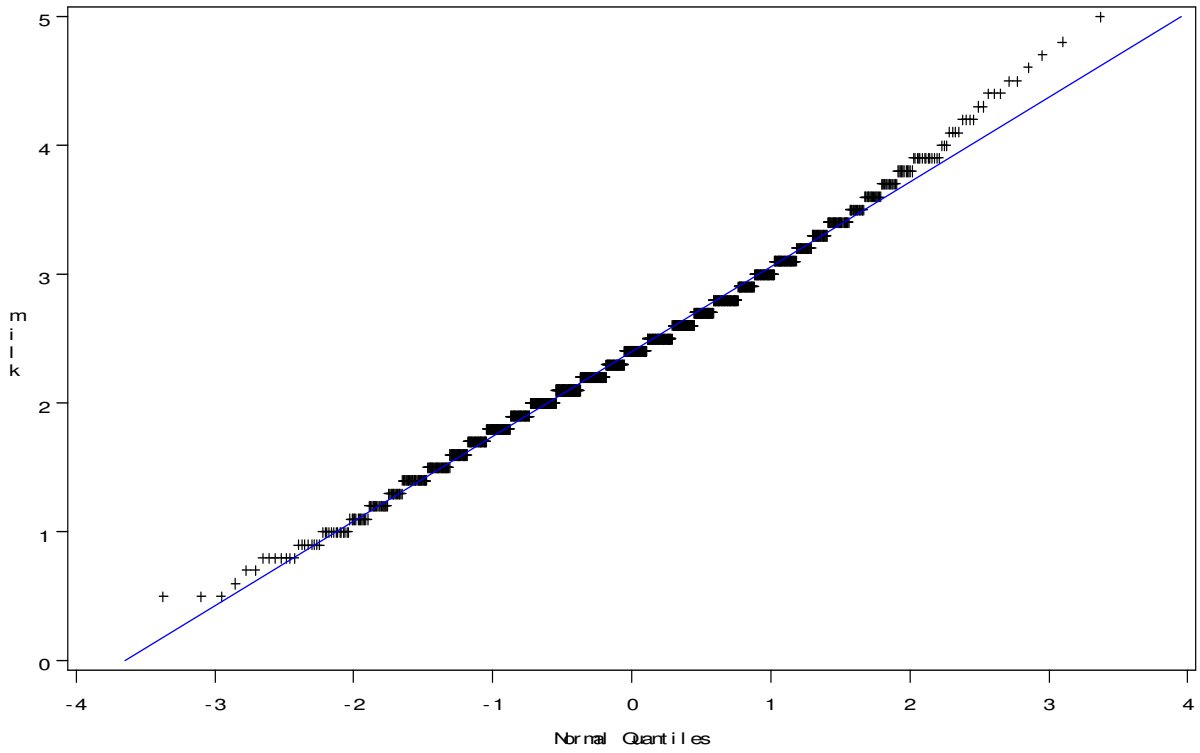

Figure 39. The Q_Q plot for test day milk yield.

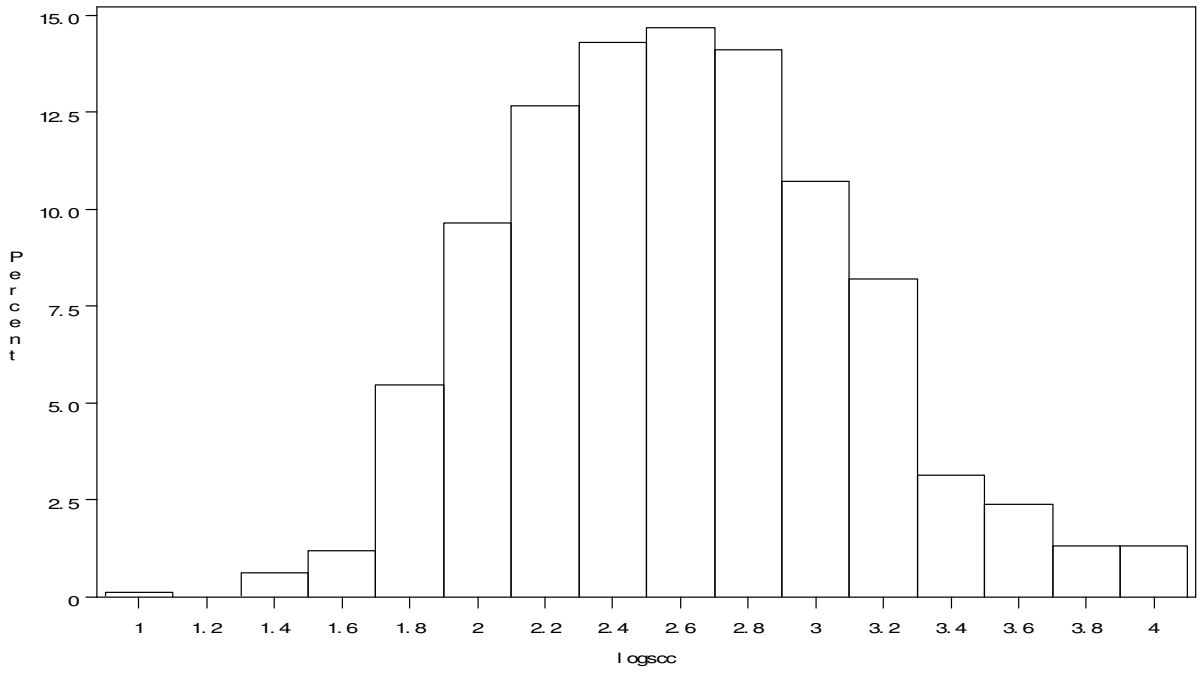

Figure 40. Histogram showing the distribution of somatic cell count (log-transformed).

