J. Dairy Sci. 100:1–15
https://doi.org/10.3168/jds.2016-12249

Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle

N. P. P. Macciotta,*[1] S. Biffani,† U. Bernabucci,‡ N. Lacetera,‡ A. Vitali,‡ P. Ajmone-Marsan,§ and A. Nardone‡[1]
*Dipartimento di Agraria, Università di Sassari, 07100 Sassari, Italy
†Associazione Italiana Allevatori, 00161 Roma, Italy
‡Dipartimento di Scienze Agrarie e Forestali, Università degli Studi della Tuscia-Viterbo, 01100 Viterbo, Italy
§Istituto di Zootecnica, Università Cattolica del Sacro Cuore, 29122 Piacenza, Italy

ABSTRACT

Heat stress represents a key factor that negatively affects the productive and reproductive performance of farm animals. In the present work, a new measure of tolerance to heat stress for dairy cattle was developed using principal component analysis. Data were from 590,174 test-day records for milk yield, fat and protein percentages, and somatic cell score of 39,261 Italian Holstein cows. Test-day records adjusted for main systematic factors were grouped into 11 temperature-humidity index (THI) classes. Daughter trait deviations (DTD) were calculated for 1,540 bulls as means of the adjusted test-day records for each THI class. Principal component analysis was performed on the DTD for each bull. The first 2 principal components (PC) explained 42 to 51% of the total variance of the system across the 4 traits. The first PC, a measure of the level at which the curve is located, was interpreted as a measure of the level at which the DTD curve was located. The second PC, which shows the slope of increasing or decreases DTD curves, synthesized the behavior of the DTD pattern. Heritability of the 2 component scores was moderate to high for level across all traits (range = 0.23–0.82) and low to moderate for slope (range = 0.16–0.28). For each trait, phenotypic and genetic correlations between level and slope were equal to zero. A genome-wide association analysis was carried out on a subsample of 423 bulls genotyped with the Illumina 50K bovine bead chip (Illumina, San Diego, CA). Two single nucleotide polymorphisms were significantly associated with slope for milk yield, 4 with level for fat percentage, and 2 with level and slope of protein percentage, respectively. The gene discovery was carried out considering windows of 0.5 Mb surrounding the significant markers and highlighted some interesting candidate genes. Some of them have been already associated with the mechanism of heat tolerance as the heat shock transcription factor (*HSF1*) and the malonyl-CoA-acyl carrier protein transacylase (*MCAT*). The 2 PC were able to describe the overall level and the slope of response of milk production traits across increasing levels of THI index. Moreover, they exhibited genetic variability and were genetically uncorrelated. These features suggest their use as measures of thermotolerance in dairy cattle breeding schemes.

Key words: heat tolerance, principal component analysis, heritability, genome-wide association study, dairy cattle

INTRODUCTION

The improvement of an animal's ability to cope with adverse environmental conditions is one of the great challenges of animal breeding for the future (Bernabucci et al., 2010). Among the traits that contribute to define animal adaptability to environmental variation, tolerance to heat stress plays a major role. Heat stress can be defined as the condition where the animal is not able to adequately dissipate the excess of endogenous or exogenous heat to maintain body thermal balance (Bernabucci et al., 2014). In dairy cattle, it is known that heat stress results in relevant economic losses due to reduced milk production and reproduction performance (Aguilar et al., 2010; Nardone et al., 2010; Biffani et al., 2016). Increasing concern about tolerance to heat stress for dairy animals in temperate areas is a consequence of both climate change and higher metabolic heat production by high-yielding animals (Kadzere et al., 2002; Hansen, 2007; Segnalini et al., 2011).

If tolerance to heat stress is a quite straightforward concept, its systematic measure remains problematic. On the other hand, a quantification of this trait is fundamental if it is to be considered a potential selection goal in breeding programs.

