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.

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

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 as-

sumption 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,

and SCS (Ali and Shook, 1980) of 39,261 Italian Holstein cows (first, second, and third parity) from 484 farms, collected from 2001 to 2007. Data were recorded by the Italian Breeders Association according to International Committee for Animal Recording standards (http://www.icar.org/index.php/publications-technical-materials/recording-guidelines/). Age at calving classes were established for each parity according to the following thresholds: 20 to 36 (17 classes), 31 to 50 (20 classes), and 42 to 65 (24 classes) months of age for first-, second-, and third-parity cows, respectively. All cows had first-lactation data and a minimum of 8 TD records per lactation (from 5–305 DIM). A minimum of 24 records per herd-year of calving were required. Cows were sired by 4,184 AI bulls.

Daily weather information were collected from 35 meteorological stations located no more than 5 km from the considered herd. The THI index (Kelly and Bond, 1971) was then calculated as:

$$THI = (1.8 \times AT + 32) - (0.55 - 0.55 \times RH) \times [(1.8 \times AT + 32) - 58],$$

where AT is the maximum daily temperature, expressed in degrees Celsius, and RH is the minimum relative humidity, expressed as a percentage.

Statistical Analysis

Test-day records were first analyzed with the following mixed linear model

$$y = month(year) + age + DIM \times parity$$
$$+ herd(year) + e, \qquad [1]$$

where y is the record for MY, FP, PP, or SCS; month(year) is the fixed effect of the month of calving (12 mo) nested within the year of calving (7 yr, 2001–2007); age is the fixed effect of age class in months (61 classes, from 20 to 65 mo); DIM × parity is the interaction between the fixed effect of the DIM class (10 intervals of 30 d each) and the fixed effect of parity (3 parities, 1–3); herd(year) is the random effect of the herd (458) nested within calving year; and e his the random residual.

Residuals of model [1] are therefore production data adjusted for main systematic factors, except from additive genetic and THI effects. On the basis of THI values, records were grouped into 11 THI classes (1 = 50–52, 2 = 53–54, …, 11 = >79). Distribution of records across THI classes is reported in Figure 1. Means of residuals were calculated for each bull and THI class for obtaining DTD (Nguyen et al., 2016). The DTD
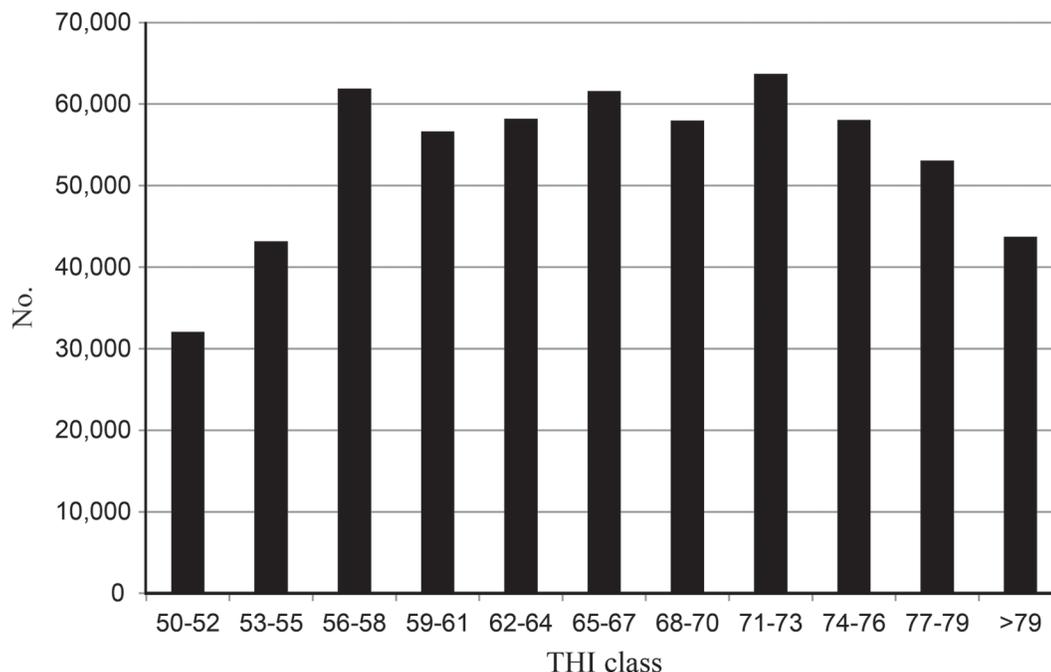
Figure 1. Distribution of test-day records across different classes of temperature-humidity index (THI).

plotted against the THI class express the sensitivity of bull's daughter production performance for increasing THI levels.

The following step was to derive a measure able to summarize the shape of these curves that could be used as dependent variable in a GWAS. A PCA was carried out on the 11 points of the DTD curves, considered as different variables; for example, the first bull for milk yield had 11 records (i.e., DTD1_MY, DTD2_MY, ..., DTD11_MY). Only bulls that had the complete set of 11 DTD values for each trait were considered (1,540 for MY, 1,513 for FP, 1,536 for PP, and 1,535 for SCS, respectively). The PCA was carried out using the SAS PRINCOMP procedure (SAS Institute, 2008). The number of principal components to be retained was based on their eigenvalue, and on their relationships with the original variables.

Principal component (**PC**) scores were then calculated for each bull and treated as new phenotypes for performing either genetic parameter estimation or GWAS. Variance components for PC scores were estimated with the following multi-trait animal model:

$$\mathbf{y} = \mu + \text{animal} + e,$$

where $\mathbf{y}$ is a vector of PC scores for MY, FP, PP, and SCS, respectively; μ is the overall mean; animal is the random additive genetic effect; and e is the residual term. The following (co)variance structure was assumed for the random effects:

$$\text{var}\begin{vmatrix} a \\ e \end{vmatrix} = \begin{vmatrix} \mathbf{A} \otimes \mathbf{G}_0 & 0 \\ 0 & \mathbf{I} \otimes \mathbf{R}_0 \end{vmatrix},$$

where $\mathbf{A}$ is the numerator relationship matrix, $\mathbf{G}_0$ is the matrix of (co)variances for additive effects, $\mathbf{R}_0$ is a diagonal matrix of residual variances corresponding to each trait, and $\mathbf{I}$ is an identity matrix. The pedigree file had 21,685 animals, including the 1,540 sires with DTD in the data set.

The model was solved using the program AIREML90 (Misztal et al., 2002). Considering that PC scores were calculated starting from average yields per bull, the ratio was

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2},$$

where σ_A^2 and σ_E^2 are the additive genetic and the residual variances, respectively; h^2 represents an approximation of the true heritability because averaging affects the variability of the response (different number of TD records per bull). Thus, obtained values have been properly called as pseudo-heritability.

Of the 1,540 bulls considered for the PC score calculation, 423 were genotyped with the Illumina 50K bovine bead chip (Illumina, San Diego, CA). Monomorphic SNP (7,140) and SNP with a call rate <95% (1,045) were discarded. In total, 45,546 SNP were retained for the analysis. Genome-wide scan was performed on PCA scores with the GenABEL R package (Aulchenko et al., 2007), using the GRAMMAR procedure. First, an additive polygenic model was fitted to obtain individual residuals using the genomic relationship matrix. Then, SNP association was tested using a linear model on residuals of the first step. The SNP statistical significance was corrected for the stratification of the population using the genomic control (**GC**) option (Amin et al., 2007). The GC-corrected P-values (GC_P_i) were further corrected for multiple testing using either (1) the Bonferroni correction, obtained as GC_P_i × m (where m is the number of performed tests); (2) and calculating the false discovery rate (**FDR**), as (GC_P_i × m)/m_0, where m_0 is the number of tests having the GC P-values lower or equal to GC_P_i. A marker was declared significantly associated with a trait when the FDR was <0.10.