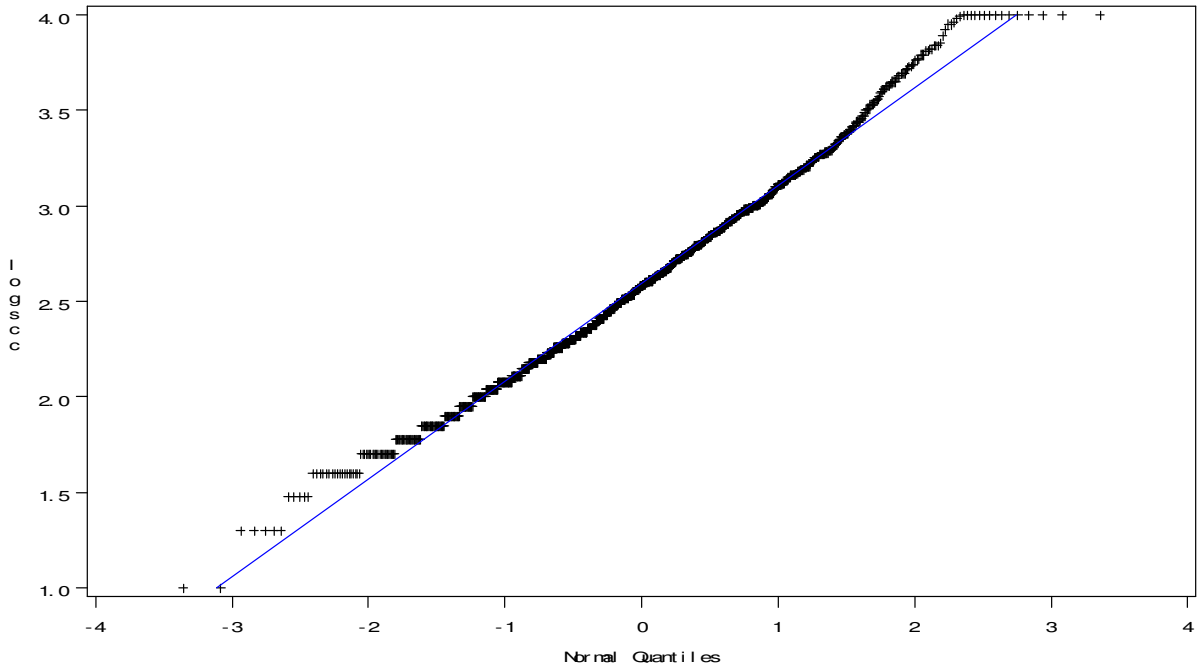

Figure 41. The Q_Q plot for somatic cell count (log-transformed).

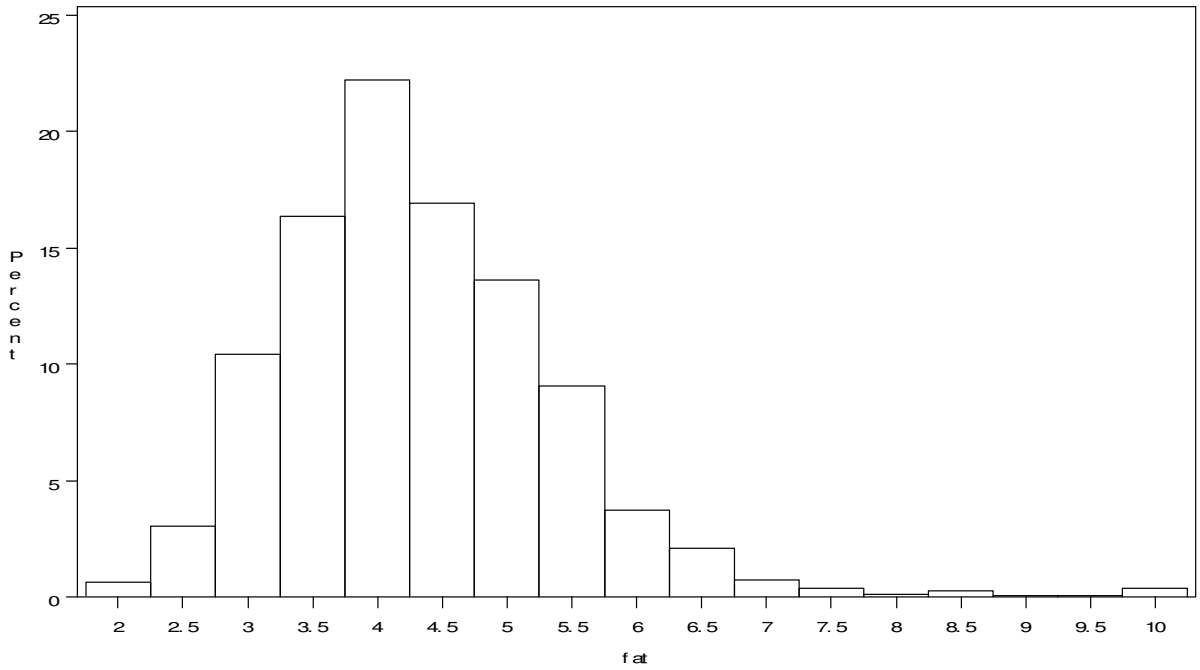

Figure 42. Histogram showing the distribution of fat percentage of test day milk.

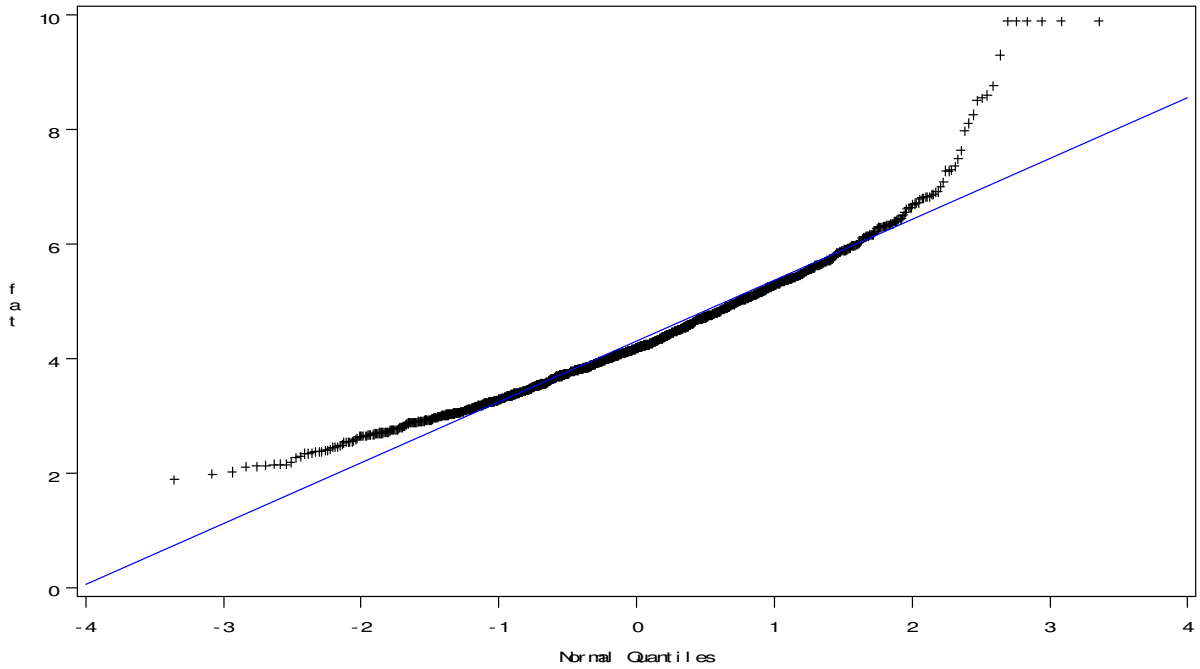

Figure 43. The Q_Q plot for test day milk fat percentage.

