# A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the Common Occurrence of Balancing Selection in Livestock

*Naveen K Kadri*[*†], Goutam Sahana[*], Carole Charlier[†], Terhi Iso-Touru[‡], Bernt Guldbrandtsen[*], Latifa Karim[†], Ulrik S Nielsen[§], Frank Panitz[*], Gert P Aamand[#], Nina Schulman[‡], Michel Georges[†], Johanna Vilkki[‡], Mogens S Lund[*], Tom
[*]Aarhus University, Tjele, Denmark, [†]University of Liège , Liège, Belgium, [‡]MTT Agrifood Research Finland, Jokioinen, Finland, [§]Danish Agricultural Advisory Service, Aarhus , Denmark, [#]Nordic Cattle Genetic Evaluation, Aarhus, Denmark

**ABSTRACT:** The spectacular increase in productivity of dairy cattle has been accompanied by the concomitant decline in fertility. It is generally assumed that this decline is primarily due to the negative energy balance of high-producing cows at the peak of lactation. We herein describe the fine-mapping of a major fertility QTL in Nordic Red cattle, and identify a 660-Kb deletion encompassing four genes as the causative variant. We show that the deletion is a recessive embryonically lethal mutation. This probably results from the loss of RNASEH2B, which is known to cause embryonic death in mice. Despite its dramatic effect on fertility, 13%, 23% and 32% of the animals carry the deletion in Danish, Swedish and Finnish Red Cattle, respectively. To explain this, we searched for favorable effects on other traits and found that the deletion has strong positive effect on milk yield. This study demonstrates that embryonic lethal mutations account for a non-negligible fraction of the decline in fertility of domestic cattle, and that associated positive effect on milk yield may account for part of the negative genetic correlation. Our study adds to the evidence that structural variants contribute to animal phenotypic variation, and that balancing selection might be more common in livestock species than previously appreciated.

**Keywords:** balancing selection, fertility, embryonic lethals

## Introduction

In the last decades, dairy cattle industry has seen an opposite genetic trend in production and fertility traits. For example the average milk yield per lactation has nearly doubled in US Holstein cows between 1960 (~6,300 kgs) and 2000 (~11,800 kgs) (Dekkers and Hospital (2002)), while in the same population the number of days between calving and first estrus has increased from 126 to 169 between 1976 and 1999 (Washburn et al. (2002)). It is generally assumed that the reduction in fertility is due to the negative energy balance of high-producing cows at the peak of lactation (e.g., Lucy (2001)).

We herein report the dissection of a fertility QTL to a 660-Kb deletion that affects fertility by causing early embryonic death of homozygous conceptuses (Kadri et al. (2014)). We also show that it is maintained at high frequency in Nordic Red breeds due to its association with positive effects on milk yield and composition. Our results thus add to the evidence that the spread of recessive embryonic lethal variants account for at least part of the decline in fertility observed in cattle. This is at least the seventh example in livestock where an allele that is deleterious at the homozygous state is maintained at high frequency in the population because of the selective advantage it confers to heterozygotes.

## Materials and Methods

**SNP Genotyping.** A total of 10,099 progeny tested bulls (4,072 Holstein-Friesian, 1,177 Jersey, 894 Danish Red, 1,714 Swedish Red, and 2,242 Finnish Ayrshire) were genotyped for 54,000 SNPs using the 50K Bovine Array (Illumina, San Diego, CA). In addition, 243 Finnish Ayrshire bulls were genotyped on the Bovine HD Genotyping BeadChip (Illumina, San Diego, CA) with 725,293 SNPs.

**Next Generation Sequencing.** Next generation Sequencing data was available for 30 Danish Red and 18 Finnish Ayrshire bulls.

**Phenotypes.** Phenotypes used for association were the bulls Estimated Breeding Values (EBV) for (i) Number of inseminations for cows/heifers (AISC & AISH), (ii) Interval between calving and first insemination (ICF), (iii) Interval between first and last insemination for cows/heifers (IFLC & IFLH), and (iv) non-return (to heat after insemination) rate at 56 days for cows/heifers (NRRC & NRRH). See www.nordicebv.info for more details.

