

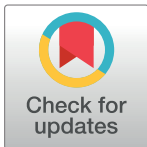
RESEARCH ARTICLE

Identification of candidate genes on the basis of SNP by time-lagged heat stress interactions for milk production traits in German Holstein cattle

Kathrin Halli , Seyi Fridaius Vanvanhossou , Mehdi Bohlouli , Sven König*, Tong Yin *

Institute of Animal Breeding and Genetics, Justus-Liebig-University Gießen, Gießen, Germany

* Sven.Koenig@agrار.uni-giessen.de (SK); Tong.Yin@agrار.uni-giessen.de (TY)



OPEN ACCESS

Citation: Halli K, Vanvanhossou SF, Bohlouli M, König S, Yin T (2021) Identification of candidate genes on the basis of SNP by time-lagged heat stress interactions for milk production traits in German Holstein cattle. PLoS ONE 16(10): e0258216. <https://doi.org/10.1371/journal.pone.0258216>

Editor: Gudrun A. Brockmann, Humboldt-Universität zu Berlin, GERMANY

Received: April 10, 2021

Accepted: September 21, 2021

Published: October 14, 2021

Copyright: © 2021 Halli et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All the datasets used in the study can be found in the repository using the DOI: <http://dx.doi.org/10.22029/jlupub-52>.

Funding: The authors gratefully acknowledge financial support provided by the German National Science Foundation, DFG, through grant number KO 3520/8-1.

Competing interests: The authors have declared that no competing interests exist.

Abstract

The aim of this study was to estimate genotype by time-lagged heat stress (HS) variance components as well as main and interaction SNP-marker effects for maternal HS during the last eight weeks of cow pregnancy, considering milk production traits recorded in the offspring generation. The HS indicator was the temperature humidity index (THI) for each week. A dummy variable with the code = 1 for the respective week for $THI \geq 60$ indicated HS, otherwise, for no HS, the code = 0 was assigned. The dataset included test-day and lactation production traits from 14,188 genotyped first parity Holstein cows. After genotype quality control, 41,139 SNP markers remained for the genomic analyses. Genomic animal models without (model VC_nHS) and with in-utero HS effects (model VC_wHS) were applied to estimate variance components. Accordingly, for genome-wide associations, models GWA_nHS and GWA_wHS, respectively, were applied to estimate main and interaction SNP effects. Common genomic and residual variances for the same traits were very similar from models VC_nHS and VC_wHS. Genotype by HS interaction variances varied, depending on the week with in-utero HS. Among all traits, lactation milk yield with HS from week 5 displayed the largest proportion for interaction variances (0.07). For main effects from model GWA_wHS, 380 SNPs were suggestively associated with all production traits. For the SNP interaction effects from model GWA_wHS, we identified 31 suggestive SNPs, which were located in close distance to 62 potential candidate genes. The inferred candidate genes have various biological functions, including mechanisms of immune response, growth processes and disease resistance. Two biological processes excessively represented in the overrepresentation tests addressed lymphocyte and monocyte chemotaxis, ultimately affecting immune response. The modelling approach considering time-lagged genotype by HS interactions for production traits inferred physiological mechanisms being associated with health and immunity, enabling improvements in selection of robust animals.

Introduction

The tripled number of days with extreme temperatures in European countries from 1950 to 2018 [1] indicates the increasing importance of studying the effect of climate change on primary and functional traits in dairy cattle. Comparing to other domesticated livestock species, high productive Holstein Friesian cattle responded very sensitive to heat stress (HS), especially during the challenging early lactation period with temperatures beyond the thermoneutral zone lower than 5°C or higher than 25°C [2]. Selection on improved milk production simultaneously increased feed intake and body weight, implying a raised metabolic heat increment and a decline of the thermoneutral temperature range [3]. Furthermore, in addition to temperature, dairy cow trait alterations were influenced by humidity, suggesting consideration of a temperature-humidity index (THI) in genetic HS studies [4].

Response to direct HS, i.e., HS close to the trait recording date, was observed when THI exceeds a certain threshold. For example, in Holstein cows, protein yield decreased at THI 68 [5]. Genetically, random regression models were applied to infer genetic (co)variance components along a continuous THI scale [6]. Such modelling approach was also used to detect possible genotype x environment interactions (GxE), and THI dependent re-rankings of sires have been observed [7]. Genomically, genetic markers significantly associated with HS response and underlying candidate genes were detected via genome-wide association studies (GWAS). Four single nucleotide polymorphisms (SNPs) contributing to a milk yield decline under HS were identified by Hayes et al. [8] in Australian Holstein Friesian cattle. The four SNPs were located on *Bos taurus* autosomes (BTA) 8, 10, 25, and 29. Among them, the SNP *BFGL-NGS-30169* on BTA29 was also identified in Jersey cattle, and the strongest annotated candidate gene for the variation of milk yield under HS was the fibroblast growth factor 4 [8]. Selection of cattle displaying the ability to regulate body temperature under HS will contribute to improved heat resistance genetically, implying only marginal detrimental effects on primary and functional traits. Thus, the genetic architecture of traits reflecting the thermoregulation ability under HS, such as rectal temperature and respiration rate, was investigated in Holstein and Gir x Holstein crossbred cattle [9, 10]. Rectal temperature was used an indicator for HS [11], and respiration rate reflects the ability to maintain body temperature through evaporative cooling [10].

Most of the genetic studies focused on immediate responses to HS, i.e., considering THI at the measuring day, or the average THI shortly before the measuring day. Phenotypically, HS during late gestation in dams, termed as maternal or in-utero HS, significantly impaired growth, metabolism, immune functions and survival in offspring [12–14], since this period plays a crucial role in fetal growth [15]. Monteiro et al. [13] studied long-term effects of in-utero HS on female fertility and performance traits recorded in offspring and identified most detrimental impact of HS from the last 6 weeks of gestation. Quantitative-genetically, Halli et al. [16] identified alterations of genetic co(variance) components for weight and growth traits of dual-purpose cattle due to HS impact on their dams during the late pregnancy period.

In this regard, the current study focusses on enhanced genomic modelling approaches to inferring time-lagged HS impact, and considers a large dataset of genotyped Holstein Friesian cows kept in large-scale contract herds. Specifically, we firstly aim on the detection of significant SNP for the direct (main) and the interaction effect in the context of time-lagged HS on milk production traits at the first official test-day after calving and on first lactation production records. The main effect represents the SNP effect for the respective trait expressed consistently, independent from time-lagged HS. The interaction component is the difference of the SNP effects between cows undergoing in-utero HS or not. Afterwards, on the basis of the identified SNP associations, we annotated potential candidate genes. Against this background, i.e.,

in the context of pronounced time-lagged HS, we postulate the identification of candidate genes affecting overall robustness, health and adaptation.

Materials and methods

Cow traits

The official records at the first test-day for the milk production traits included milk yield (TMY), fat percentage (TFP), fat yield (TFY), protein percentage (TPP), protein yield (TPY) and somatic cell score (TSCS). First lactation records were available for milk yield (LMY), fat percentage (LFP), fat yield (LFY), protein percentage (LPP) and protein yield (LPY). The first test-day records of the genotyped cows covered the period between 5 and 35 days after calving. Lactation lengths for LMY, LFP, LFY, LPP and LPY comprised 275 to 305 days in milk. Age at first calving ranged from 20 to 40 months, and calving years from 2010 to 2015. A phenotypic dataset including 14,188 first parity Holstein cows was available for subsequent GWAS. The cows are kept in 53 large-scale herds, located in the German federal states of Mecklenburg-West Pomerania and Berlin-Brandenburg. The average number of cows per herd ranged from 55 to 775. For the test-day records, cows with extreme TMY (≤ 2 kg per day or ≥ 60 kg per day) and TSCS larger than 9,999,000 cells/ml, were excluded. Descriptive statistics for the test-day and the lactation production traits are listed in Table 1.