Gene discovery analysis was carried out considering windows of 0.5 Mb surrounding the significant marker (0.25 Mb up- and downstream, respectively). Genes were derived from UCSC Genome Browser Gateway (http://genome.ucsc.edu/). Both SNP and gene positions were obtained from the UMD3.1 bovine genome assembly (Zimin et al., 2009).

RESULTS

PCA

Eigenvectors and eigenvalues of the first 2 PC extracted from all the 4 considered phenotypes are reported in Table 1. The first 2 PC explained from 42 to 51% of the total variance of the system across the 4 traits. The choice of retaining only the first 2 PC was motivated by the magnitude of single eigenvalues, even though the amount of explained variance was not particularly relevant. A common criterion used for retaining PC is that the eigenvalue should be greater than 1. In the present work, for all 4 traits only the first eigenvalue fulfilled this requirement, and the second eigenvalue was very close to 1 (Table 1).

The first principal component (**PC1**) showed positive and moderate eigenvector coefficients or loadings (ranging from 0.25 to 0.35) with all the original variables (Table 1). Thus, PC1 can be considered as a mea-

Table 1. Eigenvectors and eigenvalues of the first 2 principal components (level and slope) extracted from the correlation matrix of daughter trait deviations for milk yield (DTD_MY), fat (DTD_FP) and protein (DTD_PP) percentages, and SCS (DTD_SCS)

Temperature-humidity index interval	DTD_MY		DTD_FP		DTD_PP		DTD_SCS	
	Level	Slope	Level	Slope	Level	Slope	Level	Slope
50–52	0.25	0.45	0.24	0.76	0.26	−0.48	0.22	0.49
53–55	0.28	0.41	0.29	0.37	0.29	−0.37	0.25	0.55
56–58	0.32	0.22	0.30	0.17	0.31	−0.20	0.33	0.07
59–61	0.31	0.14	0.31	−0.25	0.32	−0.07	0.30	0.12
62–64	0.32	0.17	0.31	0.03	0.32	−0.18	0.34	0.09
65–67	0.32	0.00	0.31	−0.02	0.33	−0.06	0.30	0.01
68–70	0.32	−0.04	0.32	−0.14	0.31	0.10	0.34	−0.11
71–73	0.32	−0.13	0.32	−0.28	0.31	0.22	0.31	−0.06
74–76	0.31	−0.31	0.31	−0.02	0.31	0.11	0.33	−0.26
77–79	0.30	−0.34	0.31	−0.30	0.29	0.36	0.32	−0.14
>79	0.25	−0.55	0.28	−0.11	0.25	0.60	0.27	−0.57
Eigenvalue	4.43	0.96	4.25	0.91	4.79	0.91	3.60	0.98
Eigenvalue %	40	9	39	8	43	8	33	9

sure of the level at which the curve is located and it was named level. Bulls with large scores for this component have their DTD pattern located on a higher level. The second principal component (**PC2**) exhibited larger loadings (up to 0.76 for DTD_FP) with both positive and negative signs. In particular, PC2 showed positive values for the first part of THI interval and negative for the second part for MY, SCS, and, even if less definite, for FP; on the contrary, PC2 for PP showed the opposite trend. For its structure, PC2 defined the shape of the DTD curve. Therefore, larger values (or smaller in the case of protein content) of PC2 scores characterize DTD curves with a decreasing pattern, whereas smaller (or larger) PC2 scores indicate increasing patterns. This component was named slope. The interpretation of level and slope meaning may be better inferred from the average DTD curves for different PC1 and PC2

classes (Figures 2). Only MY data were reported for brevity, but the other traits showed the same pattern.

To simplify the comparison between slope for different traits, scores of slope for PP were multiplied by −1. Pearson correlations among PC scores (Table 2) confirm the expected orthogonality between level and slope within each trait. Sign and magnitude of correlations between PC scores of different traits confirm the meaning of the new variables extracted; for example, the negative correlations between level for milk yield and for fat and protein percentage and the positive correlation between the last 2 traits.

Genetic Parameter Estimation

All new variables exhibited genetic variability (Table 3). In particular, pseudo-heritability had a moderate
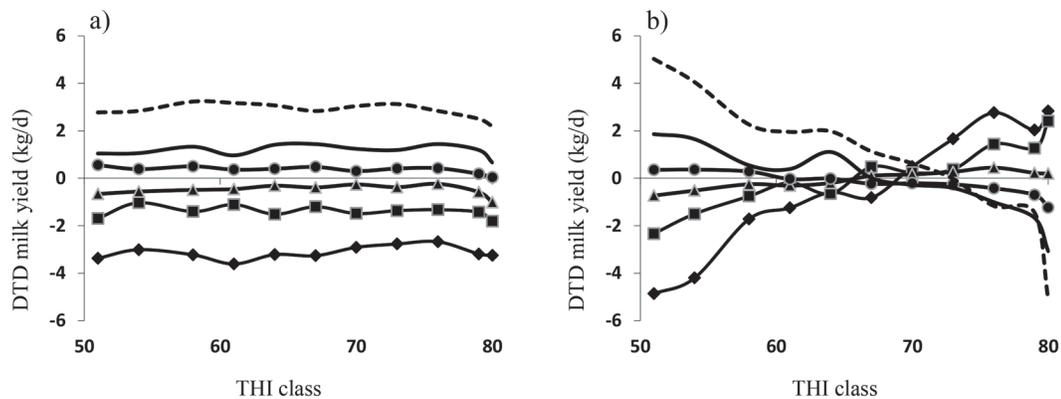
Figure 2. Average curves of daughter trait deviations (DTD) for milk yield of groups of bulls of different level (a) and slope (b) score classes (♦ = <−2; ■ = −2 to −1; ▲ = −1 to 0; ● = 0 to 1; continuous line = 1 to 2; dotted line = >2). Points are plotted for the average DIM on each test day. THI = temperature-humidity index.

Table 2. Pearson correlations between the scores of the first principal components (level and slope) extracted from the correlation matrix of daughter trait deviations for milk yield (MY), fat (FP) and protein (PP) percentage, and SCS

Item	Level MY	Slope MY	Level FP	Slope FP	Level PP	Slope PP	Level SCS	Slope SCS
Level MY	1.00	0.00	−0.37	−0.03	−0.34	0.00	−0.07	−0.07
Slope MY		1.00	0.01	−0.12	−0.01	−0.20	−0.04	−0.16
Level FP			1.00	0.00	0.52	0.02	−0.05	0.04
Slope FP				1.00	0.03	0.23	−0.02	0.05
Level PP					1.00	0.00	0.05	0.02
Slope PP						1.00	0.02	−0.15
Level SCS							1.00	0.00
Slope SCS								1.00

to high level across all traits (values ranged between 0.32 for SCS and 0.82 for FP, respectively) and low to moderate slope (ranging from 0.16 for SCS to 0.28 for PP, respectively). Moreover, values for slope of FP and PP were consistent with the proportion of variability explained by one of the canonical variables obtained for the eigendecomposition of the additive (co)variance matrix of random regression coefficients for these traits (Carabaño et al., 2014). Pseudo-genetic correlations confirm at a genetic level the substantial orthogonality between the level and slope components within each trait. Values of pseudo genetic correlation r_g between level values across the different traits were large (absolute values >0.65), with the exception of comparisons involving SCS. Large pseudo-genetic correlations were also observed between slope values for all traits, with the exception of the correlation between FP and SCS (Table 3).

GWAS

Eight SNP were significantly associated with the considered traits (FDR <0.10; Bonferroni *P*-value <0.08; Table 4). Two SNP were associated with the slope of MY. Four, including the 3 top significant SNP, were associated with the level of FP and 2 for PP (1 to level and 1 to slope). No significant association was found for principal components extracted from DTD for SCS.