**Phasing and Association models.** Genotypes were phased and clustered into 40 Ancestral Haplotypes with PHASEBOOK (Druet and Georges (2010)). A haplotype-based association study was then carried out using these Ancestral Haplotypes at each SNP position on BTA12. The presence of a QTL was tested using a linear mixed model accounting for familial relationships (random individual effect) and population stratification (4 principal components)

**Analysis of rate of insemination failure.** Four classes of matings were defined according to the 660-Kb deletion genotype: (i) non-carrier (NC) sire X daughter of NC maternal grand-sire, (ii) non-carrier (NC) sire X daughter of carrier (C) maternal grand-sire, (iii) C sire X daughter of NC maternal grand-sire, and (iv) C sire X daughter of C maternal grand-sire. The effect of mating class on insemination failure was then estimated using a mixed model with fixed effects of parity, month of insemination and random effect of Maternal grand-sire. The estimated effects were then compared with the theoretical expectation calculated from the weighted frequency (0.128) of the deletion in the Nordic Red population.

## Results

A genome-wide scan for fertility QTL in a five-breed (Holstein-Friesian, Jersey, Danish Red, Swedish Red and Finnish Ayrshire) combined dataset detected a highly significant QTL on BTA12. The QTL was segregating in the Nordic Red breeds (Danish Red, Swedish Red and

Finnish Ayrshire) but not in Holstein and Jersey.

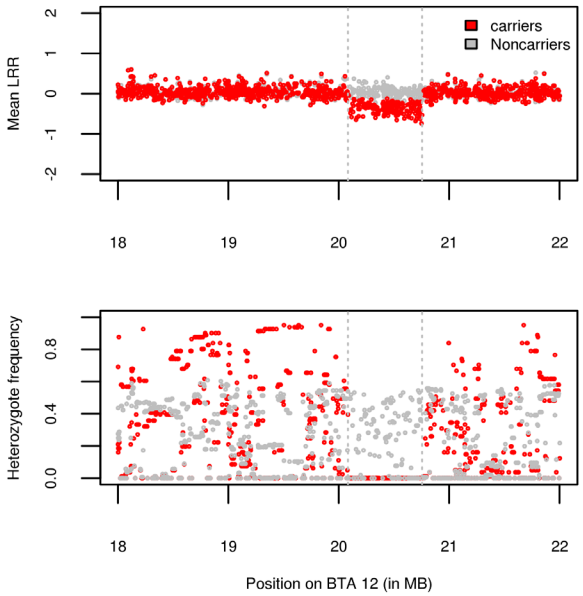


**Figure1: Average signal intensity (LRR) and mean heterozygosity per marker for carriers (red) and non carriers (gray) of Haplotype H27.**

To fine map the fertility QTL, we repeated the association analysis using a haplotype-based method in the Swedish Red and Finnish Ayrshire with the seven individual fertility traits (see M&M). Significant association signals were obtained at the expected position in both breeds for the traits relating to pregnancy success (AISC, AISH, IFLC, IFLH, NRRC and NRRH), while the traits related to estrus (HS and ICF) remained non-significant. Examination of the effects of ancestral haplotypes identified a single haplotype (referred hereafter as haplotype H27), with a strong negative effect in both the breeds. A haplotype with strong negative effect, affecting only the traits relating to insemination success (but not estrus) raised the suspicion of the variant underlying the QTL being recessive lethal. The suspicion was further supported by lack of homozygosity for the haplotype H27 (pvalues of $4.8 \times 10^{-19}$ and $2.6 \times 10^{-56}$), as expected for the recessive lethal variants. Upon examining the SNPs in the vicinity of the association peak, we identified SNPs with lack of homozygosity. To our surprise we also identified a cluster of 5 consecutive SNPs with excess homozygosity amidst the cluster. The excess homozygosity could arise in case of a deletion as the SNPs within the deletion are mis-typed. Inflation of Mendelian parent-offspring incompatibilities for these five SNPs added strength to the deletion argument.