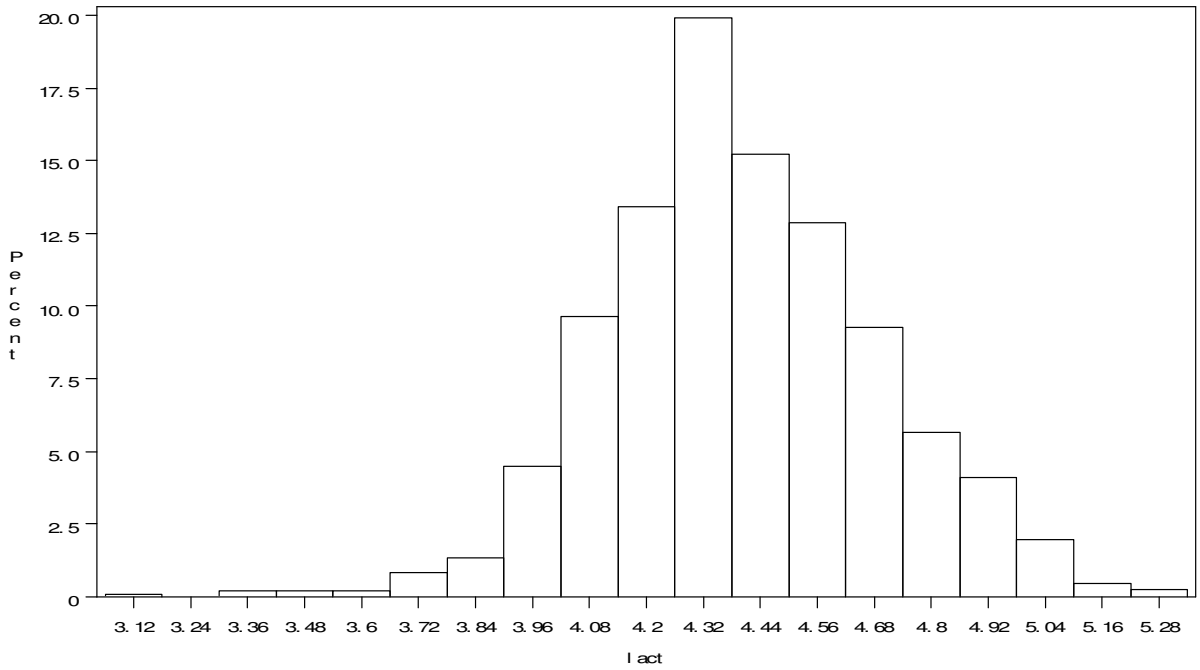

Figure 44. Histogram showing the distribution of lactose percentage of test day milk yield.

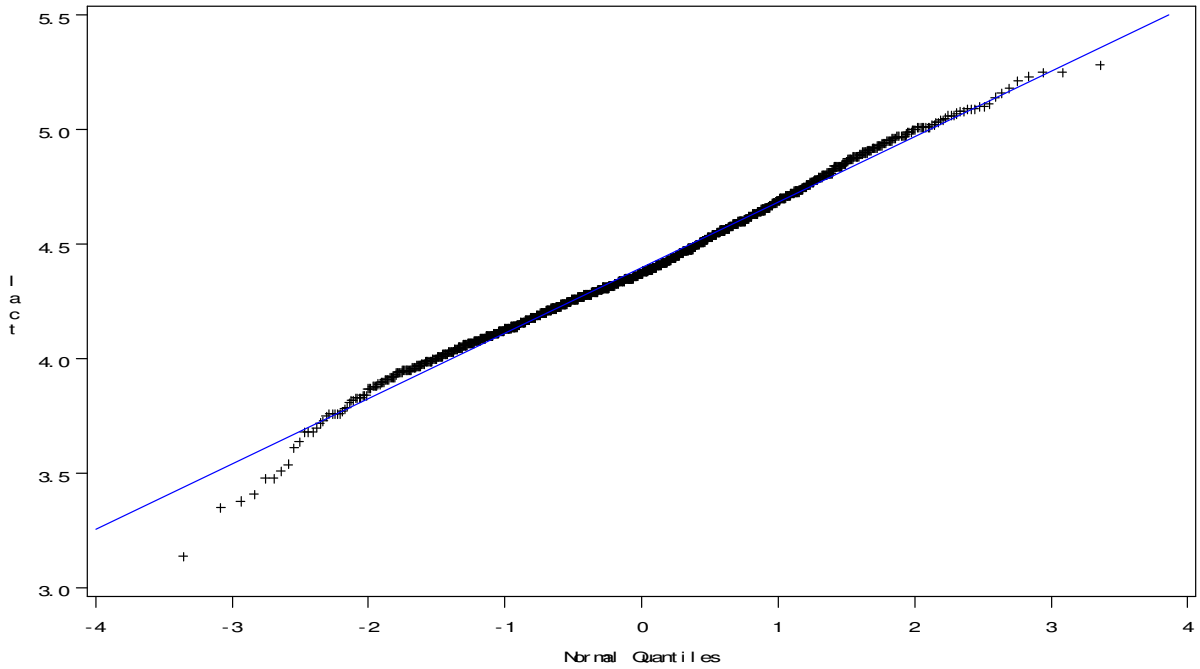

Figure 45. The Q_Q plot for test day milk lactose percentage.