Genotype data

The cows were genotyped using the *Illumina BovineSNP50 v2 BeadChip* (3,775 animals), or the *Illumina Bovine Eurogenomics 10K low-density chip* (10,413 animals). The 10K SNP genotypes were imputed to the 50K SNP panel by project partner vit (Verden, Germany), as implemented in the process for national routine genetic evaluations [17]. SNP quality controls were performed using the preGSf90 program from the BLUPf90 package [18, 19]. Filtering criteria for markers were as follows: consideration only of SNPs located on *Bos taurus* autosomes, minor allele frequency larger than 0.01, minimum animal and SNP call rate of 0.95 and no significant deviation from Hardy-Weinberg equilibrium (the difference between observed and expected heterozygous frequencies was smaller than 0.15). Furthermore, we deleted cows pairs

Table 1. Descriptive statistics of first test-day and lactation production traits.

Trait ^a	#observations	Mean	Minimum	Maximum	Standard deviation
TMY	12,333	28.78	2.00	58.40	6.31
TFP	12,333	4.14	1.62	10.17	0.78
TFY	12,333	1.18	0.07	3.62	0.28
TPP	12,333	3.24	2.14	5.67	0.33
TPY	12,333	0.92	0.09	1.90	0.18
TSCS	12,307	2.80	-1.06	9.64	1.67
LMY	12,804	9104.29	1994.00	16156.00	1476.15
LFP	12,804	3.82	2.21	5.99	0.45
LFY	12,804	345.92	90.00	571.00	51.33
LPP	12,804	3.35	2.64	4.16	0.20
LPY	12,804	303.37	68.00	528.00	44.44

^a: TMY = first test-day milk yield; TFP = first test-day fat percentage; TFY = first test-day fat yield; TPP = first test-day protein percentage; TPY = first test-day protein yield; TSCS = first test-day somatic cell score; LMY = first lactation milk yield; LFP = first lactation fat percentage; LFY = first lactation fat yield; LPP = first lactation protein percentage; LPY = first lactation protein yield.

<https://doi.org/10.1371/journal.pone.0258216.t001>

with genomic relationships [20] larger than 0.95. Finally, 41,139 SNPs from the 14,188 genotyped cows were used for the ongoing genomic analyses. The test-day data were available from 12,333 genotyped cows, and 12,804 genotyped cows had first lactation records. In total, 10,949 genotyped cows had both test-day and lactation records. The SNP positions were coordinated according to the reference assembly ARS-UCD1.2.

Meteorological data

Pairwise distances (in km) were calculated between weather stations and cow herds. The calculation based on coordinates for the respective longitude and latitude of each herd and weather station, and was performed using the GEOSPHERE package in R [21]. According to the minimal distances, we allocated 32 weather stations to 53 herds. The maximum distance between a herd and a weather station was 27.88 km, the minimum distance was 0.74 km and the average distance was 14.79 km. Hourly THI were calculated considering hourly temperature (T) and relative humidity (RH) as follows [22]:

$$\text{THI} = [1.8 \times T + 32] - [0.55 - 0.0055 \times \text{RH}] \times [1.8 \times T - 26]$$

The daily THI was computed by averaging hourly THI over 24 hours. Afterwards, we calculated the average daily THI within the following weeks during late gestation of the respective dam, or, in other words, before the birth date of the genotyped cows: 0–7 days (WK1), 8–14 days (WK2), 15–21 days (WK3), 22–28 days (WK4), 29–35 days (WK5), 36–42 days (WK6), 43–49 days (WK7) and 50–56 days (WK8). The range of the weekly THI across the eight weeks was from 10.43 to 73.10. Descriptive statistics for the weekly THI are displayed in Table 2.

Statistical models

Variance components and genomic heritabilities. A genomic animal model was applied to estimate variance components for milk production and TSCS. In this regard, we used the AIREML algorithm as implemented in AIREMLf90 from the BLUPf90 package [19]. The statistical model in matrix notation was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zg} + \mathbf{e} \text{ [model VC_nHS]}$$

where \mathbf{y} = a vector of observations for the test-day or the lactation records; \mathbf{b} = a vector of fixed effects including herd, calving year, calving month, age at first calving classes, lactation-calving-age classes for dams, classes for days in milk (for the test-day traits), lactation length classes

Table 2. Descriptive statistics of the weekly temperature humidity index before birth.

Week ^a	Mean	Minimum	Maximum	Standard deviation
1	47.88	11.47	72.44	12.10
2	47.83	11.80	73.10	12.22
3	47.84	10.43	72.52	12.16
4	47.91	14.02	72.62	12.10
5	48.01	10.43	72.64	11.95
6	48.08	11.76	72.62	11.99
7	48.10	11.76	72.33	12.03
8	48.17	14.40	73.10	11.99

^a: week before birth.

<https://doi.org/10.1371/journal.pone.0258216.t002>

(for lactation traits) and a dummy variable indicating the presence of HS during 1 to 8 weeks before birth for the genotyped cows; \mathbf{g} = a vector of random additive genetic effects following $N(0, \mathbf{G}\sigma_g^2)$, where \mathbf{G} = the genomic relationship matrix constructed according to VanRaden [20] and σ_g^2 = the genomic variance; and \mathbf{e} = a vector of random residuals following $N(0, \mathbf{I}\sigma_e^2)$, where \mathbf{I} = an identity matrix and σ_e^2 = residual variance. \mathbf{X} and \mathbf{Z} were incidence matrices for \mathbf{b} and \mathbf{g} , respectively.

Age at first calving was grouped into 6 classes (class 1: ≤ 22 months, class 2: 23–24 months, class 3: 25–26 months, class 4: 27–28 months, class 5: 29–30 months, and class 6: ≥ 31 months). Calving age and lactation (L) of the dams were combined to form 21 classes (class 1–6: the same as criteria for age at first calving in L1, class 7: ≤ 35 months in L2, class 8: 36–37 months in L2, class 9: 38–39 months in L2, class 10: 40–42 months in L2, class 11: ≥ 43 months in L2, class 12: ≤ 48 months in L3, class 13: 49–51 months in L3, class 14: 52–54 months in L3, class 15: ≥ 55 months in L3, class 16: ≤ 61 months in L4, class 17: 62–66 months in L4, class 18: ≥ 67 months in L4, class 19: ≤ 77 months in L5, class 20: ≥ 78 months in L5, and class 21: dams in L6). For DIM on the first test-day, 6 classes were defined: class 1: ≤ 10 days, class 2: 11–15 days, class 3: 16–20 days, class 4: 21–25 days, class 5: 26–30 days, and class 6: ≥ 31 days. Lactation length for lactation production traits was classified into 16 classes considering equal intervals between 275 to 305 days. The dummy variable indicating in-utero HS was 1 for the respective week for $\text{THI} \geq 60$, otherwise, a 0 was assigned. We performed separate runs for each week before calving.

Alternatively, an interaction model (as introduced by Yao et al. [23]) considering interactions between genotype and in-utero HS, was defined:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{g} + \mathbf{W}\mathbf{g}_{\text{hs}} + \mathbf{e} \text{ [model VC_wHS]}$$

where \mathbf{g}_{hs} = a vector of genotype by HS interaction effects for cows with in-utero HS following $N(0, \mathbf{G}_{\text{hs}}\sigma_{\text{g}_{\text{hs}}}^2)$, where \mathbf{G}_{hs} = the genomic relationship matrix for the cows with in-utero HS and $\sigma_{\text{g}_{\text{hs}}}^2$ = the variance of GxE; \mathbf{W} = a design matrix allocating phenotypic records to \mathbf{g}_{hs} . The remaining effects are the same as described in model VC_nHS.