MY. The 2 markers significantly associated with DTD_MY were both related to slope (Table 4 and Figure 3a), the PC that expresses the shape of individual curves for increasing levels of THI. The first SNP was located on BTA26 at approximately 22.3 Mb. A possible candidate gene located within the interval defined by this marker is the β-transducin repeat containing E3 ubiquitin protein ligase (*BTRC*; Table 5), reported to be associated with milk production (Raven et al., 2016) and leg morphology (van den Berg et al., 2014) in cattle, and with growth rate in chicken (Zhang et al., 2015). The 0.5-Mb window includes also genes involved in folliculogenesis in cattle, such as fibroblast growth factor 8 (*FGF8*), found in a selection sweep study in dairy cattle (Kemper et al., 2014), meningioma expressed antigen 5 (hyaluronidase; *MGEA5*), and taurus Kv channel interacting protein 2 (*KCNIP2*; Hatzirodos et al., 2014). Of interest is also Hermansky-Pudlak syndrome 6 (*HPS6*), a gene related to pigmentation in humans (Sturm and Duffy, 2012). A marker located on BTA26 and significantly associated with sweating rate was reported by Dikmen et al. (2013) for Holsteins, although the map position is about 2.0 Mb from the SNP identified in our study.

The second marker associated with slope for MY was located on BTA6 at approximately 35.5 Mb (Table 4). In this region maps the coiled-coil serine rich protein 1 (*CCSER1*; Table 5), a gene involved in the mechanism

Table 3. Pseudo-heritability (on the diagonal) and pseudo-genetic correlations (out of the diagonal) between the scores of the first principal components extracted from the correlation matrix of daughter trait deviations for milk yield (MY), fat (FP) and protein (PP) percentage, and SCS (SE of pseudo-heritabilities in parentheses)

Item	Level MY	Slope MY	Level FP	Slope FP	Level PP	Slope PP	Level SCS	Slope SCS
Level MY	0.52 (0.07)	0.02	−0.67	−0.05	−0.67	0.01	−0.05	−0.19
Slope MY		0.24 (0.04)	0.04	−0.56	−0.05	−0.83	−0.17	−0.85
Level FP			0.82 (0.11)	0.01	0.75	0.04	−0.10	0.14
Slope FP				0.23 (0.01)	0.06	0.88	−0.10	0.36
Level PP					0.74 (0.11)	0.00	0.11	0.10
Slope PP						0.28 (0.01)	−0.07	0.69
Level SCS							0.32 (0.07)	0.04
Slope SCS								0.16 (0.05)

Table 4. Markers significantly associated [false discovery rate (FDR) <0.10] with scores of principal components extracted from daughter trait deviations for milk yield (MY) and fat (FP) and protein (PP) contents

SNP	BTA	Position (bp)	P-GC_P Bonf[1]	FDR	Trait
ARS-BFGL-NGS-4939	14	1,801,116	0.000001	0.000001	Level FP
ARS-BFGL-NGS-57820	14	1,651,311	0.000021	0.000011	Level FP
ARS-BFGL-NGS-107379	14	2,054,457	0.000335	0.000112	Level FP
ARS-BFGL-NGS-29678	26	22,383,645	0.000973	0.000973	Slope MY
Hapmap30383-BTC-005848	14	1,489,340	0.049244	0.012311	Level FP
ARS-BFGL-NGS-19275	5	114,818,206	0.053830	0.053830	Slope PP
Hapmap32110-BTA-153952	6	35,555,247	0.077947	0.038974	Slope MY
Hapmap32435-BTC-012188	14	56,075,435	0.081155	0.081155	Level PP

[1]Level of significance of the test adjusted with Bonferroni correction.

of cell division found to be associated with birth weight in Brangus cattle (Saatchi et al., 2014) and with Na concentration in muscle of Nelore cattle (Tizioto et al., 2015). Dikmen et al. (2013, 2015) found markers significantly associated with rectal temperature and respiration rate on BTA6, but at a position about 10 Mb from the marker flagged in the present study.

FP. The 4 significant SNP detected for level of FP were located on BTA14 in an interval of approximately 0.6 Mb (Table 4). The top significant SNP (ARS-BF-GL-NGS-4939) was reported to be significantly associated with milk fat in Italian, US, and German Holsteins (Cole et al., 2011; Wang et al., 2012; Capomaccio et al., 2015), in Italian Simmental cattle (Macciotta et al., 2015), and in a multibreed population (Raven et al., 2014). The 0.5-Mb window surrounding this top marker flagged a zone characterized by a high density of annotated genes. In particular, in that region maps one of the most important genes affecting milk fat content and yield in cattle (Table 5), the diacylglycerol o-acyl-transferase 1 (*DGAT1*) (Grisart et al., 2002). However, on BTA14 at approximately 1.81 to 1.83 Mb maps the heat shock transcription factor 1 (HSF1), a protein that is involved in the mechanism of response to heat stress (Guettouche et al., 2005). Among the genes that map in the interval defined by the second marker (Table 4), of interest is the Rho GTPase activating protein 39 (*ARHGAP39*), involved in the development of the central nervous system (Ma and Nowak, 2011) (Table 5). This gene has been reported to be associated with milk fat composition in Danish (Buitenhuis et al., 2014) and North American (Nayeri et al., 2016) Holstein, and to SCS (Wang et al., 2015) in Chinese Holstein. Another interesting gene located in this region is the ribosomal protein L8 (*RPL8*), involved in the cellular mechanisms of homeostasis (Katz et al., 2016), associated with response to acute heat stress in the fish *Lates calcarifer* (Newton et al., 2012).

The third marker located on BTA14 displayed the mitogen-activated protein kinase 15 (***MAPK15***)

(Table 5), suggested as a candidate gene for SCS in a study on Chinese Holstein (Wang et al., 2015). The last significant SNP for PC1 of DTD_FP was also located on BTA14, very close to the second significant SNP. Close to this marker maps the zinc finger protein 34 (*ZNF34*), found to be associated with milk fat percentage in Chinese Holstein (Jiang et al., 2014) and Italian Simmental (Macciotta et al., 2015). No significant SNP were detected for the slope component for fat percentage (Figure 3b)

PP. A SNP significantly associated with slope of PP was found on BTA5, at approximately 114.8 Mb (Table 4 and Figure 3c). In the 0.5-Mb interval flanking this marker is an interesting gene, malonyl-CoA-acyl carrier protein transacylase (*MCAT*; Table 5), expressed in the mitochondrion and involved in fatty acid metabolism. This gene has been found to be differentially expressed in chicken embryos exposed to heat challenge (Loyau et al., 2016). Other genes of interest located in this interval are the sorting and assembly machinery component (*SAMM50*), a mitochondrial protein found associated with serum triglyceride levels in humans (Kitamoto et al., 2013), and the translocator protein (*TSPO*), a gene upregulated in atretic bovine follicles (Hatzirodos et al., 2014).

The other significant marker associated with PP was related to level. It was located on BTA14, in a region where no annotated genes have been retrieved in the UCSC genome database.