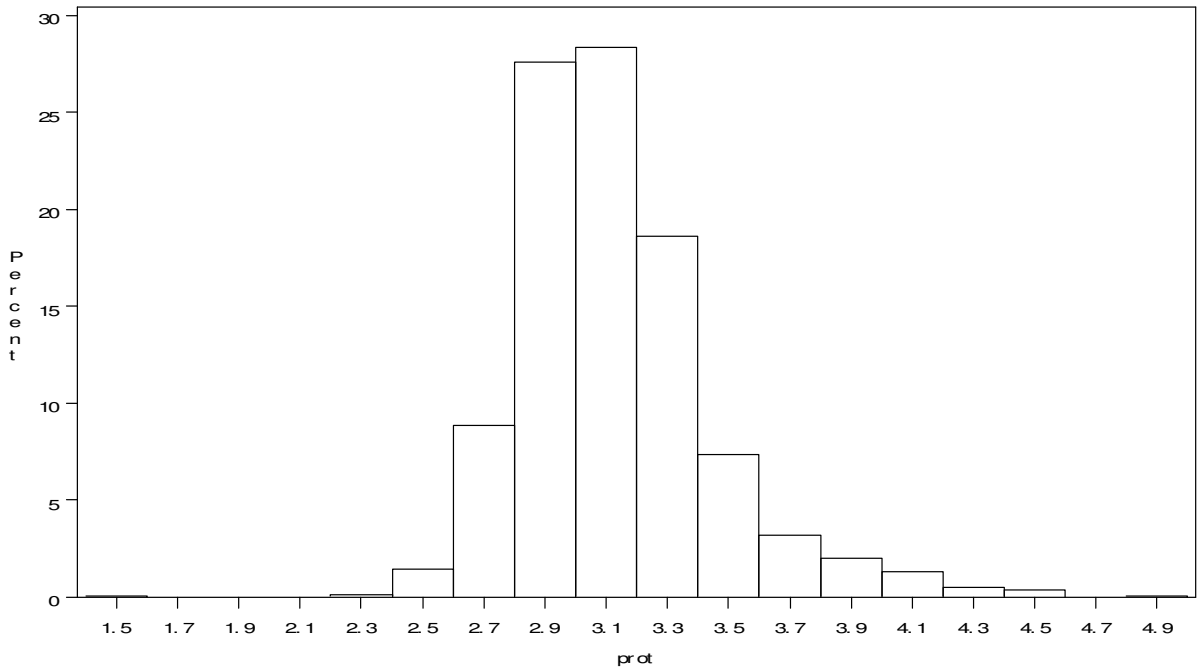

Figure 46. Histogram showing the distribution of protein percentage of the test day milk.

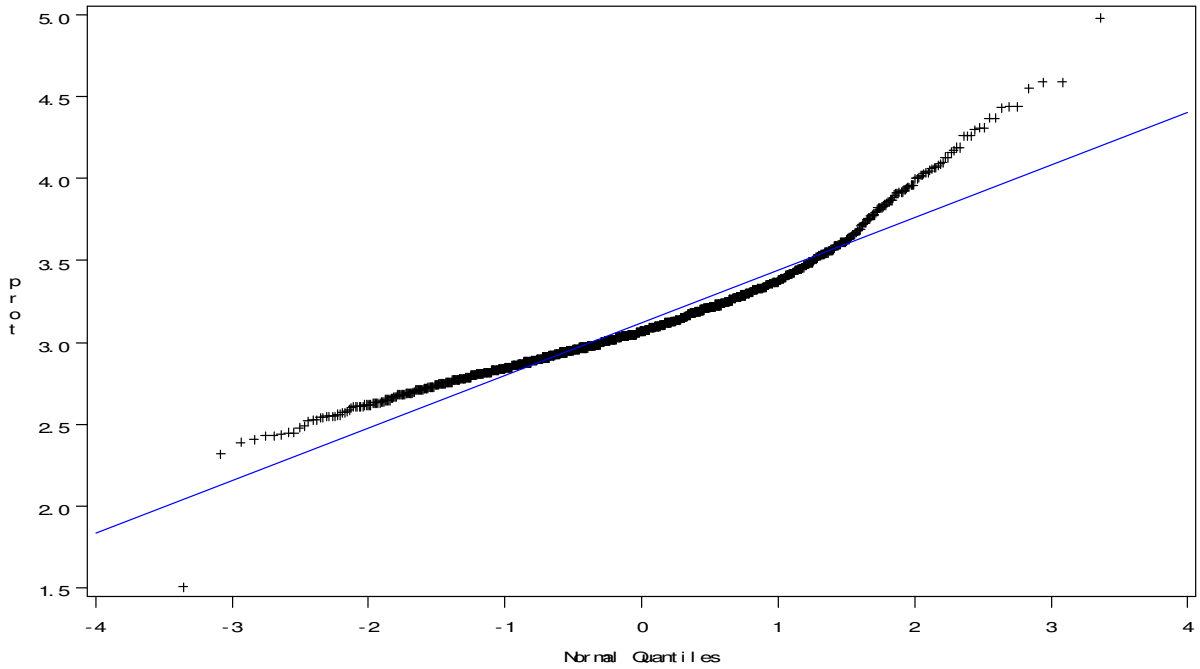

Figure 47. The Q_Q plot for test day milk protein percentage.

