J. Dairy Sci. 100:1–13
https://doi.org/10.3168/jds.2016-11770

Genome-wide association study for lactation persistency, female fertility, longevity, and lifetime profit index traits in Holstein dairy cattle

S. Nayeri,* M. Sargolzaei,†‡ M. K. Abo-Ismail,*§ S. Miller,†# F. Schenkel,† S. S. Moore,‖ and P. Stothard*[1]
*Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada, T6G 2P5
†Centre for the Genetic Improvement of Livestock, University of Guelph, Ontario, Canada, N1G 2W1
‡The Semex Alliance, Guelph, Ontario, Canada, N1H 6J2
§Department of Animal and Poultry Production, Damanhour University, Damanhour, Egypt
#Angus Genetics Inc., American Angus Association, St. Joseph, MO 64506
‖Agriculture and Food Innovation, Center for Animal Science, University of Queensland, St Lucia QLD 4072, Australia

ABSTRACT

Female fertility in Holstein cattle can decline when intense genetic selection is placed on milk production. One approach to improving fertility is to identify the genomic regions and variants affecting fertility traits and then incorporate this knowledge into selection decisions. The objectives of this study were to identify or refine the positions of the genomic regions associated with lactation persistency, female fertility traits (age at first service, cow first service to conception, heifer and cow nonreturn rates), longevity traits (herd life, indirect herd life, and direct herd life), and lifetime profit index in the North American Holstein dairy cattle population. A genome-wide association study was performed for each trait, using a single SNP (single nucleotide polymorphism) regression mixed linear model and imputed high-density panel (777k) genotypes. No associations were identified for fertility traits. Several peak regions were detected for lifetime profit index, lactation persistency, and longevity. The results overlap with previous findings and identify some novel regions for lactation persistency. Previously proposed causative and candidate genes supported by this work include *DGAT1*, *GRINA*, and *CPSF1*, whereas new candidate genes are *SLC2A4RG* and *THRB*. Thus, the chromosomal regions identified in this study not only confirm several previous findings but also highlight new regions that may contribute to genetic variation in lactation persistency and longevity-associated traits in dairy cattle.

Key words: longevity, genome-wide association study, lifetime profit index, fertility

INTRODUCTION

Milk production and female fertility are 2 important traits that contribute to the profitability of the dairy industry (Boichard, 1990). Increasing milk production in dairy cattle has been a primary focus of genetic selection (Oltenacu and Algers, 2005). This selection has caused a decline in cow fertility because of the negative genetic correlation between fertility and milk production (Kadarmideen et al., 2000; Royal et al., 2002). It has been concluded that a combination of physiology, nutrition, genetic, and management strategies should be considered to provide a long-term improvement in fertility of high-producing dairy cows (Shook, 2006). Dairy breeding programs stand to improve the overall profitability of the industry through greater emphasis on durability, health, and fertility of cows (Kulak et al., 1997).

Genomic regions explaining variation in female fertility traits in cattle have been identified in several genome-wide association studies (**GWAS**) within a variety of breeds (Höglund et al., 2009; Pryce et al., 2010; Schulman et al., 2011; Sahana et al., 2011; Hawken et al., 2012; Peñagaricano et al., 2012; Minozzi et al., 2013; Höglund et al., 2015). Significant associations have been identified on several chromosomes for age at puberty (Hawken et al., 2012), cow nonreturn rate (Holmberg and Andersson-Eklund, 2006), pregnancy rate (Ashwell et al., 2004), and calving performance (Holmberg and Andersson-Eklund, 2006). More in-depth analyses have identified several candidate genes associated with fertility traits such as pregnancy-associated plasma protein-A2 (*PAPP2-A2*) on chromosome 16 (associated with calving ease; Wickramasinghe et al., 2011) and calpastatin (*CAST*) on chromosome 7 (associated with fertility and longevity; Garcia et al., 2006) in dairy cattle (Minozzi et al., 2013).

The objective of this study was to detect genomic regions associated with several female fertility, longevity, and productivity traits in the North American Holstein

Received July 22, 2016.
Accepted October 13, 2016.
[1]Corresponding author: stothard@ualberta.ca

population. Lifetime profit index (**LPI**), which reflects the relative profitability that can be expected during the lifetime of future daughters, was included in the analysis.