DISCUSSION

Traits able to describe efficiently the response of animals to heat stress are rather problematic to be routinely measured in dairy cattle populations involved in breeding programs. Dikmen et al. (2012) estimated that 13 to 17% of the variation in rectal temperature in cows during heat stress is due to genetic differences. However, it will not be practical to select cows for heat tolerance based on rectal temperature directly because
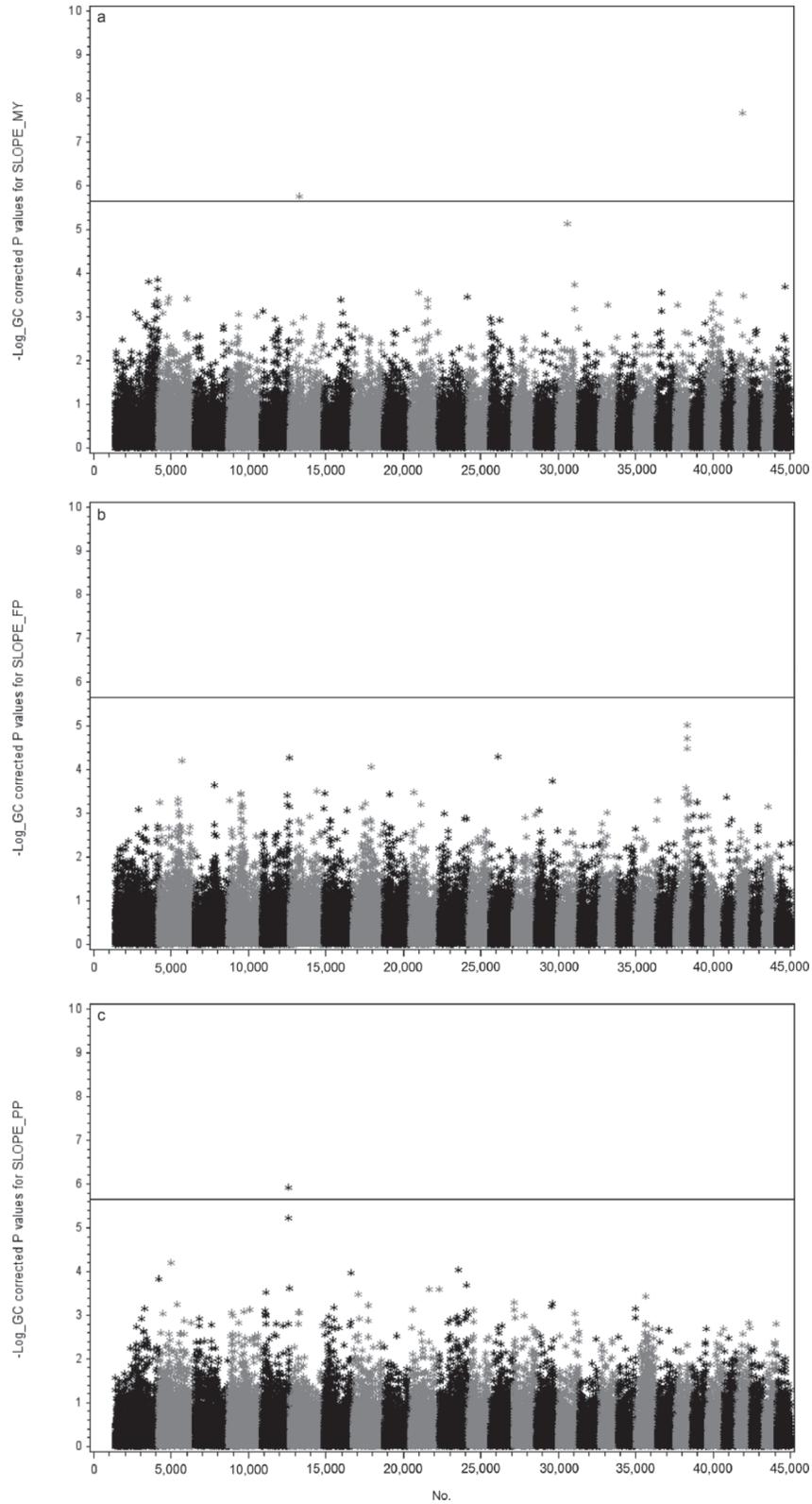
Figure 3. Genome-wide association study of the scores of the second principal component (slope) for milk yield (MY; a), fat percentage (FP; b), and protein percentage (PP; c). On the vertical axis is the negative logarithm of the P-value corrected for the stratification of the population. On the horizontal axis are the SNP ordered by their position and by chromosome. The line corresponds to a false discovery rate of 0.10. GC = genomic control.

Table 5. Putative candidate genes located in the 0.5-Mb interval surrounding the significant SNP

BTA	Position (Mb)	Gene name and symbol	Trait[1]	Function	Associations with phenotypic traits
26	22.09–22.17	β-Transducin repeat containing E3 ubiquitin protein ligase, BTRC	Slope MY	Intracellular protein degradation	Milk production, leg morphology, growth rate
26	22.37–22.38	Fibroblast growth factor 8, FGF8	Slope MY	Mitogenic and cell survival	Folliculogenesis
26	22.39–22.42	Meningioma expressed antigen 5 (hyaluronidase), MGEA5	Slope MY	Protein glycosilation	Folliculogenesis
26	22.42–22.43	TaurusKv channel interacting protein 2, KCNIP2	Slope MY	Ion transportation	Folliculogenesis
26	22.62–22.63	Hermansky-Pudlak syndrome 6, HPS6	Slope MY	Melanosome development	Pigmentation
6	35.10–35.95	Coiled-coil serine rich protein 1, CCSER1	Slope MY	Cell division	Beef traits
14	1.79–1.81	Diacylglycerol O-acyltransferase 1, DGAT1	Level FP	Fat metabolism	Milk fat content
14	1.81–1.83	Heat shock transcription factor 1, HSF1	Level FP	Response to cell stress	Heat stress tolerance, *Mycobacterium bovis* susceptibility
14	1.56–1.60	Rho GTPase activating protein 39, ARHGAP39	Level FP	Nervous system development	Milk fat content
14	1.50–1.51	Ribosomal protein L8, RPL8	Level FP	Homeostasis	Heat stress (fish)
14	2.23–2.24	Mitogen-activated protein kinase 15, MAPK15	Level FP	Cell proliferation	SCS
14	1.49–1.50	Zinc finger protein 34, ZNF34	Level FP	Transcription regulation	Milk fat
5	11.45–11.46	Malonyl-CoA-acyl carrier protein transacylase, MCAT	Slope PP	Fatty acid metabolism	Heat stress (chicken)
5	11.49–11.50	Sorting and assembly machinery component, SAMM50	Slope PP	Mitochondrial protein	Triglyceride levels (humans)
5	11.45–11.46	Translocator protein, TSPO	Slope PP	Cholesterol transportation	Folliculogenesis

[1]Level and slope of principle components for milk yield (MY), fat percentage (FP), and protein percentage (PP).

this trait is not recorded on dairy farms. On the contrary, patterns of milk production traits can be conveniently modeled to estimate tolerance to heat stress (Carabaño et al., 2014; Bernabucci et al., 2014; Hammami et al., 2015). In particular, DTD have been proposed as a proxy of individual response of bulls across increasing levels of heat load (Hayes et al., 2009). The reliability of heat tolerance proxies based on production traits was underlined in a recent study on dairy cattle subjected to heat challenge under a controlled environment, namely climate chambers (Garner et al., 2016). Heat-tolerant cows, ranked according to a genomic breeding value of heat tolerance based on milk yield, showed lower values of physiological indicators (core body temperature, rectal and intravaginal temperature) than heat-sensitive cows.

In the present work, DTD calculated for 4 different milk production traits were analyzed with PCA. This approach was able to extract 2 new variables, which explained approximately 50% of the original variance. Interestingly, they were related to the level of the DTD curve and to its slope, respectively. Our interpretation of PC meaning is in agreement with the outcome of eigendecomposition of coefficient matrix of random regression models used to estimate heat tolerance (Carabaño et al., 2014) and with previous reports on PCA carried out on milk production traits in cattle (Macciotta et al., 2006, 2015).