The genomic heritability (h_g^2) for each trait was calculated as $h_g^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$ for estimates from model VC_nHS. For estimates from model VC_wHS, the heritability of the common genomic effects (h_c^2) was $h_c^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{\text{g}_{\text{hs}}}^2 + \sigma_e^2)$ and the ratio of the variance for genotype by HS interaction effects (r_{hs}) was $r_{\text{hs}} = \sigma_{\text{g}_{\text{hs}}}^2 / (\sigma_g^2 + \sigma_{\text{g}_{\text{hs}}}^2 + \sigma_e^2)$.

Genome wide associations. Subsequently, we estimated main and interaction effects for every SNP via generalized least squares (GLS) equations according to the algorithm as introduced by Yang et al. [24]. The models to estimate main and interaction SNP effects were:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{x}_{\text{snpi}}b_{\text{snpi}} + \mathbf{x}_{\text{interi}}b_{\text{interi}} + \mathbf{Z}\mathbf{g} + \mathbf{e} \text{ [model GWA_nHS]}$$

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{x}_{\text{snpi}}b_{\text{snpi}} + \mathbf{x}_{\text{interi}}b_{\text{interi}} + \mathbf{Z}\mathbf{g} + \mathbf{W}\mathbf{g}_{\text{hs}} + \mathbf{e} \text{ [model GWA_wHS]}$$

For solving the equations of model GWA_nHS, we considered the variance components estimated from model VC_nHS (algorithm is specified below). For solving the equations of model GWA_wHS, we considered the variance components estimated from model VC_wHS (algorithm is specified below).

Matrices and vectors of models for genome wide associations were as follows: \mathbf{x}_{snpi} = a vector of centered genotypes calculated as $\mathbf{m}_{\text{snpi}} - 2\mathbf{p}_{\text{snpi}}$ (\mathbf{m}_{snpi} = a vector of genotypes for marker i considering all genotyped animals; \mathbf{p}_{snpi} = a vector of allele frequency for marker i); b_{snpi} = a regression coefficient for the i^{th} SNP (the main SNP effect); $\mathbf{x}_{\text{interi}}$ = a vector of centered

genotypes for cows undergoing in-utero HS in the respective week before birth (consecutive runs for the 8 different weeks); and b_{interi} = a regression coefficient for the i^{th} SNP (the interaction effect) under in-utero HS. The remaining effects, vectors and matrices were identical with the effects as specified for models VC_nHS and VC_wHS.

For GWAS models, $E(\mathbf{y}) = \mathbf{X}\mathbf{b} + \mathbf{x}_{snpi}b_{snpi} + \mathbf{x}_{interi}b_{interi}$. $\text{Var}(\mathbf{y})$ equals to $\text{Var}(\mathbf{Z}\mathbf{g} + \mathbf{e}) = \mathbf{Z}\mathbf{G}\sigma_g^2\mathbf{Z}' + \mathbf{I}\sigma_e^2 = \mathbf{G}\sigma_g^2 + \mathbf{I}\sigma_e^2$ for model GWA_nHS and to $\text{Var}(\mathbf{Z}\mathbf{g} + \mathbf{W}\mathbf{g}_{hs} + \mathbf{e}) = \mathbf{Z}\mathbf{G}\sigma_g^2\mathbf{Z}' + \mathbf{W}\mathbf{G}_{hs}\sigma_{ghs}^2\mathbf{W}' + \mathbf{I}\sigma_e^2 = \mathbf{G}\sigma_g^2 + \mathbf{W}\mathbf{G}_{hs}\sigma_{ghs}^2\mathbf{W}' + \mathbf{I}\sigma_e^2$ for model GWA_wHS. \mathbf{Z} was an identity matrix, because all phenotyped animals were genotyped.

In a self-written R program “GWAInter.R” (in S1 Appendix), we applied GLS to estimate the main and interaction effects for each marker, implying a loop with 41,139 repetitions for all SNP markers. The detailed procedure was:

1. Setting $\mathbf{b}_1 = [\mathbf{b} \ b_{snpi} \ b_{interi}]$ and $\mathbf{X}_1 = [\mathbf{X} \ \mathbf{x}_{snpi} \ \mathbf{x}_{interi}]$, then $\widehat{\mathbf{b}}_1 = (\mathbf{X}'_1\mathbf{V}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{V}^{-1}\mathbf{y}$ (the last two fixed effects in the solution from GLS were main SNP and interaction effects for marker i , respectively) and $\text{var}(\widehat{\mathbf{b}}_1) = (\mathbf{X}'_1\mathbf{V}^{-1}\mathbf{X}_1)^{-1}$, $\mathbf{V} = \mathbf{G}\sigma_g^2 + \mathbf{I}\sigma_e^2$ or $\mathbf{V} = \mathbf{G}\sigma_g^2 + \mathbf{W}\mathbf{G}_{hs}\sigma_{ghs}^2\mathbf{W}' + \mathbf{I}\sigma_e^2$ (the last two variances were the variances for main SNP and interaction effects for marker i , respectively);
2. The test statistic for the i^{th} main SNP effect was: $\chi_{snpi}^2 = \frac{\widehat{b_{snpi}}^2}{\text{var}(b_{snpi})}$, with 1 degree of freedom (df);
3. The test statistic for the i^{th} interaction effect was: $\chi_{interi}^2 = \frac{\widehat{b_{interi}}^2}{\text{var}(b_{interi})}$, with 1 df.

The inflation factor (λ) was calculated based on the chi-squared statistics as:

$\widehat{\lambda} = \frac{\text{median}(\chi_i^2)}{\chi_{0.5,df=1}^2}$, where $\chi_i^2 = \chi_{snpi}^2$ or χ_{interi}^2 , and $\chi_{0.5,df=1}^2 = 0.4549$ (the statistic for a probability of 0.5 from a chi-squared distribution with 1 df). The P -value of the main and interaction effects for each SNP were determined by χ_{snpi}^2 and χ_{interi}^2 , respectively. Significantly associated SNPs were detected according to the Bonferroni correction calculated as $P_{Bonf} = 0.05/(\text{number of SNPs}) = 0.05/41,139 = 1.22 \times 10^{-6}$ ($-\log_{10}(P_{Bonf}) = 5.92$). Additionally, a less stringent correction was applied and defined as suggestive threshold with $P_{sugg} = 0.05/(\text{number of independent SNPs}) = 0.05/3873 = 4.76 \times 10^{-6}$ ($-\log_{10}(P_{sugg}) = 4.89$). The number of independent SNPs was calculated based on restrictions for linkage disequilibrium ($R^2 \leq 0.15$) for consecutive genomic windows including 500 SNPs (calculated in PLINK 2.0 [25, 26]).

Gene annotation

Suggestive SNPs according to P_{sugg} were annotated to potential candidate genes as listed in the Ensembl genome database [27] for main and interaction SNP effects. In this regard, we used the BiomaRt R package [28, 29]. Only genes located within a window of ± 100 kb of suggestive SNP were considered as potential candidate genes. Afterwards, we submitted the identified potential candidate genes to gene ontology (GO) overrepresentation tests [30], using the GO web-tool [31, 32]. The false discovery rate of $P < 0.05$ was considered to identify overrepresented GO terms for biological processes and reactome pathways. In addition, the windows with suggestive SNPs were mapped with the bovine QTL database through the online Data Analysis Tools [33] to elucidate the phenotypic contributions of the genomic segments. According to Hu et al. [33], the phenotypes were concisely

classified into milk, meat and carcass, growth performance, reproduction, exterior and health trait categories.

