J. Dairy Sci. 101:1–15
https://doi.org/10.3168/jds.2017-14067

Genome-wide association study for milking speed in French Holstein cows

Andrew Marete,*†[1] Goutam Sahana,† Sébastien Fritz,*‡ Rachel Lefebvre,* Anne Barbat,* Mogens Sandø Lund,† Bernt Guldbrandtsen,† and Didier Boichard*
*INRA, UMR 1313 GABI, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France
†Aarhus University, Center for Quantitative Genetics and Genomics, 8830 Tjele, Denmark
‡ALLICE, 75595 Paris, France

ABSTRACT

Using a combination of data from the BovineSNP50 BeadChip SNP array (Illumina, San Diego, CA) and a EuroGenomics (Amsterdam, the Netherlands) custom single nucleotide polymorphism (SNP) chip with SNP pre-selected from whole genome sequence data, we carried out an association study of milking speed in 32,491 French Holstein dairy cows. Milking speed was measured by a score given by the farmer. Phenotypes were yield deviations as obtained from the French evaluation system. They were analyzed with a linear mixed model for association studies. We identified SNP on 22 chromosomes significantly associated with milking speed. As clinical mastitis and somatic cell score have an unfavorable genetic correlation with milking speed, we tested whether the most significant SNP on these 22 chromosomes associated with milking speed were also associated with clinical mastitis or somatic cell score. Nine hundred seventy-one genome-wide significant SNP were associated with milking speed. Of these, 86 were associated with clinical mastitis and 198 with somatic cell score. The most significant association signals for milking speed were observed on chromosomes 7, 8, 10, 14, and 18. The most significant signal was located on chromosome 14 (*ZFAT* gene). Eleven novel milking speed quantitative trait loci (QTL) were observed on chromosomes 7, 10, 11, 14, 18, 25, and 26. Twelve candidate SNP for milking speed mapped directly within genes. Of these 10 were QTL lead SNP, which mapped within the genes *HMHA1*, *POLR2E*, *GNB5*, *KLHL29*, *ZFAT*, *KCNB2*, *CEACAM18*, *CCL24*, and *LHPP*. Limited pleiotropy was observed between milking speed QTL and clinical mastitis.

Key words: milking speed, mastitis, bovine, genome-wide association study, pleiotropy

INTRODUCTION

Milking speed (**MS**) measured as the time taken to milk a cow potentially affects the labor cost per cow. Different devices integrated into milking machines [e.g., LactoCorder (Hoefelmayr and Faerber WMB AG, Balgach, Switzerland, 2007), Bou-Matic (Bou-Matic, Madison, WI), or voluntary milking systems (Davis et al., 2008)] provide objective measures of MS, such as average and maximum milk flow rate, and total milking time. Alternatively, MS can be subjectively scored by the farmer. This system is quite reliable in small- to medium-size herds because of the repeated daily milking work. In France, MS is measured subjectively (Rupp and Boichard, 1999) for genetic evaluation purposes.

The relationship between MS with clinical mastitis (**CM**) and SCS is complex. A previous study in French Holstein reported positive genetic correlations between MS and CM (0.18) and between CM and SCS (0.70; Govignon-Gion et al., 2016). Similarly, a positive genetic correlation between MS and SCC was reported by several authors [0.44 (Rupp and Boichard, 1999), 0.60 (Gäde et al., 2007), and 0.46 (Samoré et al., 2010)], indicating the existence of an unfavorable relationship between MS and mastitis. Genetic correlation between MS and SCS was similar over 2 lactations, with an average correlation of 0.62 (Boettcher et al., 1998). These results support an association between MS and CM. This relationship is plausible because both traits are affected by the anatomy of the teat canal. Animals with a wider teat canal and sphincter on average show a higher milk flow, but the main disadvantage is that teat canals with a greater diameter simultaneously facilitate access to pathogens (Gäde et al., 2007; Sewalem et al., 2011). Another mechanism explaining this positive correlation is that fast milking extracts more alveolar milk rich in somatic cells (Ferneborg and Svennersten-Sjaunja, 2015). But fast milking cows have a complete flush of milk along with bacteria, which may contribute to preventing CM. Slow milking could lead to increased mastitis incidence due to incomplete

Received October 30, 2017.
Accepted March 16, 2018.
[1]Corresponding author: andrew.marete@inra.fr

milk-out, or irritated teat ends because of extended milking time (Dodd and Neave, 1951). Milking speed, therefore, should not be too fast or too slow, as both instances would probably lead to an increased incidence of CM; rather, MS should be at an intermediate optimum (Wiggans et al., 2007).

