Over-dominance in Breeding: Case of Deletion in Exon12 of *CSN1S1* in Norwegian Goats

B. S. Dagnachew[*], G. Thaller[†], S. Lien[*‡] and T. Ådnøy[*]

Introduction

Casein proteins in goat milk are encoded by four closely linked genes within 250kb on *caprine* chromosome 6. Polymorphism of casein genes in goats have been characterized as complex and reported to influence both dairy performance and milk technological properties (Grosclaude and Martin (1997)). In the Norwegian dairy goat population, 39 DNA-polymorphic sites have been identified throughout the four casein loci (Hayes *et al.* (2006)). A deletion in exon12 of *CSN1S1,* so far reported only in Norwegian goats, has been found in high frequency (0.73, Hayes *et al.* (2006)). This deletion is believed to contribute to unusually high frequency (70%, Vegarud *et al.* (1999)) of "null" α_{S1}-CN phenotype in Norwegian goats' milk. Three DNA variants have been identified at this position: deletion (allele D: CTG**AAAAA**TAC) and non-deletion (allele G: CTG**AAGAAA**TAC or allele A: CTG**AAAAAA**TAC).

The deletion, allele D, has been found associated with reduced level of dry matter (DM) content of milk as well as influencing technological properties of the milk (Ådnøy *el al.* (2003); Hayes *et al.* (2006)). The national breeding goal is to increase dry matter content in milk (against the deletion effect), so it is difficult to explain the high frequency (i.e. 73%) of the allele in the population. The purpose of this study is to verify the additive effect of casein SNPs and examine dominance interaction of casein SNPs in Norwegian goat population.

Material and methods

Genotype data: Blood samples of 575 goats were collected and DNA was isolated from the samples following standard procedure. Genotyping of possible casein SNPs were accomplished with the Sequenom MassARRAY genotyping platform.

Phenotype data: Recordings from the Norwegian Dairy Goat Control on milk production traits in 2005 were used as phenotypes. 3194 test-days were available for analysis for daily milk yield (DMY), and 2236 samples of fat content (FC), protein content (PC) and lactose content (LC). The genetic relationship matrix (A) was calculated using 7325 pedigree records form the Dairy Control.

[*] Department of Animal and Aquacultural Sciences, UMB, Box 5003, N-1432 Ås, Norway
[†] Institute of Animal Breeding and Husbandry, CAU, D-24098 Kiel, Germany
[‡] Center for Integrative Genetics, UMB, Box 5003, N-1432 Ås, Norway

Statistical analysis: single trait test-day mixed models with fixed effect of single SNP's additive and dominance effects were fitted. Polygenic effects were also included in the model to account for effects of genes other than casein genes.

$$y = X\beta + Qq + Zu + Zpe + e$$

Where: **y** is the vector of phenotypic observations, **X** is a design matrix of fixed effects, other than SNP effect, **Q** is design matrix of SNP (additive and dominance), **β** is a vector of fixed non-genetic effects, **q** is a vector of fixed SNP effects (additive and dominant), **Z** is an incidence matrix relating individuals' phenotypes to breeding values **u** and permanent environment effect **pe** and **e** is the vector of residual error associated with each observation. **Q** was set for **additive** effect as 1 if the SNP is homozygous for one allele, 0 if the SNP is heterozygous and -1 if the SNP is homozygous for the other allele; and for dominance effect as 1 if the SNP is heterozygous and 0 otherwise.

The three alleles at exon12 of CSN1S1 (SNP14) were treated as deletion or non-deletion in the model. Dominance effects of SNP2, SNP11, SNP18, SNP19, SNP20, SNP24 and SNP29 were not estimated due to the very low number of homozygous goats for rare alleles of these SNPs. The estimation of effects was done one SNP at a time (i.e. 38 times) and Bonferroni correction was applied to account for multiple testing.

Estimated additive (a) and dominance (d) effects were collected from the model and gene substitution effects (α) were calculated ($\alpha = a + (1-2p)d$); where p is frequency of the most common allele at each SNP position.

Results and discussion

The additive effects of the most frequent alleles at each SNP position are plotted in Figure 1 and dominance effects in Figure 2. The deletion at exon12 of CSN1S1 is represented by SNP14.

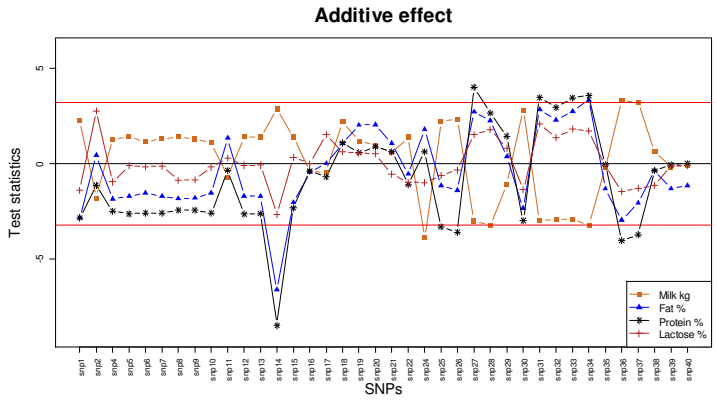

Figure 1: Casein SNPs additive effect. The horizontal lines indicate 10% experimental wise threshold level and test statistic value above the top-line or bellow the bottom line are taken as significant.