Results and discussion

Variance components

The genomic heritabilities from a model without consideration of in-utero HS (model VC_nHS) as well as common genomic, interaction and residual variances estimated from a model with in-utero HS effects (model VC_wHS) are plotted in Fig 1. Basically, common genomic and residual variances were quite constant for all studied weeks during late gestation. Moreover, both variance components from model VC_wHS were approximately the same when compared to the respective estimates from model VC_nHS. Hence, the genomic heritabilities from model VC_nHS and the common genomic heritabilities from model VC_wHS were 0.13 for TMY, 0.19 for TFP, 0.15 for TFY, 0.24 for TPP, 0.10 for TPY, and 0.08 for TSCS.

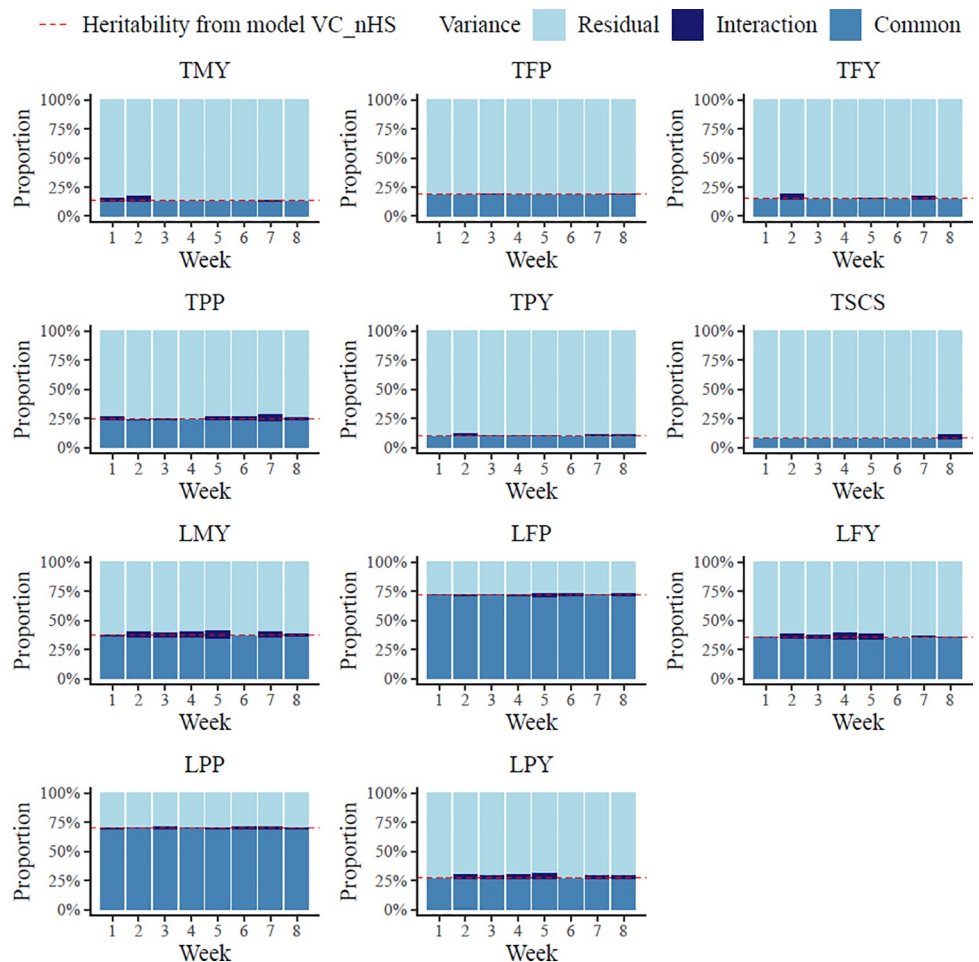


Fig 1. Common genomic, interaction and residual variances of milk production traits and somatic cell score estimated from an animal model with genotype by heat stress interactions during the last eight weeks of pregnancy (model VC_wHS). Dotted line = heritability estimated from an animal model without interaction (model VC_nHS); TMY = first test-day milk yield; TFP = first test-day fat percentage; TFY = first test-day fat yield; TPP = first test-day protein percentage; TPY = first test-day protein yield; TSCS = first test-day somatic cell score; LMY = first lactation milk yield; LFP = first lactation fat percentage; LFY = first lactation fat yield; LPP = first lactation protein percentage; LPY = first lactation protein yield.

<https://doi.org/10.1371/journal.pone.0258216.g001>

Minor differences in heritability estimates from model VC_nHS and VC_wHS were observed for lactation production traits. In detail, the h_c^2 from model VC_wHS ranged between 0.35 and 0.37 for LMY, between 0.70 and 0.72 for LFP, between 0.34 and 0.36 for LFY, between 0.69 and 0.70 for LPP, and between 0.26 and 0.28 for LPY. The h_g^2 from model VC_nHS for the same traits were 0.37, 0.72, 0.36, 0.71, and 0.28, respectively. However, interaction variances varied to a large extent when HS during last eight weeks of pregnancy was considered in consecutive runs, leading to diverse in-utero HS ratios across the eight weeks. Among all traits, LMY displayed the largest r_{hs} in a range from 0.00 (WK6) to 0.07 (WK5). The r_{hs} estimates for TFP were consistently very close to zero in response to climatic impact from all eight last weeks of gestation. For all traits, the interaction variances explained small proportions of the total phenotypic variance. In contrast to small interaction variances reported in our study, environment-specific genomic variances for feed efficiency traits [23] were larger than the common genomic variances from the present studies. However, in this study [23], production and management environments different substantially, because datasets from three countries located in two different continents were merged. Herd specific effects explained very large proportions up to 30% of phenotypic variations in cow production traits in North America [34]. The maximum r_{hs} of 0.07 in our study indicate that the effect of in-utero HS during late pregnancy is smaller than the herd effect due to geographical impact across latitudes. The time period from in-utero HS measurements until the trait recording date is longer for lactation records, but the r_{hs} for lactation production traits were slightly larger than for the respective test-day traits. Nevertheless, the stronger impacts might be due to the accumulative in-utero HS effects during the whole lactation.

In general, regarding in-utero HS, WK2 was most important for test-day yield traits, i.e., TMY ($r_{hs} = 0.05$), TFY ($r_{hs} = 0.05$) and TPY ($r_{hs} = 0.03$). For lactation yield traits, WK4 and WK5 displayed the largest r_{hs} , with 0.06 and 0.07 for LMY, 0.06 and 0.05 for LFY, and 0.05 for LPY.

Inflation factors

The inflation factors for the main effects from model VC_nHS and VC_wHS ranged from 0.74 for LFP to 1.05 for TSCS (S1 Fig). An inflation factor of value 1 represents sufficient correction for the population structure in GWAS, and a value larger than 1.05 indicates overestimation with inflated false positive results [35]. Consideration of the interaction variance in model GWA_wHS had no impact on inflation factors for the main SNP effects (when compared to inflation factors from model GWA_nHS). However, a larger number of values exceeding 1.05 were observed for the SNP interaction effects from model GWA_nHS (Fig 2). Generally, with regard to model GWA_nHS, the inflation factors indicated more false positives for lactation compared to test-day traits. For test-day traits, less than 30% of all runs had a $\lambda > 1.05$, but 67.5% of all runs for lactation milk-production traits displayed such an increased inflation factor. Interestingly, when considering the interaction variances for in-utero HS in model GWA_wHS, inflated results decreased to 12.5% for test-day traits and to 5% for lactation traits. Due to the improvements from model GWA_wHS (i.e., the preferable inflation factors), only significant and suggestive SNPs for main and interaction effects from model GWA_wHS were considered in the ongoing gene annotations.