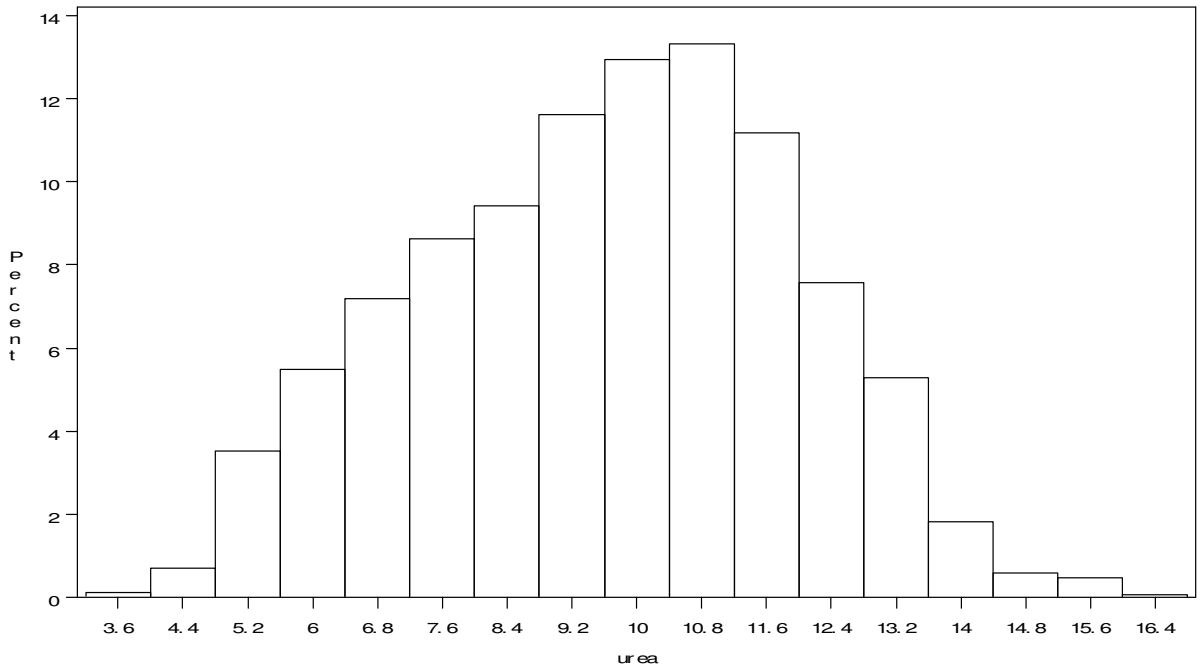

Figure 48. Histogram showing the distribution of urea content of the test day milk.

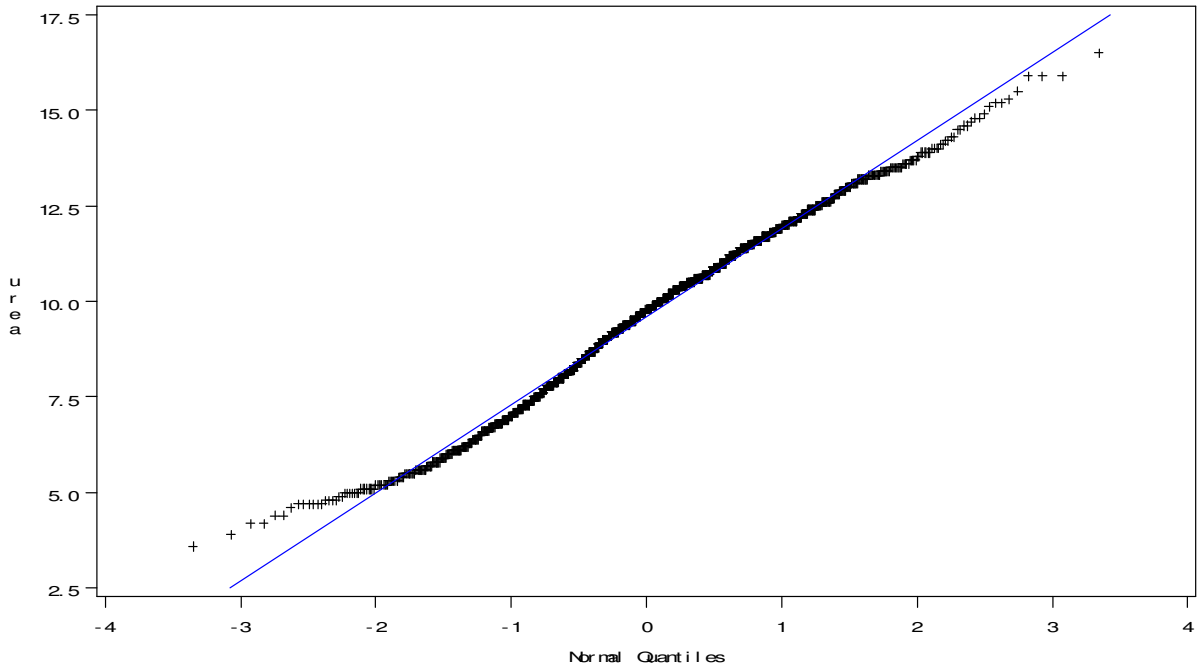

Figure 49. The Q_Q plot for test day milk urea content.

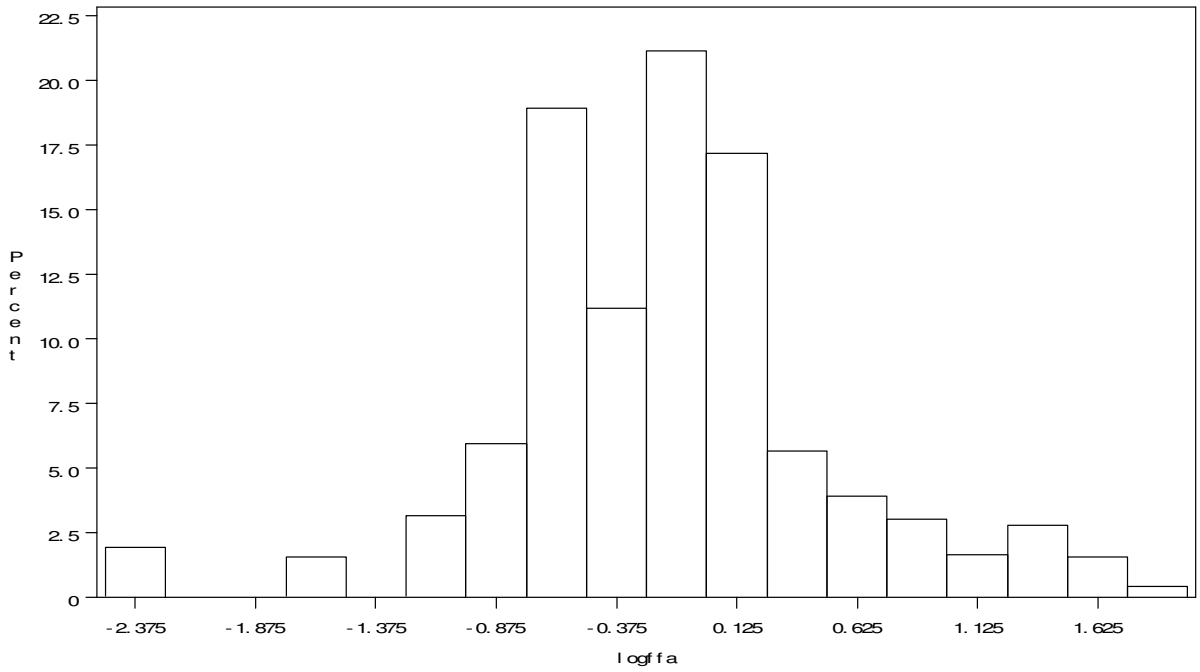

Figure 50. Histogram showing the distribution of free fatty acids content of test day milk.