Most of the genetic models used to study the heat stress effect on dairy traits use the reaction norm model approach (Kolmodin and Bijma, 2004; Shariati et al., 2007). In particular, for each animal the THI effect is fitted as a general intercept plus a slope (Hayes et al., 2009; Sánchez et al., 2009; Carabaño et al., 2014). The results of the model-free PCA approach used in the present study basically confirm theoretical assumptions of the reaction norm model. Also, the predominance of level over slope in terms of variance explained is in agreement with previous reports. In particular, slope eigenvalue (about 10%) is not far from values reported for the second eigenfunction for fat and protein yield obtained from the decomposition of (co)variance matrix of random regression models by Carabaño et al. (2014). These suggest that the behavior of milk production traits across increasing THI levels can be partitioned into 2 main components, one basically related to the overall production genetic potential of the animal and the other to the individual specific response.

The slope component could be proposed as a measure of individual tolerance to heat stress; however, the interpretation of its values deserves further discussion. From Figures 3b it can be noted that animals with negative slope scores exhibited, on average, lower DTD in the first part of the THI scale (i.e., in the comfort

Figure 4. Individual patterns of daughter trait deviations (DTD) for milk yield, which have scores >2 and <−2 for the level and slope components, respectively. THI = temperature-humidity index.

points mostly having values higher than zero or showing great variability.

Another point of agreement between the results of the present study and those of other studies on heat tolerance based on the broken line model can be found in the structure of PC eigenvectors. It is worth noting that the inversion of eigenvector coefficient sign of slope (Table 1) occurs in the THI class 68 to 70 for DTD of MY, PP, and SCS, and in the class 65 to 67 for FP. These values basically agreed with estimates of TH_0 threshold reported by some authors. Sánchez et al. (2009) estimated a THI threshold of 72 for milk yield in US Holsteins using hierarchical models. Carabaño et al. (2016) estimated TH_0 thresholds of 72 to 73 for milk yield in Holstein populations of different European countries. The lower threshold of fat percentage compared with the other traits reported in the present study also agrees with previous reports. In our previous study, THI thresholds of 76, 73, and 74 were estimated in first-, second-, and third-parity Holstein cows, respectively, for MY (Bernabucci et al., 2014).

In view of a possible implementation of PCA-derived measures of tolerance to heat stress in breeding programs, the 2-stage approach suggested in the current study has the limitation that only bulls having complete records (i.e., the 11 points of the DTD curve) could be considered. Such a requirement strongly reduced the number of animals for which PC could be computed (for example, from 4,184 to 1,540 for MY in the present study). To overcome this problem, the correlation matrix between different points of the DTD curve could be reconstructed by using a random regression model (**RRM**) in which records at different THI classes are treated as repeated measures for each sire. Thus, DTD_MY were fitted with a mixed model having the same structure of [1] but with the sire effect fitted as random. The covariance within animal was accounted for by an 11×11 unstructured matrix of between sire effects for each THI level. To avoid convergence problems, bulls having 7 or more records

zone) and higher in the second part. Thus, selection in favor of this kind of response pattern would result in less productive animals under comfort conditions (i.e., the most frequent at least in temperate areas). However, it should be remembered that graphs depicted in Figures 2 represent deviations from the average response pattern (Carabaño et al., 2016). Thus, animals exhibiting a curve as the one reported in Figure 2b are not expected to respond with a decrease of production to lower THI levels, but to have a quite constant average level of production in this point of the curve. Moreover, provided the orthogonality between the 2 traits, the crossed distribution of individual patterns among the classes of the 2 PC (Table 6) shows values in all the cells. Although intermediate classes (i.e., those having values between −1 and 1 for both the components) are the most abundant, 10 bulls belong to the class having the most positive and negative values for level and slope, respectively. The DTD_MY patterns of some of these bulls are reported in Figure 4; they exhibit the typical ascending pattern of this slope class, with the

Table 6. Absolute frequencies of individual patterns of daughter trait deviations for milk yield across different classes of level and slope principal component scores

Level	Slope					
	≤−2	−2 to −1	−1 to 0	0 to 1	1 to 2	≥2
≤−2	10	24	72	81	21	6
−2 to −1[1]	4	15	82	80	25	4
−1 to 0	5	32	138	137	31	4
0 to 1	4	27	128	137	28	3
1 to 2	7	21	90	84	16	8
≥2	10	32	70	62	27	15

[1]Lower limit of the class is included.

Table 7. Pearson correlations among daughter trait deviations for milk yield (above the diagonal) and estimated unstructured correlation matrix estimated with random regression model (under the diagonal)[1]

Item	THI1	THI2	THI3	THI4	THI5	THI6	THI7	THI8	THI9	THI110	THI11
THI1		0.41	0.37	0.35	0.36	0.37	0.33	0.29	0.31	0.30	0.23
THI2	0.43		0.42	0.40	0.40	0.40	0.37	0.36	0.35	0.33	0.25
THI3	0.38	0.44		0.43	0.46	0.44	0.43	0.41	0.43	0.36	0.26
THI4	0.35	0.41	0.43		0.41	0.44	0.42	0.42	0.40	0.34	0.33
THI5	0.35	0.40	0.44	0.41		0.42	0.44	0.41	0.43	0.37	0.30
THI6	0.34	0.39	0.42	0.42	0.42		0.44	0.44	0.42	0.43	0.35
THI7	0.32	0.37	0.41	0.41	0.43	0.40		0.43	0.45	0.43	0.34
THI8	0.28	0.35	0.38	0.40	0.37	0.41	0.39		0.42	0.41	0.43
THI9	0.27	0.30	0.35	0.36	0.39	0.39	0.44	0.42		0.45	0.37
THI10	0.23	0.26	0.33	0.27	0.31	0.37	0.43	0.42	0.47		0.41
THI11	0.18	0.22	0.22	0.26	0.26	0.34	0.34	0.41	0.41	0.43	

[1]THI = temperature-humidity index.

were considered. A total of 35,992 records belonging to 3,697 bulls were used. From Table 7, it can be seen that observed Pearson correlations among DTD_MY and the correlation matrix estimated by the RRM are very similar. Finally, PCA carried out on the RRM estimated correlation matrix yielded basically the same results (Table 8) of the 2-step approach, both in terms of variance explained by the first 2 PC and of their eigenvector structure (i.e., the first can be considered a level PC, the second a slope). Such a concordance of results opens interesting perspectives on the possible use of this heat tolerance indices on large scale.

Previous studies on the genetic basis of tolerance to heat stress indicated that this trait has a moderate genetic variability and that it could be split into an intercept and a slope component that are genetically related. The PC-based measure of heat tolerance proposed in the present work does show genetic variability. In particular, the pseudo-heritability of the slope component is similar to heritability values reported in other studies for the slope parameter of the reaction norm model. Thus, the 2 PC extracted from DTD of 4 milk traits could be considered as possible breeding criteria when selecting for improved heat tolerance in cattle. However, compared with previous measurements of tolerance to heat stress based on milk production data, a distinguishing feature of the PC is their phenotypic and genetic orthogonality. The absence of any genetic relationship between level and slope (Table 2) suggests that an independent selection of the 2 main aspects of DTD patterns (i.e., level of production and heat tolerance) should be feasible. A simultaneous selection for improving both heat tolerance and dairy traits could be achieved also by implementing a selection index in which suitable economic weight have to be determined (Nguyen et al., 2016). However, provided the unfavorable genetic correlation between heat tolerance and production, a smaller selection response is expected for each single trait in comparison with the use of the 2 uncorrelated level and slope components.

The amount of variance accounted for by the slope parameter in reaction norm models is used as an indicator of the genotype × environment interaction (Kolmodin and Bijma, 2004; Shariati et al., 2007). In the present study, slope and level additive variance ratios were 0.10 for MY, 0.05 for FP, 0.06 for PP, and 0.13 for SCS. These values confirm results of Santana et al. (2017), who concluded that a genotype × environment interaction due to heat stress is more relevant for milk yield and SCS than for fat and protein percentage.