Significant and suggestive SNPs for main effects

For the main effects, suggestive SNPs from model GWA_nHS overlapped to a large extent with suggestive SNPs from model GWA_wHS. Therefore, only results from GWA_wHS are presented. A total of 380 suggestive SNPs and 237 significant SNPs (including repetitive SNPs

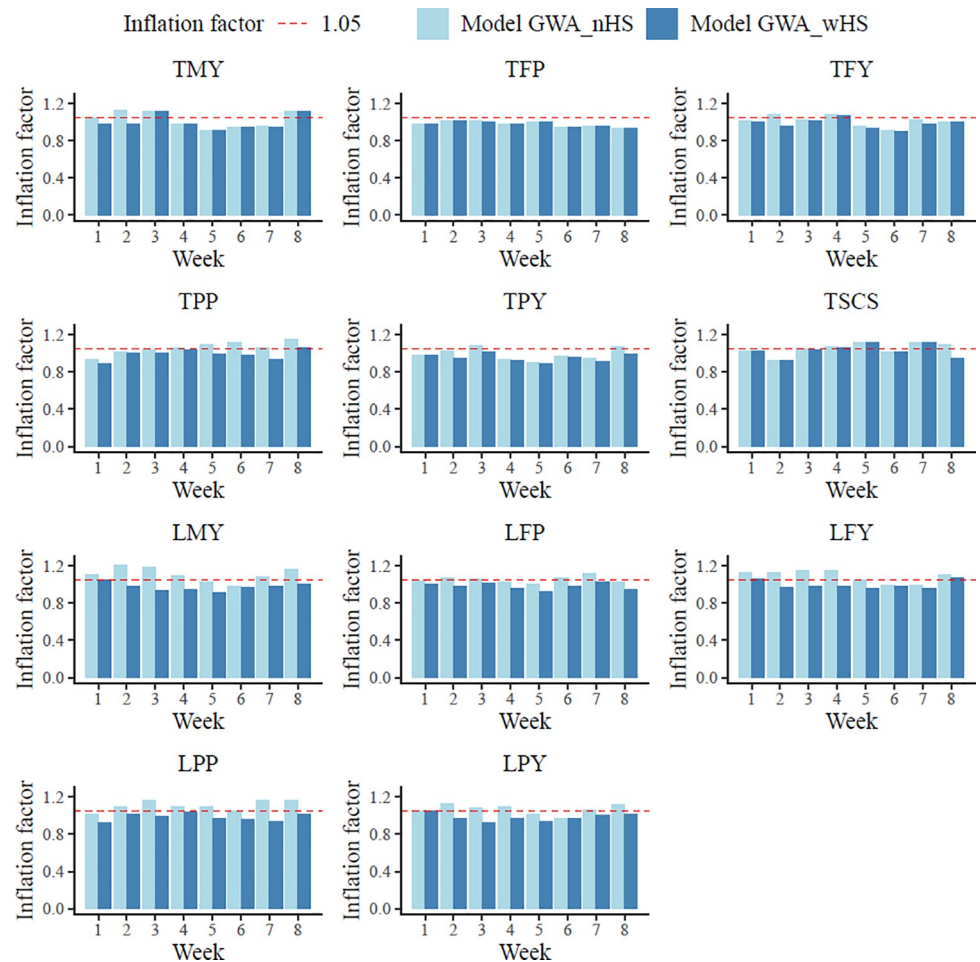


Fig 2. Inflation factors for interaction effects from GWAS models with (GWA_wHS) and without genotype by heat stress interaction (GWA_nHS) during the last eight weeks of pregnancy. Dotted line = inflation factor of 1.05; TMY = first test-day milk yield; TFP = first test-day fat percentage; TFY = first test-day fat yield; TPP = first test-day protein percentage; TPY = first test-day protein yield; TSCS = first test-day somatic cell score; LMY = first lactation milk yield; LFP = first lactation fat percentage; LFY = first lactation fat yield; LPP = first lactation protein percentage; LPY = first lactation protein yield.

<https://doi.org/10.1371/journal.pone.0258216.g002>

for different traits) exceeded the respective thresholds (Table 3). Number of suggestive and significant SNPs equaled to 171 and 98, respectively, when considering only unique SNPs. Detailed information for the suggestive SNPs is given in S1 Table. Among the eleven analyzed traits, LPP displayed the largest number of suggestive SNPs (91), while only three suggestive SNPs were detected for TSCS, indicating the strong polygenic genetic architecture of TSCS compared to milk production traits [36]. For both test-day and lactation records, results for fat and protein percentages indicated a likewise oligo-genic genetic architecture, as previously reported in genomic studies comparing percentage and yield traits [37]. Hence, the number of suggestive SNPs for milk composition traits including TFP, TPP, LFP, and LPP were always larger than for the respective yield traits TFY, TPY, LFY and LPY. It is well known that fat and protein percentages are controlled by some major genes with large effects, and many unknown genes with small effects. In this regard, the gene *DGAT1* on BTA14 strongly contributed to milk fat content [38], and the bovine *ABCG2* gene on BTA6 had a major effect on milk protein content [39]. Accordingly, BTA14 comprised the most suggestive SNPs for all traits, apart

Table 3. Number of SNPs with suggestive main effects associated with milk production traits and somatic cell score.

Trait ^a	#sug. SNPs	Chromosome (#sug. SNPs on chromosome)
TMY	9	6 (3); 14 (4); 16 (1); 20 (1)
TFP	51	3 (1); 14 (40); 27 (10)
TFY	27	6 (2); 14 (23); 20 (1); 27 (1)
TPP	30	1 (1); 3 (3); 6 (9); 10 (1); 14 (9); 20 (1); 23 (6)
TPY	5	9 (1); 10 (1); 11 (2); 20 (1)
TSCS	3	5 (1); 24 (2)
LMY	34	5 (4); 6(5); 14 (25)
LFP	72	3 (1); 5 (24); 14 (43); 20 (4)
LFY	37	11 (1); 14 (33); 15 (1); 18 (1); 26 (1)
LPP	91	3 (6); 5 (9); 6 (9); 10 (7); 11 (3); 14 (38); 15 (2); 20 (12); 21 (1); 25 (1); 29 (3)
LPY	21	6 (8); 11 (2); 14 (9); 20 (1); 24 (1)

^a. TMY = first test-day milk yield; TFP = first test-day fat percentage; TFY = first test-day fat yield; TPP = first test-day protein percentage; TPY = first test-day protein yield; TSCS = first test-day somatic cell score; LMY = first lactation milk yield; LFP = first lactation fat percentage; LFY = first lactation fat yield; LPP = first lactation protein percentage; LPY = first lactation protein yield.

<https://doi.org/10.1371/journal.pone.0258216.t003>

from TPY and TSCS. SNPs on BTA6 were suggestively associated with six traits, including TMY, TFY, TPP, LMY, LPP and LPY. A larger number of SNPs were detected for lactation than for corresponding test-day milk production traits, reflecting the larger variation for accumulated lactation than for single test-day traits.

In general, for milk production traits, the identified suggestive SNPs for main effects, in particular associated SNPs on BTA3, 5, 6, 10, 14, and 20, were in agreement with the associated SNPs as reported in previous studies [40, 41]. A quite large proportion (66.67%) of the suggestively associated SNPs was mapped by significant SNPs influencing milk production traits [40, 41]. For TSCS, Meredith et al. [40, 42] identified dozens of SNPs, but only one QTL region on BTA24 (spanning from 59.67 Mb to 59.76 Mb) was in close proximity to the SNP detected in our study. Meredith et al. [42] used daughter yield deviations for TSCS as phenotypes for sires in their studies. When considering original phenotypes in a cow population as dependent traits, the number of identified SNPs for TSCS decreased to nine [41].