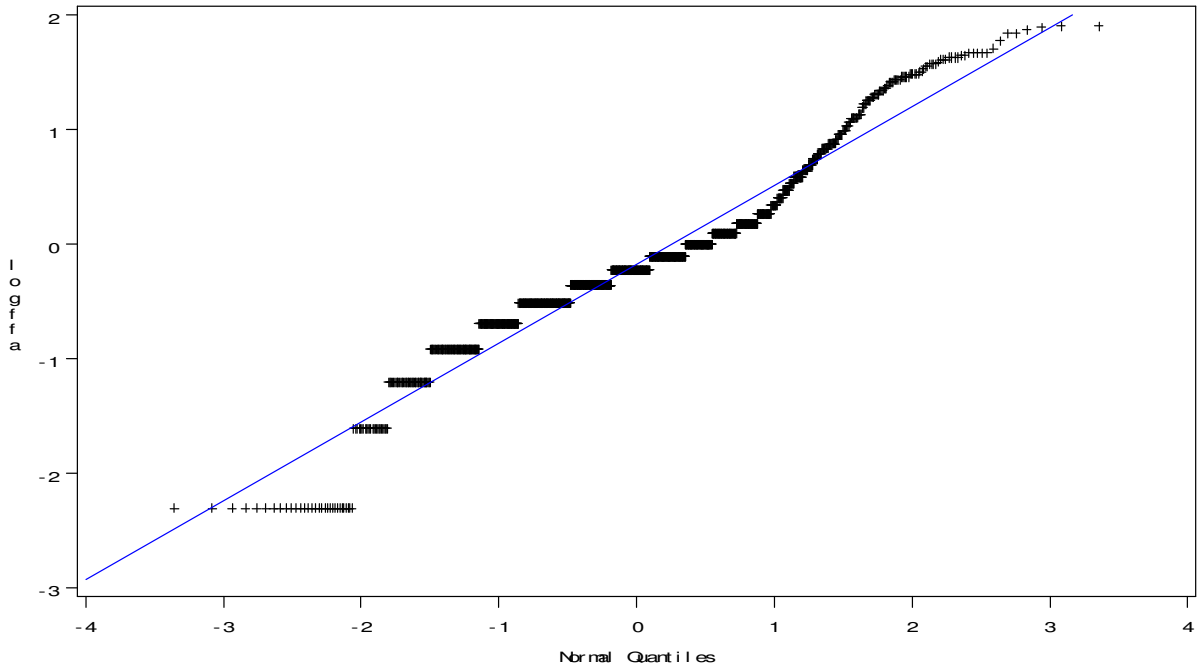

Figure 51. The Q_Q plot for test day milk free fatty acids content.

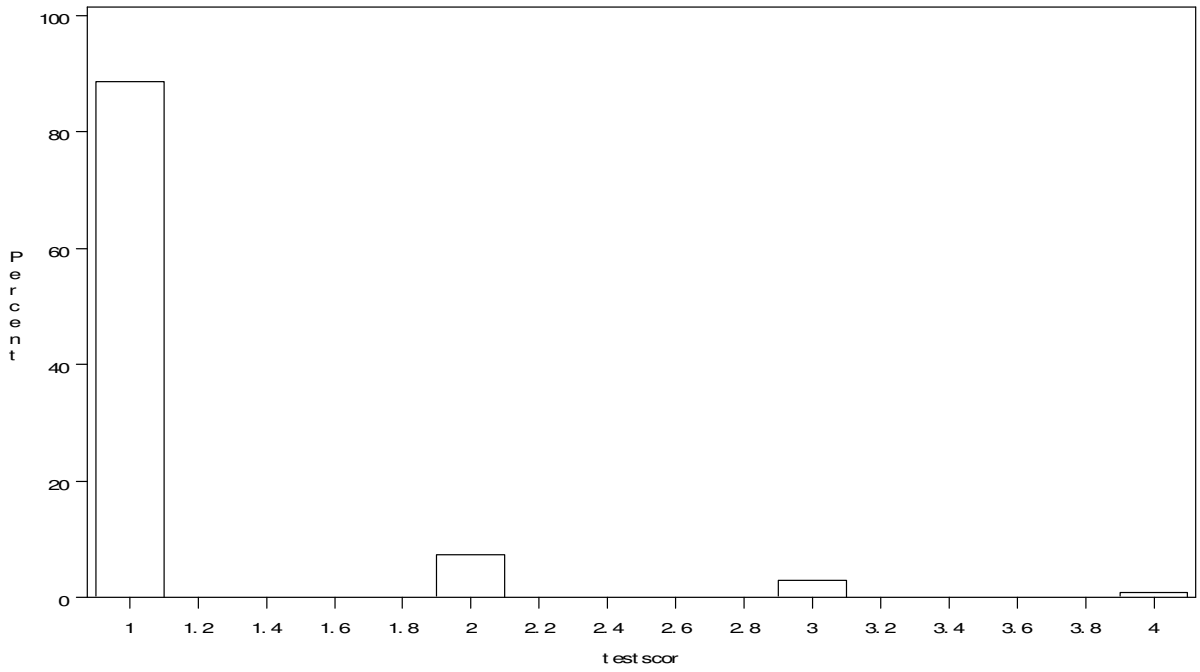

Figure 52. Histogram showing the distribution of taste score of the test day milk.