We then took advantage of the high density (777K) SNP genotypes available for 243 Finnish Ayrshire animals (82 animals carrying H27). The presence of a deletion was evident from the reduced signal intensity (Log R ratio) and complete homozygosity for 174 consecutive SNPs (20,101,696 to 20,755,193) in carriers of H27 compared to non carriers (Figure 1). Next we searched for the deletion in the next generation sequencing data (for 30 Danish Red and 18 Finnish Ayrshire) and found evidences for deletion (halved sequence depth and incongruent mapping of paired ends separated by approximately 660-Kb) in carriers (1 Danish Red and 6 Finnish Ayrshire) of haplotype H27. A detailed analysis of individual sequences allowed us to map the breakpoints of the deletion at position 20,100,649 to 20,763,116 (662,463 bp) in the UMD3.1 assembly (Zimin et. al (2009)). The deletion was finally confirmed by two independent PCR reactions. Examination of the annotation of the orthologous region in mammals suggested that the deletion encompasses three protein-encoding genes (RNASEH2B, GUCY1B2 and 3 out of 4 exons of FAM124A), one gene with uncertain coding potential (DLEU7) and two non-coding RNA genes (DLEU7-AS1 and LINC00371).

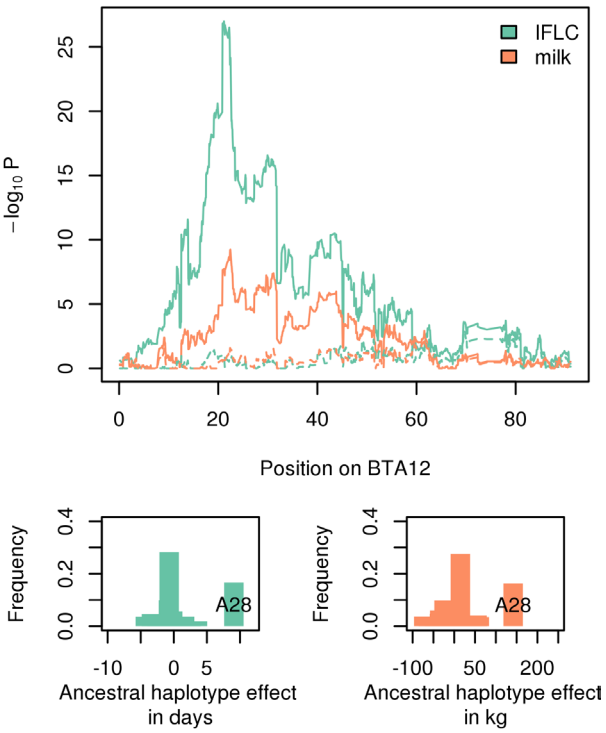


**Figure 2: (i) The haplotype based analysis for IFLC and milk in Finnish Ayrshire (top). Full and dashed lines represent QTL mapping with and without correction for haplotype A28. (ii) Effect and frequency of the 40 ancestral haplotypes on IFLC and milk (bottom).**

The 660-Kb deletion spans five SNPs in the data used for phasing. This might have compromised phasing accuracy and hence mapping accuracy. So we re-phased the data excluding the deleted SNP and repeated the association analyses. The new results complied with the old results, the peak of association signal was at the deletion and among the newly fitted ancestral haplotypes one haplotype (A28) showed a strong negative effect. Fitting A28 in the association model completely annihilated the association signal in both breeds (Figure 2, results shown only for IFLC in Finnish Ayrshire). The estimated r2 between the genotypes for A28 and deletion was 0.96, indicating that the 660-Kb deletion is most likely the causative variant

underlying the fertility QTL.

The 660-Kb deletion disrupts 4 genes, including RNASEH2B, known to be embryonic lethal when knocked out in mice (Reijns et al. (2012); White et al. (2013)). This suggests that the deletion could be embryonic lethal (EL) at homozygous state and thus affects fertility. This hypothesis implies that the insemination failure should only be restricted to matings between carriers of the deletion. To test this, we looked into the rates of insemination failure in 4 types (see M&M) of matings at 35, 56, 100 and 156 days post insemination (Figure 3). At all four time points, the increased insemination failure was only observed in mating classes (iii and iv) where both the parents have a non-zero probability of carrying the deletion proving the that the deletion is recessive lethal.