Heritability of MS is moderate: $h^2 = 0.17$ for 305-d farmer scores in French Holstein (Rupp and Boichard, 1999) and 0.14 in Canadian Holstein (Sewalem et al., 2011). It is higher for measured milking time (e.g., 0.38 in German Holstein; Gäde et al., 2007). These values are moderate and high enough for genetic selection to be effective. Nevertheless, little selection has been done on MS due to potential detrimental consequences on mastitis. Mastitis has negative financial effects on the farmer and biological effects on the cow. Financial effects include milk price penalty for high cell counts, treatment cost, loss of milk sales (due to CM and treatment), loss of production over the rest of the lactation, and increased risk of culling. Biological effects on the cow include udder injury and incomplete udder draining, which contribute negatively to cows' welfare. In the breeding objective, it would be desirable to improve MS without increasing mastitis. The QTL with effects not in line with the genetic correlation (i.e., improving MS without affecting mastitis) would be especially useful to reach this objective. In this study, we report on a genome-wide association study (**GWAS**) for MS, test the SNP associated with MS for associations with CM and SCS, and explore novel MS QTL and candidate genes associated with MS in French Holstein cattle.

MATERIALS AND METHODS

Study Population

A total of 32,491 cows with all phenotypes (MS, SCS, and CM) and genotypes were included in this study. Milking speed score (1 = slow to 5 = fast) is a subjective appraisal given by the farmer and recorded by the type classifier during classification visit. There was only one score for each cow, obtained in the first half of the first lactation. Somatic cell counts were obtained at each monthly test-day. Somatic cell score was defined in the usual way as $SCS = 3 + \log_2(SCC/100{,}000)$ and averaged over the lactation. The CM events were declared by the farmer and recorded by the technician at each test-day. The phenotype was recorded as one if the cow had at least one CM incidence in the lactation and zero otherwise. Yield deviations (**YD**, VanRaden and Wiggans, 1991), that is, performances adjusted for nongenetic effect, were obtained from the French national evaluation system (Boichard et al., 2012b). As SCS and CM might have repeated records, YD was an average over the lactations. Finally, each cow received one phenotype for each trait. Technically, the phenotypes corresponding to MS, SCS, and CM are trait deviations, and not YD (as would be for milk production traits); however, to conform with the original terminology by VanRaden and Wiggans (1991), we use the term YD.

Cows were genotyped with different types of SNP chip, so imputation process was required to recover complete genotype information. For imputation, the whole genotyped French Holstein population (males (young and old) and females (with and without phenotypes) was used as a reference. Holstein males and females were genotyped with either Illumina BovineSNP50 Beadchip (50k, Illumina, San Diego, CA), or the customized EuroGenomics SNP chip (LD-chip, Amsterdam, the Netherlands). Two successive imputation steps were carried out. In the first step, 43,800 markers from the 50k were imputed, using all 50k genotypes as a reference. This step was carried out as part of the national evaluation procedure, before this study. In the second step carried out for the present study, additional functional variants were imputed, using animals genotyped for the functional variants included in the LD-chip as a reference. The LD-chip is composed of 2 parts: (1) ~8,000 generic (and supposedly neutral) SNP mainly from BovineLD Genotyping BeadChip (Boichard et al., 2012a) and the 50k chip and (2) a custom part selected from whole-genome sequence variants based on different functional arguments: (1) known genetic variants described in literature, (2) potential regulatory variants located in the promoter regions of genes, (3) nonsynonymous variants with strongly deleterious effect on the function of the encoded protein as predicted by the variant effect predictor (McLaren et al., 2010), (4) breakpoints of structural variants affecting genes as described in Boussaha et al. (2015), and (5) variants corresponding to peaks in GWAS analysis of several economic traits in cattle. All SNP with minor allele frequency (**MAF**) lower than 0.5%, with a call rate lower than 95%, or deviating from Hardy-Weinberg proportions ($P < 10^{-4}$) were deleted. Four versions of this LD-chip were used, with partial overlap between custom parts. A total of 6,035 variants in LD-chip passed quality checks. Imputation was carried out using FImpute (Sargolzaei et al., 2014), for 187,025 genotyped Holstein animals (males and females, with or without performance records). Imputation errors may affect the 50k SNP not present on the LD-chip, and the candidate variants not present on the 50k. However, this consequence was likely small, as the loss in imputation error rate as expressed by allelic correlation was <0.01. After imputation, 49,835 SNP distributed over 29 BTA remained.