Some physiological traits are related to the ability of the animal to cope with heat stress. For example, rectal

---

Received November 2, 2016.
Accepted February 5, 2017.
[1] Corresponding authors: macciott@uniss.it and nardone@unitus.it

temperature and respiration rate increase when animals are exposed to warm environment (Dikmen et al., 2012; Perano et al., 2015; Garner et al., 2016). These traits exhibit a genetic component; for example, a moderate heritability and associations with SNP and candidate genes have been reported for rectal temperature (Dikmen et al., 2013, 2015). However, the inclusion of these heat tolerance indicator traits in large-scale phenotype recording systems for selecting thermotolerant animals appears rather problematic in terms of logistics and costs. An alternative is to evaluate heat tolerance by measuring changes of milk production traits under warm environmental conditions (Hammami et al., 2015; Carabaño et al., 2016; Nguyen et al., 2016). In dairy cattle populations involved in selection programs, milk production data could be easily retrieved from dairy recording systems and associated with climate data provided by weather stations. The variable most frequently used to evaluate heat stress conditions is the temperature-humidity index (**THI**). The approach commonly used to evaluate heat tolerance relies on the so-called broken line model (Ravagnolo and Misztal, 2000). It assumes the existence of a comfort zone limited by an upper threshold value (**TH_0**), beyond which the production linearly decreases as THI increases (Bernabucci et al., 2014; Carabaño et al., 2014).

In statistical models, tolerance to heat stress might be fitted according to a reaction norm model (Kolmodin and Bijma, 2004), where the phenotype is expressed as a linear function of an environmental variable (for example THI or temperature). Very often, the environmental variable effect is a dummy variable, set to zero when $THI < TH_0$ and to $THI - TH_0$ when $THI > TH_0$ (Bernabucci et al., 2014). Some studies adopted the fixed value of 72 for TH_0 (Ravagnolo and Misztal, 2000; Bohmanova et al., 2008; Aguilar et al., 2009), but recently different TH_0 have been estimated across traits, parities and geographical regions (Bernabucci et al., 2014; Biffani et al., 2016).

Some studies on tolerance of heat stress have used individual production curves along different THI levels corrected for fixed factors as a measure of heat tolerance (Hayes et al., 2009; Carabaño et al., 2016). Average curves of bull progeny for milk production traits across different THI levels, named as daughter trait deviations (**DTD**), have been recently used as phenotypes in a genomic selection study on tolerance to heat stress (Nguyen et al., 2016).

For genetic purposes, individual effects for heat tolerance are usually fitted with an intercept and a slope, representing the overall level of production and the response of the animal to heat stress, respectively. Main concerns about these approaches are on the use of a common threshold across all animals and the assumption of linearity for the production decay after TH_0 (Bernabucci et al., 2014; Carabaño et al., 2014). On the other hand, estimation of individual thresholds (Sánchez et al., 2009) is more realistic though it is more computationally demanding. Individual change points of production patterns for increasing THI levels have been fitted also with Legendre polynomials in random regression models (Brügemann et al., 2011; Carabaño et al., 2014, 2016)

Several papers that evaluated the effect of heat stress on milk reported an unfavorable genetic relationship between production and heat tolerance (Sánchez et al., 2009; Bernabucci et al., 2014; Hammami et al., 2015). These results were confirmed also by the strong negative correlations (−0.85 and −0.75) between genomic breeding value for milk DTD-derived heat tolerance and EBV for milk yield in Australian Holsteins and Jerseys respectively (Nguyen et al., 2016). Such correlations are the result of the increased metabolic heat production that occurs in high-producing cows and that exacerbate the effects of the external heat. This represents a severe constraint to an efficient selection for improving heat tolerance without negative consequences on production. The aggregation of the 2 traits into a selection index may help selection, even though the definition of optimal economic weights could remain a theoretical issue and the negative correlation undoubtedly will reduce the selection response on each individual trait. An alternative could be the use of a measure of tolerance to heat stress that is not correlated with production levels. The use of a model-free approach, able to disentangle main features of DTD without imposing specific constraints, is an appealing option for assessing proper variables to study tolerance to heat stress. Principal component analysis (**PCA**) is a multivariate statistical technique able to synthesize complex patterns as the lactation curves for dairy traits in 2 variables with a clear technical meaning (Macciotta et al., 2006, 2015). Principal component analysis can, therefore, be conveniently used to analyze DTD curves for extracting new variables able to synthesize the pattern.

In the present work, a PCA approach was tested to derive indicator variables of tolerance to heat stress from milk production data in dairy cattle. Moreover, a genome-wide association study (**GWAS**) using a medium-density (50K) SNP panel was used for investigating the genetic determinism of these new variables.

MATERIALS AND METHODS

Data

Data were 590,174 test-day (**TD**) records for milk yield (**MY**), fat (**FP**) and protein (**PP**) percentages,