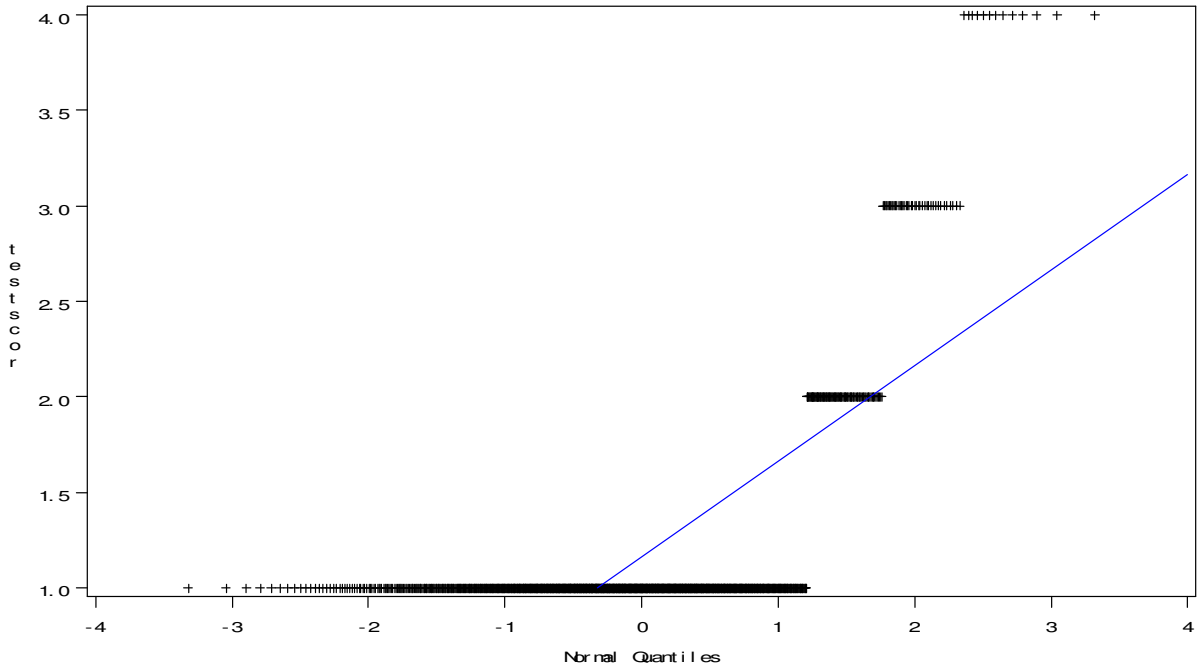

Figure 53. The Q_Q plot for test day milk taste score.

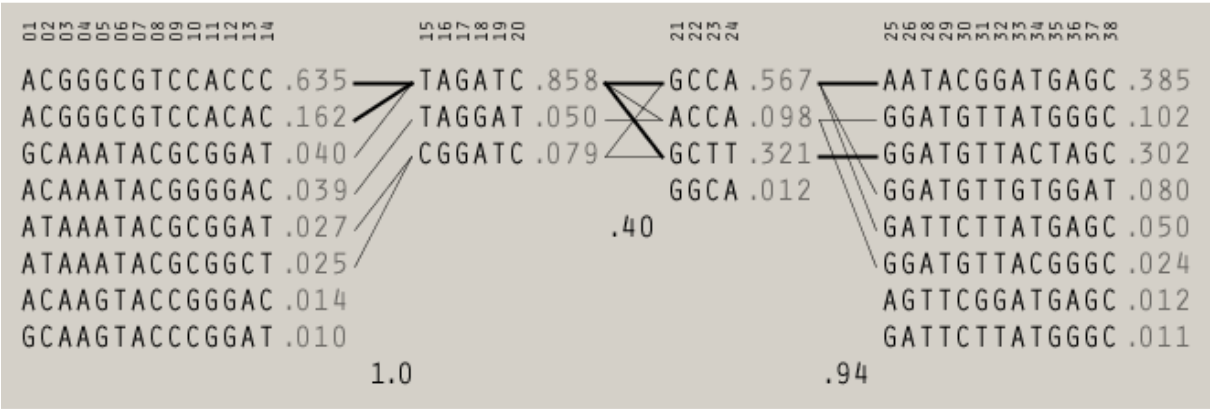

Figure 54. Individual casein genes haplotypes display for the Norwegian goats.

Casein loci SNP haplotype combinations; *CSN1S1* (marker 1 to 14), *CSN2* (marker 15 to 20), *CSN1S2* (marker 21 to 24) and *CSN3* (marker 25 to 38), and the grey numbers indicate the frequency of each haplotype.

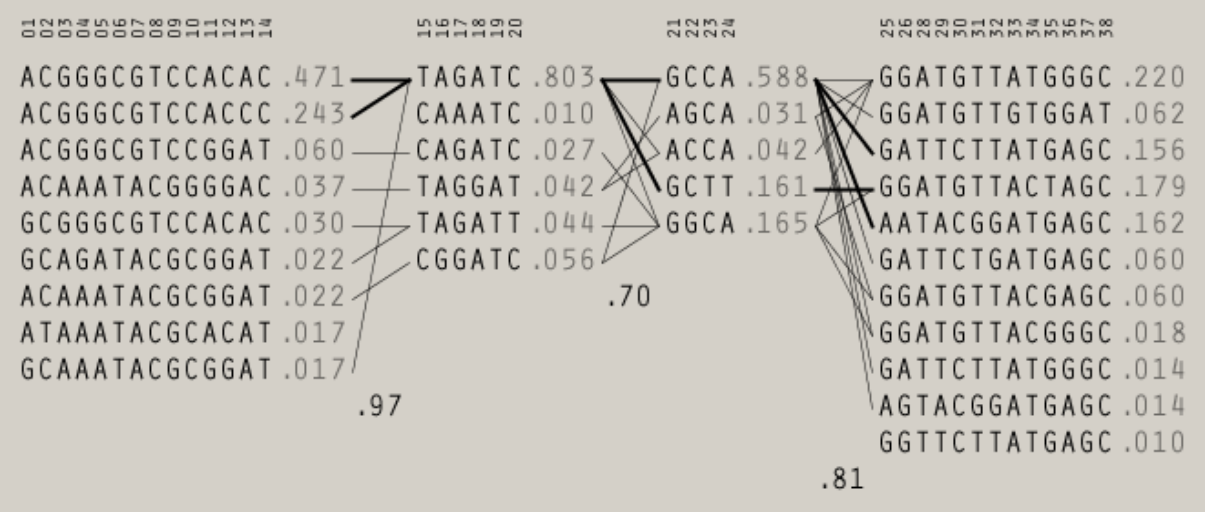

Figure 55. Individual casein genes haplotypes display for the crosses.

Table 19. Marker information for the Norwegian goats.