Significant and suggestive SNPs for interaction effects

With regard to interaction effects, model GWA_nHS generated more suggestive SNPs (51) than model GWA_wHS. This was especially the case for lactation production traits, because the inflation factors for the interaction effects of lactation traits substantially decreased when switching from GWA_nHS to GWA_wHS (Fig 2). The suggestive SNPs from model GWA_nHS contained all suggestive SNPs from model GWA_wHS. Due to the improved inflation factors from GWA_wHS, only suggestive SNPs from GWA_wHS were considered in ongoing gene annotations. The number of suggestive and significant SNPs (31 and 4) for the interaction effects was smaller than for the main effects (Table 4), reflecting the variance proportions for the interaction and common genomic effects, respectively. In contrast to main effects, the number of suggestive and significant SNPs for the interaction effects (estimations from model GWA_wHS) was larger for test-day than for lactation traits. Three from the 31 suggestively associated SNP (model GWA_wHS) were associated with functional traits in Holstein cattle, including *ARS-BFGL-NGS-88748* (at 61.48 Mb on BTA19) with body size [43],

Table 4. SNPs with suggestive interaction effects on first test-day milk yield (TMY), fat percentage (TFP), fat yield (TFY), protein percentage (TPP), protein yield (TPY) and somatic cell score (TSCS), and on lactation milk yield (LMY), lactation fat percentage (LFP), lactation fat yield (LFY), lactation protein percentage (LPP) and lactation protein yield (LPY).

Trait	#sug. SNPs	SNP	Week	Chr.	Base pair	MAF	Beta.inter	SE	P-value
TMY	2	ARS-BFGL-BAC-33086	8	19	14537900	0.27	-0.91	0.19	1.47E-06
		ARS-BFGL-NGS-88748	7	19	61476059	0.18	-1.00	0.22	7.99E-06
TFP	7	ARS-BFGL-NGS-3794	5	1	89914813	0.13	0.15	0.03	1.37E-06
		Hapmap39006-BTA-116188	1	2	30384603	0.35	-0.10	0.02	6.49E-06
		Hapmap47201-BTA-121648	2	4	91525544	0.39	-0.09	0.02	1.11E-05
		Hapmap22875-BTA-155031	2	4	112899067	0.25	0.11	0.02	3.49E-06
		Hapmap44288-BTA-81147	5,7	8	48282567	0.19	-0.11	0.03	2.61E-07
		ARS-BFGL-NGS-17797	5	18	51967260	0.15	0.14	0.03	1.33E-06
		ARS-BFGL-NGS-17754	5	18	52115086	0.15	0.14	0.03	1.33E-06
TFY	3	Hapmap51787-BTA-101971	8	2	54937229	0.23	0.04	0.01	6.09E-06
		Hapmap26991-BTA-151076	8	2	82959890	0.44	0.04	0.01	1.33E-06
		BTA-104962-no-rs	4	6	8553220	0.16	-0.05	0.01	4.31E-06
TPP	6	ARS-BFGL-NGS-3821	3	1	60736616	0.32	-0.04	0.01	7.37E-07
		BTB-01195369	8	3	90945010	0.39	-0.04	0.01	7.09E-06
		ARS-BFGL-NGS-73518	8	3	90978652	0.37	0.04	0.01	7.70E-06
		ARS-BFGL-NGS-34063	8	5	118085231	0.33	0.04	0.01	4.97E-06
		ARS-BFGL-NGS-62753	3	12	66123957	0.11	0.06	0.01	7.69E-06
		Hapmap33939-BES5_Contig460_1314	1	24	44888957	0.29	0.04	0.01	9.45E-06
TPY	3	ARS-BFGL-NGS-65689	2	8	23528913	0.08	-0.05	0.01	2.00E-06
		Hapmap51393-BTA-113111	2	8	23814719	0.07	0.05	0.01	4.73E-06
		ARS-BFGL-BAC-33086	8	19	14537900	0.27	-0.03	0.01	6.19E-06
TSCS	4	Hapmap38434-BTA-74956	6	5	106953248	0.19	-0.29	0.06	5.02E-06
		ARS-BFGL-NGS-4427	6	11	81840635	0.14	0.32	0.07	5.48E-06
		ARS-BFGL-NGS-17721	6	11	82176836	0.16	0.30	0.07	1.02E-05
		ARS-BFGL-NGS-20982	1	27	23032376	0.30	0.30	0.05	1.93E-08
LMY	3	ARS-BFGL-NGS-103742	7,8	10	40203321	0.07	308.38	69.72	9.33E-07
		ARS-BFGL-NGS-106479	4	11	79750751	0.19	-230.38	48.15	1.72E-06
		ARS-BFGL-NGS-54155	8	27	7920380	0.37	168.80	38.47	1.15E-05
LFP	1	BTB-01985049	3	9	21565055	0.11	-0.07	0.02	7.47E-06
LFY	1	ARS-BFGL-NGS-53850	8	21	28216224	0.06	11.58	2.62	1.02E-05
LPP	1	BTB-01130157	4	7	56381886	0.25	-0.02	0.00	1.28E-05
LPY	3	ARS-BFGL-NGS-103742	8	10	40203321	0.07	10.16	2.09	1.24E-06
		ARS-BFGL-NGS-54155	8	27	7920380	0.37	5.11	1.16	1.07E-05
		ARS-BFGL-NGS-34282	5	29	33704537	0.41	5.04	1.13	8.03E-06

Week = week before birth; SNP = SNP name; Chr. = chromosome; MAF = minor allele frequency; Beta.inter = interaction effect; SE = standard error.

Bold = SNP with suggestive interaction effects for two traits.

<https://doi.org/10.1371/journal.pone.0258216.t004>

BTB-01195369 (at 90.95 Mb on BTA3) with endocrine fertility traits [44], and *Hapmap51393-BTA-113111* (at 23.81 Mb on BTA8) with cow conception rate [45]. Another three SNPs, i.e., *ARS-BFGL-NGS-17754* (at 52.12 Mb on BTA18), *ARS-BFGL-NGS-73518* (at 90.98 Mb on BTA3) and *ARS-BFGL-NGS-34063* (at 118.09 Mb on BTA5), had an effect on weight, height at withers, and weaning weight, respectively, in Hanwoo cattle [46]. *Hapmap22875-BTA-155031* (at 112.90 Mb on BTA4) was significantly associated with tick resistance in Brazilian Braford and Hereford cattle [47] and *ARS-BFGL-NGS-106479* (at 79.75 Mb on BTA11) with intramuscular fat percentage in Australian beef cattle [48]. Furthermore, the

Holstein haplotype of the gene *APOB*, detected as a causal mutation for cholesterol deficiency, embraces the SNP *ARS-BFGL-NGS-106479* [49]. Interestingly, for the interaction effects, the eight SNPs as reported in the aforementioned studies are more relevant for functional than for primary traits in cattle, although our analyses mainly focused on milk production. Results for the suggestive SNP interactions and their associated traits indicate that robust cattle with superior functionality, e.g., favorable fertility and conformation, might be more tolerant to HS during late pregnancy than cattle with breeding focus on milk production. Mapping all suggestive SNPs for interaction effects with the bovine QTL database revealed that the two most important QTL regions were associated with TPP and TPY.