Association analysis highlighted a limited number of significant markers. This was not an unexpected outcome due to the severe correction needed to account for multiple testing when high-throughput platforms are used, to the complex biology underlying the physiological response to heat stress, and to the limited size of the sample of animals genotyped.

Table 8. Eigenvectors and eigenvalues of the first 2 principal components (PC) extracted from the correlation matrix estimated by fitting a random regression model to daughter trait deviations for milk yield (DTD_MY)

Temperature humidity index interval	PC1	PC2
50–52	0.26	−0.40
53–55	0.30	−0.36
56–58	0.31	−0.27
59–61	0.31	−0.21
62–64	0.31	−0.18
65–67	0.32	−0.03
68–70	0.32	0.07
71–73	0.31	0.19
74–76	0.31	0.30
77–79	0.29	0.42
>79	0.27	0.51
Eigenvalue	4.66	1.15
Eigenvalue %	42	10

Half of the significant SNP were associated with the level variable for FP and located on BTA14. This result is quite common in genomic studies on Holstein cattle, mainly due to the genetic architecture of the trait, which is largely influenced by a single segregating gene, *DGAT1* (Grisart et al., 2002). The relevant influence of this gene on FP was previously observed in the Italian Holstein population (Fontanesi at al., 2014).

However, it is worth mentioning that a gene encoding for a heat shock transcription factor (*HSF1*) maps on BTA14 very close to the top significant marker detected in the present study. In humans, *HSF1* mediates the expression of the heat shock and stress proteins in response to physical and chemical stresses (Guettoche et al., 2005). This gene has been recognized to have a central role in coordinating thermal tolerance in cattle (Collier et al., 2008). Differences in the expression of *HSF1* in the liver were found between cows calving in spring and summer, respectively (Shahzad et al., 2015). Moreover, associations between this gene and thermotolerance have been detected in Chinese Holstein (Li et al., 2011). An overexpression of *HSF1* has been found in buffalo during summer under tropical environment (Kumar et al., 2015) and is reported to be associated with genetic susceptibility to *Mycobacterium bovis* infection in dairy cattle (Richardson et al., 2016). Differential expression of heat shock proteins has been related to in vitro fertilization rate and blastocyst rate of bovine embryos (Zhang et al., 2011). The *HSP70.1* polymorphism has been associated with cellular thermos tolerance in Holstein lactating cows (Basiricò et al., 2011)

The relationships between tolerance to heat stress and fertility has been confirmed by associations found in the present study. Candidate genes found in the 0.5-Mb intervals defined by the significant markers for the slope component are involved in cellular regulation mechanism, fertility, and weight at birth. Nguyen et al. (2016), using the DTD deviation as indicators of heat stress in Australian Holsteins, found favorable correlation between DTD and fertility.

Among putative candidate genes that have been detected in the present study for the SLOPE trait, one (*MCAT*) is involved in fatty acid metabolism. This result is in agreement with findings by Hammami et al. (2015), who highlighted a relationship between milk fatty acid content and tolerance to heat stress in cattle. It should pointed out that milk fatty acid profile is an index of animal energy balance (Bastin et al., 2011) and is strongly related to the diet, which may be affected by climatic conditions. Nardone et al. (1997) found greater proportions of long-chain fatty acids in colostrum produced by heifers under heat stress conditions. Those authors demonstrated that the higher proportion of long-chain fatty acids was due to the reduced synthesis of short- and medium-chain fatty acids in the mammary gland cells. Thus, the role of milk FA as potent biomarkers for evaluating individual thermotolerance could be hypothesized.

Previous GWAS studies carried out on DTD and rectal temperature have highlighted genomic regions associated with heat stress tolerance. These results were not confirmed in the present study, even though significant markers were seen for slope for milk yield on BTA 5, 6, and 26 [i.e., on the same chromosomes where Dikmen et al. (2013) found significant associations for rectal temperature]. The detection of a limited number of significant markers and a poor repeatability of results across studies and populations is a major issue in GWAS studies carried out on livestock species. Sample size, genetic differences among populations (i.e., level of linkage disequilibrium, allelic SNP frequencies) were mentioned as main reasons for such a lack of concordance between experiments. Moreover, the severe corrections of significance levels due to huge number of repeated tests further reduce the number of detected markers. Finally, a more general issue that has been raised in GWAS carried out in humans is that not all the genetic variation of a trait is captured by available markers (i.e., the so-called problem of missing heritability; Gusev et al., 2013). On the other hand, in the case of GWAS for heat tolerance traits, the role of the phenotypes should be carefully considered. Dikmen et al. (2015) found that 1 out of 4 significant SNP previously detected using a phenotype derived from milk yield was also associated with a physiological indicator of heat stress (i.e., sweating rate). However, SNP validation is problematic also within the same trait. Hayes et al. (2009), on a second independent data set, confirmed only 2 out of 42 SNP significantly associated with the slope component of a reaction norm model fitted to milk yield in a Holstein population. A further issue in genetic studies of tolerance to heat stress is represented by the methodologies used to obtain the environmental variable. The THI calculations, for example, could differ in the kind of variable used (i.e., daily maximum or average temperature, minimum or average relative humidity) and in the time lag with the day of the test. Such heterogeneity of measures could be one of the reasons for the differences between studies in the estimates of the THI upper threshold for the comfort zone. The existence of all these sources of variation that may possibly affect results of studies on the genetic dissection of heat tolerance traits should lead scientists to make efforts to increase the power of their experiments, validate results in independent

populations, and in harmonize methodologies for calculating environmental variables. Furthermore, alternatives to traditional methods for measuring physiological indicators of heat tolerance should be found. Recent achievements in precision farming, such as the use of microequipment directly installed on the animal or the use of indirect predictors of physiological traits (as for example midinfrared milk spectra), could provide interesting tools.

CONCLUSIONS

The PCA was able to derive 2 new variables able to describe the overall level and the slope of response of milk production traits across increasing levels of THI index. These 2 new phenotypes are uncorrelated and may therefore provide an option for overcoming the problem of the negative correlation between heat tolerance and production level that has been found within the context of milk and weather record analysis. The genetic background of the 2 PC was investigated and some putative candidate genes were proposed. A useful numerical property of the 2 extracted variables is their orthogonality. This feature makes the use of the second PC as a measure of thermotolerance for breeding purposes, which is particularly appealing because many authors have stressed the need for using measures of tolerance to heat stress uncorrelated from the production level. Moreover, this variable could be derived from data that are routinely recorded in breeding programs and, therefore, used on large scale could be proposed.

ACKNOWLEDGMENTS

REFERENCES

Aguilar, I., I. Misztal, and S. Tsuruta. 2009. Genetic components of heat stress for dairy cattle with multiple lactations. J. Dairy Sci. 92:5702–5711.

Aguilar, I., I. Misztal, and S. Tsuruta. 2010. Short communication: Genetic trends of milk yield under heat stress for US Holsteins. J. Dairy Sci. 93:1754–1758.

Ali, A. K. A., and G. E. Shook. 1980. An optimum transformation for somatic cell concentration in milk. J. Dairy Sci. 63:487–490.

Amin, N., C. M. van Duijn, and Y. S. Aulchenko. 2007. A genomic background based method for association analysis in related individuals. PLoS One 2:e1274.

Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. van Duijn. 2007. GenABEL: An R package for genome-wide association analysis. Bioinformatics 23:1294–1296.

Basiricò, L., P. Morera, V. Primi, N. Lacetera, A. Nardone, and U. Bernabucci. 2011. Cellular thermotolerance is associated with heat shock protein 70.1 genetic polymorphisms in Holstein lactating cows. Cell Stress Chaperones 16:441–448.

Bastin, C., N. Gengler, and H. Soyeurt. 2011. Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. J. Dairy Sci. 94:4152–4163.