SNP	Name	ObsHET	PredHET	HWpval	MAF	Alleles
1	CSN1S1prom_264	0.121	0.113	0.764	0.06	A:G
2	CSN1S1prom_866	0.116	0.117	1	0.062	C:T
4	CSN1S1prom_888	0.244	0.286	0.0365	0.173	G:A
5	CSN1S1prom_1105	0.255	0.291	0.0883	0.177	G:A
6	CSN1S1prom_1169	0.176	0.214	0.033	0.122	G:A
7	CSN1S1prom_1379	0.247	0.286	0.0606	0.173	C:T
8	CSN1S1prom_1470	0.28	0.309	0.1944	0.191	G:A
9	CSN1S14Ex4_6075	0.263	0.311	0.0308	0.192	T :C
10	CSN1S14Ex4_6091	0.147	0.206	3.00E–04	0.117	C:G
11	CSN1S14Ex9_9889	0.107	0.108	1	0.057	C:D
12	CSN1S1In8_9918	0.268	0.316	0.0423	0.196	A:G
13	CSN1S1Ex10_10673	0.262	0.279	0.5695	0.168	C :G
14	CSN1S1Ex12	0.244	0.448	3.45E–11	0.339	1:3 6
15	CSN1S1Ex17_16860	0.216	0.228	0.528	0.131	C:T
16	CSN2Ex7_11801	0.136	0.146	0.4091	0.079	T:C
17	CSN2prom_2071	0.104	0.12	0.1349	0.064	A:G
18	CSN2prom_1653	0.005	0.005	1	0.003	G:A
19	CSN2prom_1009	0.113	0.114	1	0.061	A:G
20	CSN2prom_862	0.103	0.105	1	0.056	T:A
21	CSN2prom_760	0.091	0.094	0.9305	0.05	C:T
22	CSN1S2Ex3_510	0.175	0.179	0.9057	0.1	G:A
23	CSN1S2In15_273	0.025	0.024	1	0.012	C:G
24	CSN1S2In15_682	0.435	0.432	1	0.315	C:T
25	CSN1S2In15_987	0.439	0.435	1	0.32	A:T
26	CSN3prom_677	0.5	0.483	0.6829	0.408	G:A
27	CSN3prom_833	0.455	0.495	0.3262	0.448	G:A
28	CSN3prom_852	0.016	0.016	1	0.008	G:A
29	CSN3prom_942	0.54	0.5	0.245	0.485	A:T
30	CSN3prom_1074	0.451	0.476	0.475	0.391	T:A
31	CSN3prom_1140	0.541	0.5	0.2447	0.486	G:C
32	CSN3prom_1191	0.494	0.485	0.8963	0.414	T:G
33	CSN3prom_1338	0.41	0.483	0.0634	0.408	T:G
34	CSN3prom_1499	0.105	0.13	0.0351	0.07	A:G
35	CSN3prom_1550	0.369	0.436	0.0225	0.321	T:C
36	CSN3prom_1935	0.44	0.426	0.8096	0.308	G:T
37	CSN3prom_2134	0.284	0.332	0.0387	0.21	A:G
38	CSN3prom_2136	0.134	0.151	0.1622	0.082	G:A
39	CSN3Ex4_146	0.136	0.153	0.168	0.083	C:T

Table 20. Marker information for the crosses.

SNP	Name	ObsHET	PredHET	HWpval	MAF	Alleles
1	CSN1S1prom_264	0.183	0.189	0.7877	0.106	A:G
2	CSN1S1prom_866	0.07	0.093	0.0048	0.049	C:T
4	CSN1S1prom_888	0.3	0.269	0.078	0.16	G:A
5	CSN1S1prom_1105	0.266	0.247	0.3211	0.144	G:A
6	CSN1S1prom_1169	0.221	0.204	0.4015	0.115	G:A
7	CSN1S1prom_1379	0.306	0.274	0.0631	0.164	C:T
8	CSN1S1prom_1470	0.314	0.279	0.0409	0.167	G:A
9	CSN1S14Ex4_6075	0.325	0.296	0.1411	0.18	T:C
10	CSN1S14Ex4_6091	0.195	0.189	0.8724	0.105	C:G
11	CSN1S14Ex9_9889	0.1	0.095	0.9327	0.05	C:D
12	CSN1S1In8_9918	0.356	0.341	0.6833	0.218	A:G
13	CSN1S1Ex10_10673	0.369	0.315	0.0314	0.196	C:G
14	CSN1S1Ex12	0.418	0.369	0.0481	0.244	1:3:6
15	CSN1S1Ex17_16860	0.187	0.304	2.24E–09	0.187	C:T
16	CSN2Ex7_11801	0.198	0.178	0.0921	0.099	T:C
17	CSN2prom_2071	0.109	0.103	0.8172	0.055	A:G
18	CSN2prom_1653	0.036	0.035	1	0.018	G:A
19	CSN2prom_1009	0.107	0.107	1	0.057	A:G
20	CSN2prom_862	0.11	0.111	1	0.059	T:A
21	CSN2prom_760	0.137	0.169	0.0106	0.093	C:T
22	CSN1S2Ex3_510	0.15	0.139	0.3491	0.075	G:A
23	CSN1S2In15_273	0.321	0.317	1	0.198	C:G
24	CSN1S2In15_682	0.281	0.271	0.7297	0.161	C:T
25	CSN1S2In15_987	0.279	0.28	1	0.168	A:T
26	CSN3prom_677	0.376	0.305	5.27E–06	0.188	G:A
27	CSN3prom_833	0.369	0.478	0.0013	0.395	G:A
28	CSN3prom_852	0	0	1	0	G:G
29	CSN3prom_942	0.458	0.495	0.231	0.45	A:T
30	CSN3prom_1074	0.332	0.277	4.00E–04	0.166	T:A
31	CSN3prom_1140	0.463	0.494	0.317	0.446	G:C
32	CSN3prom_1191	0.379	0.307	7.38E–06	0.19	T:G
33	CSN3prom_1338	0.21	0.335	1.20E–05	0.213	T:G
34	CSN3prom_1499	0.112	0.112	1	0.06	A:G
35	CSN3prom_1550	0.237	0.377	6.29E–09	0.252	T:C
36	CSN3prom_1935	0.302	0.283	0.5242	0.171	G:T
37	CSN3prom_2134	0.282	0.436	4.81E–08	0.322	A:G
38	CSN3prom_2136	0.104	0.134	0.0038	0.072	G:A
39	CSN3Ex4_146	0.127	0.119	0.592	0.063	C:T