Potential candidate genes

For test-day milk production traits, a total of 373 unique potential candidate genes are located in regions spanning 100 kb upstream to downstream of suggestive main SNPs from model GWA_wHS (S1 Table). In addition to the *DGAT1* gene, the annotated genes *CPSF1*, *FOXH1*, *ARHGAP39*, *PPP1R16A*, *GRINA* and *MROH1* were frequently associated with milk production traits [41, 50, 51]. Because of the quite large genetic correlations (≥ 0.75 ; in S2 Table) between test-day and corresponding lactation traits, the percentage of overlapping candidate genes was larger than 78%. Interestingly, the genes *GHR* and *ABCG2* were only detected as candidate genes for lactation, but not for test-day milk production traits. Only three candidate genes were located in close distance to suggestive SNPs for TSCS. However, no study reported a relationship between the three genes and TSCS. The gene *ABCC9* harbors the most significant SNPs for production traits [41], the gene *WDR7* is associated with sperm motility and concentration in Holstein-Friesian bulls [52], and the gene *TXNL1* has an effect on moisture content in rainbow trout [53]. The response to oxygen-containing compound and the G-protein-coupled receptors signaling pathway were significantly detected based on all candidate genes for main effects. The first pathway represents biological processes inducing changes in cell or an organism activities (in terms of movement, secretion, enzyme production or gene expression), due to an oxygen-containing compound stimulus [54]. The second pathway recognizes a broad array of extracellular mediators including cationic amines, lipids, peptides, proteins and sensory agents [55].

Due to the limited number of suggestive SNPs for interaction effects from model GWA_wHS, a smaller number including 62 unique candidate genes in the vicinity of the SNP locations, were identified (± 100 Kb, S3 Table). In human, the members of *ABCA6*-like transporters, *ABCA6*, *ABCA8*, *ABCA9* and *ABCA10*, are all integrated in cholesterol-related pathways, either directly or based on their dynamic regulation upon cholesterol application [56]. Chemokines, including *CCL14* and *CCL16*, play important roles in regulating bovine endometrial functions during early pregnancy [57]. The genes *CCL5*, *HEATR9*, *MMP28*, *GIMAP* family and *VEGFC* activate immune responses [58–62]. Proteins encoded by the genes *GAS2L2* and *NRIP2* participate in growth process [63, 64]. Mutations in the *SCN1A* and *TTC21B* genes cause several diseases [65, 66]. Sanglard et al. [67] reported a strong correlation between *ETNPPL* variants and the top 20 differentially expressed genes in beef calves in response to energy restriction during late gestation. *GPC4* from the same family of the identified *GPC5* and the directly identified gene *RASL10B* were favorably associated with *ETNPPL* [67]. The genes *LSAMP* and *GRM8* are involved in emotional and motivational functions [68, 69]. The *MDGA2* gene encodes novel proteins which regulate neuronal migration and neural development [70]. The *GPC5* gene was identified by Naderi et al. [71], who focused on selection signatures in subpopulations of local dual-purpose black and white cattle from Germany. As a result from another selection signature analyses, the gene *FOCAD* was identified by Alshawi et al.

[72] in Iraqi cattle. The gene *C8H9orf85* was detected as a candidate gene for milk production traits in water buffalo [73]. The genes *TULP3*, *IQSEC3* and *FAM49A* were identified as potential candidate genes for milk production traits in US Holstein cattle [41] and for milking speed in French Holstein cattle [74]. The genes *RHNO1* and *FOXMI* are linked to a variety of cancers and are defined as DNA damage repair regulators [75]. Demetriou et al. [76] explained the role of the *FKBP4* gene in recurrent fetal losses in humans. Bourneuf et al. [77] identified *de novo* deleterious mutations in the *FAM189A1* gene. The zinc-finger proteins constitute one of the most abundant groups of proteins in the mammalian genome and are involved in several cellular processes, differentiations of several tissues, development of several diseases, as well as in tumorigenesis, cancer progression and metastasis formation [78].

From the overrepresentation test, two biological processes, including lymphocyte chemotaxis and monocyte chemotaxis, were excessively represented, referring to all genes from the bovine genome. Lymphocytes contribute to the development of immunity and to allergic inflammation of the lower respiratory tract [79]. Monocytes participate in both innate and adaptive immune responses, because of their phagocytic activity [80]. The identified annotations indicate the biological complexity of in-utero HS on productivity in offspring. Nevertheless, sensitivity to in-utero HS in offspring seems to be correlated with the health status of their dams.

In conclusion, the r_{hs} in the range from 0 to 0.07 indicate quite small effects of HS during late pregnancy on test-day and lactation milk production traits. The accumulative effects due to in-utero HS may explain the generally larger interaction variances for lactation compared to test-day traits. Superiority in inflation factors for the SNP interaction effects from the model GWA_wHS over the model GWA_nHS suggests modelling of interaction variances when applying a GWAS with GxE interactions. Suggestive SNPs for the main effects and their corresponding annotated potential candidate genes reflect findings from previous GWAS focussing on milk production traits. For the SNP interaction effects, eight out of 31 suggestive SNPs are known to effect cattle functionality, indicating a possible favourable correlation between functional traits and heat tolerance. Furthermore, two biological processes as inferred on the basis of the identified 62 candidate genes located in close proximity to the SNPs with suggestive interaction effects, contribute to immune response mechanisms.

Supporting information

S1 Table. Suggestive SNPs and potential candidate genes associated with main effects for test-day and lactation production traits.

(XLSX)

S2 Table. Genetic correlations between test-day and lactation production traits.

(DOCX)

S3 Table. Potential candidate genes associated with interaction effects for test-day and lactation production traits. ^a: TMY = first test-day milk yield; TFP = first test-day fat percentage; TFY = first test-day fat yield; TPP = first test-day protein percentage; TPY = first test-day protein yield; TSCS = first test-day somatic cell score; LMY = first lactation milk.

(XLSX)

S1 Appendix. An R script to perform genome wide associations considering interaction between each marker and a discrete environmental descriptor with two classes.

(R)

S1 Fig. Inflation factors for main effects from GWAS models with (GWA_wHS) and without genotype by heat stress interaction (GWA_nHS) during the last eight weeks of pregnancy. Dotted line = inflation factor of 1.05; TMY = first test-day milk yield; TFP = first test-day fat percentage; TFY = first test-day fat yield; TPP = first test-day protein percentage; TPY = first test-day protein yield; TSCS = first test-day somatic cell score; LMY = first lactation milk yield; LFP = first lactation fat percentage; LFY = first lactation fat yield; LPP = first lactation protein percentage; LPY = first lactation protein yield. (TIF)

Author Contributions

Conceptualization: Tong Yin.

Formal analysis: Kathrin Halli.

Funding acquisition: Sven König.

Methodology: Seyi Fridaius Vanvanhossou.

Resources: Mehdi Bohlouli.

Software: Tong Yin.

Supervision: Tong Yin.

Writing – original draft: Kathrin Halli, Tong Yin.

Writing – review & editing: Sven König.