Bernabucci, U., S. Biffani, L. Buggiotti, A. Vitali, N. Lacetera, and A. Nardone. 2014. The effect of heat stress in Italian Holstein dairy cattle. J. Dairy Sci. 97:471–486.

Bernabucci, U., L. H. Lacetera, R. P. Baumgard, B. Rhoads, B. Ronchi, and A. Nardone. 2010. Metabolic and hormonal acclimation to heat stress in domesticated ruminants. Animal 4:1167–1183.

Biffani, S., U. Bernabucci, A. Vitali, N. Lacetera, and A. Nardone. 2016. Short communication: Effect of heat stress on nonreturn rate of Italian Holstein cows. J. Dairy Sci. 99:5837–5843.

Bohmanova, J., I. Misztal, S. Tsuruta, H. D. Norman, and T. J. Lawlor. 2008. Short communication: Genotype by Environment Interaction Due to Heat Stress. J. Dairy Sci. 91:840–846.

Brügemann, K., E. Gernand, U. U. von Borstel, and S. König. 2011. Genetic analyses of protein yield in dairy cows applying random regression models with time-dependent and temperature x humidity-dependent covariates. J. Dairy Sci. 94:4129–4139.

Buitenhuis, B., L. L. G. Janss, N. A. Poulsen, L. B. Larsen, M. K. Larsen, and P. Sørensen. 2014. Genome-wide association and biological pathway analysis for milk-fat composition in DanishHolstein and Danish Jersey cattle. BMC Genomics 15:1112.

Capomaccio, S., M. Milanesi, L. Bomba, K. Cappelli, E. L. Nicolazzi, J. L. Williams, P. Ajmone-Marsan, and B. Stefanon. 2015. Searching new signals for production traits through gene-based association analysis in three Italian cattle breeds. Anim. Genet. 46:361–370.

Carabaño, M. J., K. Bachaga, M. Ramon, and C. Diaz. 2014. Modeling heat stress effect on Holstein cows under hot and dry conditions: selection tools. J. Dairy Sci. 97:7889–7904.

Carabaño, M. J., B. Logar, J. Bormann, J. Minet, M. L. Vanrobays, C. Diaz, B. Tychon, N. Gengler, and H. Hammami. 2016. Modeling heat stress under different environmental conditions. J. Dairy Sci. 99:3798–3814.

Cole, J. B., G. R. Wiggans, L. Ma, T. S. Sonstegard, T. J. Lawlor Jr., B. A. Crooker, C. P. Van Tassell, J. Yang, S. Wang, L. K. Matukumalli, and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S.Holstein cows. BMC Genomics 12:408.

Collier, R. J., J. L. Collier, R. P. Rhoads, and L. H. Baumgard. 2008. Invited review: Genes involved in the bovine heat stress response. J. Dairy Sci. 91:445–454.

Dikmen, S., J. B. Cole, D. J. Null, and P. J. Hansen. 2012. Heritability of rectal temperature and genetic correlations with production and reproduction traits in dairy cattle. J. Dairy Sci. 95:3401–3405.

Dikmen, S., J. B. Cole, D. J. Null, and P. J. Hansen. 2013. Genome-wide association mapping for identification of quantitative trait loci for rectal temperature during heat stress in Holstein cattle. PLoS One 8:e69202.

Dikmen, S., X. Z. Wang, M. S. Ortega, J. B. Cole, D. J. Null, and P. J. Hansen. 2015. Single nucleotide polymorphisms associated with thermoregulation in lactating dairy cows exposed to heat stress. J. Anim. Breed. Genet. 132:409–419.

Fontanesi, L., D. G. Calò, G. Galimberti, R. Negrini, R. Marino, A. Nardone, P. Ajmone-Marsan, and V. Russo. 2014. A candidate gene association study for nine production traits in Italian Holstein sires. Anim. Genet. 45:576–580.

Garner, J. B., M. L. Douglas, S. R. O. Williams, W. J. Wales, L. C. Marett, T. T. T. Nguyen, C. M. Reich, and B. J. Hayes. 2016. Genomic selection improves heat tolerance in dairy cattle. Sci. Rep. https://doi.org/10.1038/srep34114.

Grisart, B., W. Coppieters, F. Farnir, L. Karim, C. Ford, P. Berzi, N. Cambisano, M. Mni, S. Reid, P. Simon, R. Spelman, M. Georges, and R. Snell. 2002. Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1

gene with major effect on milk yield and composition. Genome Res. 12:222–231.

Guettouche, T., F. Boellmann, W. S. Lane, and R. Voellmy. 2005. Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. BMC Biochem. 6:4. https://doi.org/10.1186/1471-2091-6-4.

Gusev, A., G. Bhatia, N. Zaitien, B. J. Vilhjalmsson, D. Diogo, E. A. Stahl, P. K. Gregersen, J. Worthington, L. Klareskog, S. Rayachaudhri, R. M. Plenge, B. Pasanius, and A. L. Price. 2013. Quantifying missing heritability at known GWAS loci. PLoS Genet. 9:e1003993.

Hammami, H., J. Vandenplas, M.-L. Vanrobays, B. Rekik, C. Bastin, and N. Gengler. 2015. Genetic analysis of heat stress effects on yield traits, udder health, and fatty acids of Walloon Holstein cows. J. Dairy Sci. 98:4956–4968.

Hansen, P. J. 2007. Exploitation of genetic and physiological determinants of embryonic resistance to elevated temperature to improve embryonic survival in dairy cattle during heat stress. Theriogenology 68:S242–S249.

Hatzirodos, N., H. F. Irving-Rodgers, K. Hummitzsch, M. L. Harland, S. E. Morris, and R. J. Rodgers. 2014. Transcriptome profiling of granulosa cells of bovine ovarian follicles during growth from small to large antral sizes. BMC Genomics 15:24.

Hayes, B. J., P. J. Bowman, A. J. Chamberlain, K. Savin, C. P. van Tassell, T. S. Sostengard, and M. E. Goddard. 2009. A validated genome wide association study to breed cattle adapted to an environment altered by climate changes. PLoS One 4:e6676.

Jiang, L., X. Liu, J. Yang, H. Wang, J. Jiang, L. Liu, S. He, X. Ding, J. Liu, and Q. Zhang. 2014. Targeted resequencing of GWAS loci reveals novel genetic variants for milk production traits. BMC Genomics 15:1105.

Kadzere, C. T., M. R. Murphy, N. Silanikove, and E. Maltz. 2002. Heat stress in lactating dairy cows: a review. Livest. Prod. Sci. 77:59–91.

Katz, M. J., L. Gandara, A. L. De Lella Ezcurra, and P. Wappner. 2016. Hydroxylation and translational adaptation to stress: Some answers lie beyond the STOP codon. Cell. Mol. Life Sci. 73:1881–1893.

Kelly, C. F., and T. E. Bond. 1971. Bioclimatic factors and their measurements. Page 7 in A Guide to Environmental Research in Animals. Natl. Acad. Sci., Washington, DC.

Kemper, K. E., S. J. Saxton, S. Bolormaa, B. J. Hayes, and M. E. Goddard. 2014. Document selection for complex traits leaves little or no classic signatures of selection. BMC Genomics 15:246.

Kitamoto, T., A. Kitamoto, M. Yoneda, H. Hyogo, H. Ochi, T. Nakamura, H. Teranishi, S. Mizusawa, T. Ueno, K. Chayama, A. Nakajima, K. Nakao, A. Sekine, and K. Hotta. 2013. Genome-wide scan revealed that polymorphisms in the PNPLA3, SAMM50, and PARVB genes are associated with development and progression of nonalcoholic fatty liver disease in Japan. Hum. Genet. 132:783–792.

Kolmodin, R., and P. Bijma. 2004. Response to mass selection when the genotype by environment interaction is modeled as a linear reaction norm. Genet. Sel. Evol. 36:435–454.