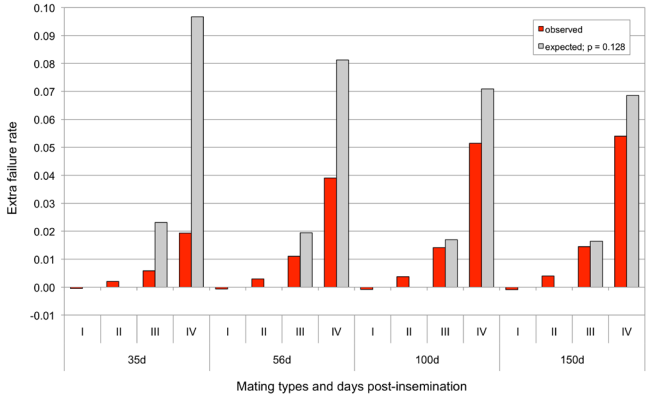


**Figure 3: Increased reproductive failure rate (expected (gray) and observed (red)) in mating types sorted according the genotype of sire and maternal grand-sire for the 660 Kb deletion.**

Despite the deleterious effect of the deletion, the frequency of carriers was intriguingly high in the Nordic Red population (32%, 23% and 13% in Finnish Ayrshire, Swedish Red and Danish Red respectively). We reasoned this to be a consequence of direct/indirect positive effect of the deletion on other traits. So we scanned BTA 12 for milk yield and composition traits in the three Nordic Red breeds and found a significant QTL, maximizing close to the 660-Kb deletion. The QTL was entirely driven by a single haplotype A28, previously found associated with the fertility. Fitting A28 in the association model completely annihilated the signal for milk and fat (Figure 2, results only shown for analysis of milk in Finnish Ayrshire). Taken together this suggests that the deletion, despite its negative effect on fertility is maintained at moderate to high frequency due to its positive effect (direct/indirect) on milk yield and composition traits.

## Discussion

We herein report dissection of a fertility QTL segregating in Nordic Red cattle to a 660-Kb deletion on BTA 12 and show that its effect on fertility is due to the death of the homozygous embryos. Our work adds to the evidence that EL are at least partly responsible for increase in insemination failures. Despite its deleterious effect, the 660-Kb deletion segregates at high frequencies in the Nordic Red cattle population. We show that this is a consequence of the positive selection of the deletion due to its associated positive effect on milk yield and composition traits. This is at least the seventh example of balancing selection in livestock. In all these cases, a single mutation is responsible for both the desirable effect in heterozygous state and the deleterious effect in the homozygous state. However in our case we cannot rule out the possibility of a distinct variant in high LD with the deletion affecting the production traits.

Further studies are required to determine which gene/genes are responsible for the effect on fertility and production. RNASEH2B (ribonuclease H2, subunit B) is a strong candidate causative gene for the embryonic lethality as knocking it out causes embryonic death in mice (Reijns et al. (2012); White et al. (2013)). RNASEH2B codes for the non-catalytic subunit of RNase H2, an endonuclease that specifically degrades the RNA of RNA:DNA hybrids and participates in DNA replication. It remains possible, however, that one or the two other coding genes included in the deletion (GUCY1B2 and FAM124A) or even DLEU7 and the two non-coding RNA genes (DLEU7-AS1; LINC00371) contribute to the embryonic lethality as well. We can also not exclude the possibility that the deletion perturbs the expression of genes lying outside of it, and that this also affects embryonic development. Genes responsible for effect on production traits also remain to be determined.

For a more complete description of this study see Kadri et al. 2014.

## Literature Cited

Dekkers, JC. and Hospital, F. (2002) .Nat Rev Genet 3: 22–32.

Druet, T. and Georges, M. (2010). Genetics 184: 789–798.

Kadri, NK., Sahana, G., Charlier., C., et al. (2014). PLoS genetics, 10(1), e1004049.

Lucy, MC. (2001). J Dairy Sci 84: 1277–1293.

Reijns, MA., Rabe, B., Rigby., et al. (2012). Cell 149: 1008–1022.

Sun, C., Madsen, P., Lund, MS., et al. (2010). J Anim Sci 88: 871–878.

Washburn, SP., Silvia, WJ., Brown, CH., et al., (2002). J Dairy Sci 85: 244–251.

White, JK., Gerdin, AK., Karp, NA., et al. (2013). Cell 154: 452–464.

Zimin, A.V., Delcher, A.L., Florea, L., et al. (2009). Genome Biology, 10:R42