MATERIALS AND METHODS

Animals and Data

Animal Care and Use Committee approval was not obtained for this study because analyses were performed on existing data obtained under standard farm management from commercial dairy farmers and breeders. All dairy farmers in Canada must follow "The Code of Practice for the Care and Handling of Dairy Cattle" developed by the National Farm Animal Care Council of Canada (http://www.nfacc.ca/).

A population of Holstein bulls registered and used in North America was used in this study for 9 fertility- and profitability-related traits, including LPI, lactation persistency (**LP**), herd life (**HL**), indirect herd life (**IHL**), direct herd life (**DHL**), cow first service to conception (**FSTCc**), age at first service (**AFS**), heifer 56-d nonreturn rate (**NRRh**), and cow 56-d nonreturn rate (**NRRc**). The LPI consists of 3 main components: production (yield traits and milk components), durability (herd life, mammary system, feet and leg, dairy strength), and health and fertility (daughter fertility). The emphases given to components are 51% for production, 34% for durability, and 15% for health and fertility (trait definitions according to www.cdn.ca).

Lactation persistency is the average of expected milk yield of a bull's daughter at d 280 in lactation compared with that at d 60 in lactation. Herd life is directly measured as the survival (direct daughter survival) of each cow at 5 time points during their productive life; IHL is evaluated based on a combination of conformation traits, reproduction traits, and udder health. Cow first service to conception is the interval in days between the first service and conception. Age at first service is the age in days at which a heifer was inseminated for the first time. The NRRh and NRRc are the fraction of heifers and cows for which an insemination in a period of 56 d leads to conception.

The Canadian Dairy Network (**CDN**) provided genotypes, available pedigree information, and official evaluations for proven bulls born between 1956 and 2009. Individuals were genotyped using either the BovineSNP50K (50k) panel (44,369 SNP, 3,729 bulls) or the high-density (**HD**, 777k) SNP panel (774,605 SNP, 2,387 bulls; Illumina Inc., San Diego, CA). The 50k panel SNP genotypes were subjected to the standard quality control measures that are used by CDN (Wiggans et al., 2009). Quality control was performed on the HD genotyping data using the snp1101 (Sargolzaei, 2014) software. This step excluded 116,619 SNP, including 46,433 SNP from a sex chromosome or misplaced SNP, 3,566 SNP with high Mendelian error rate (>0.05), 6,446 SNP with low call rate (<0.9), 61,577 SNP with low minor allele frequency (<0.000001), and 90 SNP with excess heterozygosity (>0.15; removing SNP that have a higher heterozygosity compared with expected heterozygosity from Hardy-Weinberg equilibrium, as explained by Wiggans et al., 2009). The purpose of choosing this threshold (>0.15) was to exclude only the most outlying SNP in the data, eliminating possible genotyping errors (Wiggans et al., 2009). The number of SNP remaining for downstream imputation was 40,666 SNP for the 50k panel and 657,986 SNP for the HD panel.

The 3,729 50k genotypes were imputed to the HD panel, using the 2,387 HD panel genotypes as the reference and the FImpute V2.2 software (Sargolzaei et al., 2014). After imputation, an additional quality control step was performed on the imputed data. A total of 55,817 SNP with minor allele frequency <1% and 74 SNP with a Mendelian error rate >5% (74 SNP) were excluded. After quality control, 602,095 SNP remained for use in the subsequent association analysis.

Deregressed Proofs Calculation

In this study, deregressed genetic evaluations of Holstein bulls were used as independent variables to test the association with the HD panel. In this genetic evaluation, a bull's published EBV is a weighted mean of his daughters' deviations (**DD**) and his parental average (**PA**; VanRaden et al., 2009). The deregressed bull proofs were computed by CDN as shown below (VanRaden et al., 2009):

$$DE_{prg} = \frac{Rel_{EBV}}{1 - Rel_{EBV}} - \frac{Rel_{PA}}{1 - Rel_{PA}},$$

$$Rel_{DD} = \frac{DE_{prg}}{DE_{prg} + 1},$$

$$DEBV = PA + \frac{(EBV - PA)}{Rel_{DD}},$$

where DE_{prg} is the daughter equivalent from progeny information, and Rel_{EBV} and Rel_{PA} are the reliabilities of EBV and PA, respectively; Rel_{DD} is the reliability of DD, and $DEBV$ is the deregressed bull proof.