Kumar, A., S. Ashrafa, T. Sridhar Gouda, A. Grewalc, S. V. Singha, B. R. Yadavb, and R. C. Upadhyaya. 2015. Expression profiling of major heat shock protein genes during different seasons in cattle (Bos indicus) and buffalo (Bubalus bubalis) under tropical climatic condition. J. Therm. Biol. 51:55–64.

Li, Q. L., Z. H. Ju, J. M. Huang, J. B. Li, R. L. Li, M. H. Hou, C. F. Wang, and J. F. Zhong. 2011. Two novel SNPs in HSF1 gene are associated with thermal tolerance traits in Chinese Holstein cattle. DNA Cell Biol. 30:247–254.

Loyau, T., C. Hennequet-Antier, V. Coustham, C. Berri, M. Leduc, S. Crochet, M. Sannier, M. J. Duclos, S. Mignon-Grasteau, S. Tesseraud, A. Brionne, S. Métayer-Coustard, M. Moroldo, J. Lecardonnel, P. Martin, S. Lagarrigue, S. Yahav, and A. Collin. 2016. Thermal manipulation of the chicken embryo triggers differential gene expression in response to a later heat challenge. BMC Genomics 17:329.

Ma, S., and F. V. Nowak. 2011. The RhoGAP domain-containing protein, Porf-2, inhibits proliferation and enhances apoptosis in neural stem cells. Mol. Cell. Neurosci. 46:573–582.

Macciotta, N. P. P., G. Gaspa, L. Bomba, D. Vicario, C. Dimauro, M. Cellesi, and P. Ajmone-Marsan. 2015. Genome-wide association analysis in Italian Simmental cows for lactation curve traits using a low-density (7K) SNP panel. J. Dairy Sci. 98:8175–8185.

Macciotta, N. P. P., D. Vicario, and A. Cappio-Borlino. 2006. Use of multivariate analysis to extract latent variables related to level of production and lactation persistency in dairy cattle. J. Dairy Sci. 89:3188–3194.

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Pages 743–744 in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. INRA, France.

Nardone, A., N. Lacetera, U. Bernabucci, and B. Ronchi. 1997. Composition of colostrum from dairy heifers exposed to high air temperatures during late pregnancy and the early postpartum period. J. Dairy Sci. 80:838–844.

Nardone, A., B. Ronchi, N. Lacetera, M. S. Ranieri, and U. Bernabucci. 2010. Effects of climate changes on animal production and sustainability of livestock systems. Livest. Sci. 130:57–69.

Nayeri, S., M. Sargolzaei, M. K. Abo-Ismail, N. May, S. P. Miller, F. Schenkel, S. S. Moore, and P. Stothard. 2016. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. BMC Genet. 17:75.

Newton, J. R., C. De Santis, and D. R. Jerry. 2012. The gene expression response of the catadromous perciform barramundi Latescalcarifer to an acute heat stress. J. Fish Biol. 81:81–93.

Nguyen, T. T. T., P. J. Bowman, M. Haile-Mariam, J. E. Pryce, and B. J. Hayes. 2016. Genomic selection for tolerance to heat stress in Australian dairy cattle. J. Dairy Sci. 99:2849–2862.

Perano, K. M., J. G. Usack, L. T. Angenent, and K. G. Gebremedhin. 2015. Production and physiological responses of heat-stressed lactating dairy cattle to conductive cooling. J. Dairy Sci. 98:5252–5261.

Ravagnolo, O., and I. Misztal. 2000. Genetic component of heat stress in dairy cattle, parameter estimation. J. Dairy Sci. 83:2126–2130.

Raven, L. A., B. G. Cocks, and B. J. Hayes. 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genomics 15:62.

Raven, L. A., B. G. Cocks, K. E. Kemper, A. J. Chamberlain, M. E. Goddard, and B. J. Hayes. 2016. Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. Mamm. Genome 27:81–97.

Richardson, I. W., D. P. Berry, H. L. Wiencko, I. M. Higgins, S. J. More, J. McClure, D. J. Lynn, and D. G. Bradley. 2016. A genome-wide association study for genetic susceptibility to Mycobacterium bovis infection in dairy cattle identifies a susceptibility QTL on chromosome 23. Genet. Sel. Evol. 48:19.

Saatchi, M., R. D. Schnabel, J. F. Taylor, and D. J. Garrick. 2014. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC Genomics 15:442–458.

Sánchez, J. P., I. Misztal, I. Aguilar, B. Zumbach, and R. Rekaya. 2009. Genetic determination of the onset of heat stress on daily milk production in the US Holstein cattle. J. Dairy Sci. 92:4035–4045.

Santana, M. L., A. B. Bignardi, R. J. Pereira, G. Stefani, and L. El Faro. 2017. Genetics of heat tolerance for milk yield and quality in Holsteins. Animal 11:4–14. https://doi.org/10.1017/S1751731116001725.

SAS Institute. 2008. SAS/STAT version 9.2. SAS Inst. Inc., Cary, NC.

Segnalini, M., A. Nardone, U. Bernabucci, A. Vitali, B. Ronchi, and N. Lacetera. 2011. Dynamics of the temperature-humidity index in the Mediterranean basin. Int. J. Biometeorol. 55:253–263.

Shahzad, K., H. Akbar, M. Vailati-Riboni, L. Basiricò, P. Morera, S. L. Rodriguez-Zas, A. Nardone, U. Bernabucci, and J. J. Loor. 2015. The effect of calving in the summer on the hepatic transcrip-

tome of Holstein cows during the peripartal period. J. Dairy Sci. 98:5401–5413.

Shariati, M. M., G. Su, P. Madsen, and D. Sorensen. 2007. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. J. Dairy Sci. 90:5759–5766.

Sturm, R. A., and D. L. Duffy. 2012. Human pigmentation genes under environmental selection. Genome Biol. 13:248.

Tizioto, P. C., J. F. Taylor, J. E. Decker, C. F. Gromboni, M. A. Mudadu, R. D. Schnabel, L. L. Coutinho, G. B. Mourão, P. S. N. Oliveira, M. M. Souza, J. M. Reecy, R. T. Nassu, F. A. Bressani, P. Tholon, T. S. Sonstegard, M. M. Alencar, R. R. Tullio, A. R. A. Nogueira, and L. C. A. Regitano. 2015. Detection of quantitative trait loci for mineral content of Nelore longissimus dorsi muscle. Genet. Sel. Evol. 47:15.

van den Berg, I., S. Fritz, S. Rodriguez, D. Rocha, M. Boussaha, M. S. Lund, and D. Boichard. 2014. Concordance analysis for QTL detection in dairy cattle: A case study of leg morphology. Genet. Sel. Evol. 46:31.

Wang, X., P. Ma, J. Liu, Q. Zhang, Y. Zhang, X. Ding, L. Jiang, Y. Wang, Y. Zhang, D. Sun, S. Zhang, G. Su, and Y. Yu. 2015. Ge-nome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. BMC Genet. 16:111.

Wang, X., C. Wurmser, H. Pausch, S. Jung, F. Reinhardt, J. Tetens, G. Thaller, and R. Fries. 2012. Identification and dissection of four major QTL affecting milk fat content in the German Holstein-Friesian population. PLoS One 7:e40711.

Zhang, B., F. Peñagaricano, A. Driver, H. Chen, and H. Khatib. 2011. Differential expression of heat shock protein genes and their splice variants in bovine preimplantation embryos. J. Dairy Sci. 94:4174–4182.

Zhang, G. X., Q. C. Fan, T. Zhang, J. Y. Wang, W. H. Wang, Q. Xue, and Y. J. Wang. 2015. Genome-wide association study of growth traits in the Jinghai Yellow chicken. Genet. Mol. Res. 14:15331–15338.

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, T. S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J. A. Yorke, and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow. Genom. Biol. 10:R42.