We found that the most frequent SNPs within *CSN1S1* have opposite additive effects on milk kg (per test-day) and milk composition (fat %, protein % and lactose %). SNP 14 of *CSN1S1* and a cluster of SNPs at *CSN3* had significant additive effects on milk kg, fat % and protein %. However, no significant additive effect of these SNPs on lactose % was observed (Figure 1). The deletion allele, SNP14, had a positive over-dominance effect on milk kg, negative over-dominance on lactose % and a negative dominance effect on fat % and protein % (Figure 1 and 2). Previously, it has been reported that the deletion allele has negative effect on DM content of milk (Ådnøy *et al.* (2003)) and our results also confirmed that the deletion significantly reduced the protein and fat content of the milk (Figure 1).

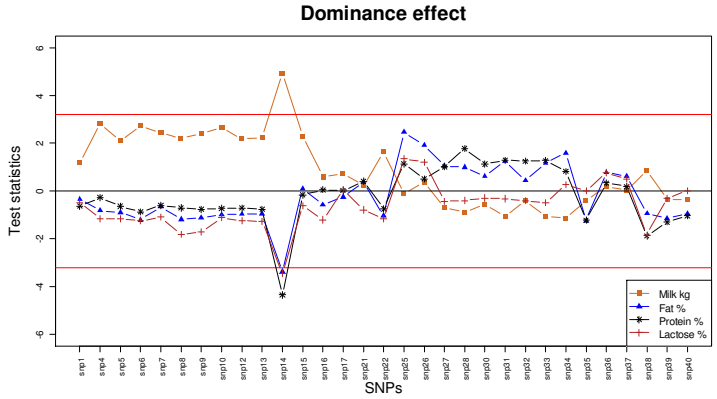

Figure 2: Casein SNPs dominance effect. The horizontal lines indicate 10% experimental wise threshold level and test statistic value above the top line or bellow the bottom line are taken as significant

In practical breeding sire evaluation is based on their daughters' performance and only utilizes the gene substitution effect variance for a polymorphic locus (Falconer and Mackay (1996)). If a gene is pure additive, gene substitution is simply equal to additive effect of the gene. With dominance the gene substitution effect is no longer equal to additive effect, but a function of dominance effect and frequency of the gene in the population (Lynch and Walsh (1998)). The deletion, allele D of SNP14, showed marked dominance effect on protein and fat percentages and exhibited over-dominance on lactose % and milk kg (Figure 1 & 2). Figure 3 shows gene substitution (α) of the deletion mutation for the fixed additive (a) and dominance (d) values. The gene substitution effect is reduced when the frequency of the allele is increasing. At current frequency of the allele in the population, 0.73, the gene substitution effect is reduced for all traits mentioned here and this might influence the selection pressure of conventional selection on the allele.

Dodds *et al.* (2007) has explained that in a population where major gene exhibits non-additive variation, and the favorable allele is found at low frequency, and selection candidate individuals can't be measured directly for the traits, traditional selection has low or no pressure on the major gene segregating in the population. In case of the deletion allele in Norwegian goat population, the observed over-dominance and dominance effects reduced the gene substitution effect (additive genetic variance available for selection) of the deletion at the current frequency of the allele in the population, 0.73. The reduced selection pressure of

standard selection on the allele might explain why the allele frequency has remained high despite selection is against the effect of the allele.

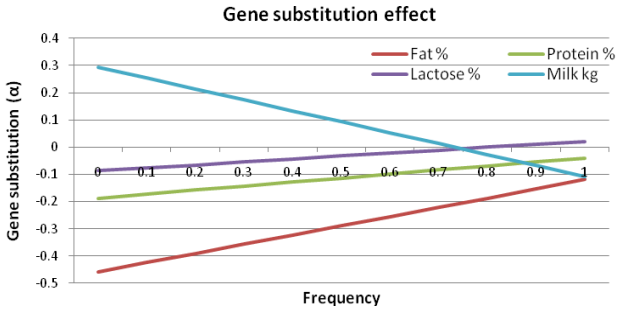

Figure 3: Gene substitution effect of the deletion allele for the fixed values of additive effect (a) and dominance (d)

Conclusion

The observed non-additive effect of the deletion allele and its high frequency in the population reduced the additive genetic variance of the allele available for selection. This limits the selection pressure of conventional breeding on the allele, and inclusion of molecular information in the national breeding scheme could help to reduce the frequency of this deletion in the population.

In the study, single SNP effects are found in separate models, modelling only one SNP at a time. This would be adequate if the SNPs were in linkage equilibrium. In other words, the model neglects the fact that neighbor SNPs are collinear and might overestimate their effects. For future we will consider multivariate analysis to overcome SNP collinear problem.

References

Ådnøy, T., Vegarud, G., Devold, T.G. *et al.* (2003). In proc. *IWMGQSG,* CD-ROM Communication No. 2-17.

Dodds, K.G., McEwan, J.C. & Davis, G.H. (2007). *Small Ruminant Res.*, 70:32-41

Falconer, D.S., & Mackay, T.F.C. (1996). *Introduction to Quantitative Genetics*, 4th ed. Longman Scientific and Technical, NY, USA.

Grosclaude, F., & Martin, P. (1997). *Int. Dairy Fed.*, No.9702:241-253

Hayes, B., Hagesæther, N., Ådnøy, T. *el al.* (2006). *Genetics*, 174:455-464

Lynch, M. & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc., Sunderland, MA, USA.

Vegarud, G.E., Devold, R., Opheim, E. *et al.* (1999). *Int. Dairy J.*, 9:367-368