References

1. Lorenz R, Stalhandske Z, Fischer EM. Detection of a Climate Change Signal in Extreme Heat, Heat Stress, and Cold in Europe From Observations. *Geophys Res Lett*. 2019; 46:8363–8374. <https://doi.org/10.1029/2019GL082062>
2. Roenfeldt S. You can't afford to ignore heat stress. *Dairy Management*. 1998; 35:6–12.
3. Kadzere C, Murphy M, Silanikove N, Maltz E. Heat stress in lactating dairy cows: a review. *Livest Prod Sci*. 2002; 77:59–91. [https://doi.org/10.1016/S0301-6226\(01\)00330-X](https://doi.org/10.1016/S0301-6226(01)00330-X)
4. Bohmanova J, Misztal I, Cole JB. Temperature-humidity indices as indicators of milk production losses due to heat stress. *J Dairy Sci*. 2007; 90:1947–1956. <https://doi.org/10.3168/jds.2006-513> PMID: 17369235.
5. Gernand E, König S, Kipp C. Influence of on-farm measurements for heat stress indicators on dairy cow productivity, female fertility, and health. *J Dairy Sci*. 2019; 102:6660–6671. <https://doi.org/10.3168/jds.2018-16011> PMID: 31128870.
6. Bohlouli M, Alijani S, Naderi S, Yin T, König S. Prediction accuracies and genetic parameters for test-day traits from genomic and pedigree-based random regression models with or without heat stress interactions. *J Dairy Sci*. 2019; 102:488–502. <https://doi.org/10.3168/jds.2018-15329> PMID: 30343923.
7. Bernabucci U, Biffani S, Buggiotti L, Vitali A, Lacetera N, Nardone A. The effects of heat stress in Italian Holstein dairy cattle. *J Dairy Sci*. 2014; 97:471–486. <https://doi.org/10.3168/jds.2013-6611> PMID: 24210494.
8. Hayes BJ, Bowman PJ, Chamberlain AJ, Savin K, van Tassell CP, Sonstegard TS, et al. A validated genome wide association study to breed cattle adapted to an environment altered by climate change. *PLOS ONE*. 2009:e6676. <https://doi.org/10.1371/journal.pone.0006676> PMID: 19688089.
9. Dikmen S, Cole JB, Null DJ, Hansen PJ. Genome-wide association mapping for identification of quantitative trait loci for rectal temperature during heat stress in Holstein cattle. *PLOS ONE*. 2013:e69202. <https://doi.org/10.1371/journal.pone.0069202> PMID: 23935954.
10. Otto PI, Guimarães SEF, Verardo LL, Azevedo ALS, Vandenplas J, Sevilano CA, et al. Genome-wide association studies for heat stress response in *Bos taurus* × *Bos indicus* crossbred cattle. *J Dairy Sci*. 2019:8148–8158. <https://doi.org/10.3168/jds.2018-15305> PMID: 31279558.

11. de Andrade Ferrazza R, Mogollón Garcia HD, Vallejo Aristizábal VH, de Souza Nogueira C, Veríssimo CJ, Sartori JR, et al. Thermoregulatory responses of Holstein cows exposed to experimentally induced heat stress. *J Therm Biol.* 2017; 66:68–80. Epub 2017/03/29. <https://doi.org/10.1016/j.jtherbio.2017.03.014> PMID: 28477912.
12. Monteiro APA, Guo J-R, Weng X-S, Ahmed BM, Hayen MJ, Dahl GE, et al. Effect of maternal heat stress during the dry period on growth and metabolism of calves. *J Dairy Sci.* 2016; 99:3896–3907. <https://doi.org/10.3168/jds.2015-10699> PMID: 26947308.
13. Monteiro APA, Tao S, Thompson IMT, Dahl GE. In utero heat stress decreases calf survival and performance through the first lactation. *J Dairy Sci.* 2016:8443–8450. <https://doi.org/10.3168/jds.2016-11072> PMID: 27522427.
14. Tao S, Monteiro APA, Thompson IM, Hayen MJ, Dahl GE. Effect of late-gestation maternal heat stress on growth and immune function of dairy calves. *J Dairy Sci.* 2012:7128–7136. <https://doi.org/10.3168/jds.2012-5697> PMID: 23021751.
15. Tao S, Dahl GE. Invited review: heat stress effects during late gestation on dry cows and their calves. *J Dairy Sci.* 2013:4079–4093. <https://doi.org/10.3168/jds.2012-6278> PMID: 23664343.
16. Halli K, Brügemann K, Bohlouli M, Yin T, König S. Heat stress during late pregnancy and postpartum influences genetic parameter estimates for birth weight and weight gain in dual-purpose cattle offspring generations. *J Anim Sci.* 2021:Forthcoming. <https://doi.org/10.1093/jas/skab106> PMID: 33822077
17. Segelke D, Chen J, Liu Z, Reinhardt F, Thaller G, Reents R. Reliability of genomic prediction for German Holsteins using imputed genotypes from low-density chips. *J Dairy Sci.* 2012; 95:5403–5411. <https://doi.org/10.3168/jds.2012-5466> PMID: 22916947.
18. Aguilar I, Misztal I, Tsuruta S, Legarra A, Wang H. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 program. Proceedings of the 10th World Congress on Genetics Applied to Livestock Production. 2014:Vancouver, Canada.
19. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. BLUPF90 and related programs (BGF90). Proceedings of the 7th World Congress on Genetics Applied to Livestock Production. 2002:Montpellier, France.
20. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci.* 2008; 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980> PMID: 18946147.
21. Hijmans RJ, Williams E, Vennes C. Package 'geosphere'. 2019 [cited 2020 Nov 19]. Available from: <https://cran.r-project.org/web/packages/geosphere/vignettes/geosphere.pdf>.
22. NRC. A guide to environmental research on animals. National Academy Science, Washington, D.C., United States; 1971.
23. Yao C, de los Campos G, VandeHaar MJ, Spurlock DM, Armentano LE, Coffey M, et al. Use of genotype × environment interaction model to accommodate genetic heterogeneity for residual feed intake, dry matter intake, net energy in milk, and metabolic body weight in dairy cattle. *J Dairy Sci.* 2017; 100:2007–2016. <https://doi.org/10.3168/jds.2016-11606> PMID: 28109605.
24. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet.* 2014; 46:100–106. <https://doi.org/10.1038/ng.2876> PMID: 24473328.
25. Purcell S, Chang CC. PLINK 2.0. 2021 [Cited 2021 Jun 15]. Available from: www.cog-genomics.org/plink/2.0/.
26. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015; 4:7. Epub 2015/02/25. <https://doi.org/10.1186/s13742-015-0047-8> PMID: 25722852.
27. Howe KL, Achuthan P, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2021. *Nucleic Acids Res.* 2021; 49:D884–D891. <https://doi.org/10.1093/nar/gkaa942> PMID: 33137190.
28. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics.* 2005; 21:3439–3440. <https://doi.org/10.1093/bioinformatics/bti525> PMID: 16082012.
29. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc.* 2009; 4:1184–1191. <https://doi.org/10.1038/nprot.2009.97> PMID: 19617889.
30. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* 2019; 47: D419–D426. <https://doi.org/10.1093/nar/gky1038> PMID: 30407594.

31. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000; 25:25–29. <https://doi.org/10.1038/75556> PMID: 10802651.
32. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res.* 2021; 49:D325–D334. <https://doi.org/10.1093/nar/gkaa1113> PMID: 33290552.
33. Hu Z-L, Park CA, Reedy JM. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* 2019; 47:D701–D710. <https://doi.org/10.1093/nar/gky1084> PMID: 30407520.
34. Schultz NE, Weigel KA. Inclusion of herdmate data improves genomic prediction for milk-production and feed-efficiency traits within North American dairy herds. *J Dairy Sci.* 2019; 102:11081–11091. <https://doi.org/10.3168/jds.2019-16820> PMID: 31548069.
35. Power RA, Parkhill J, Oliveira T de. Microbial genome-wide association studies: lessons from human GWAS. *Nat Rev Genet.* 2017; 18:41–50. <https://doi.org/10.1038/nrg.2016.132> PMID: 27840430.
36. Abdel-Shafy H, Bortfeldt RH, Tetens J, Brockmann GA. Single nucleotide polymorphism and haplotype effects associated with somatic cell score in German Holstein cattle. *Genet Sel Evol.* 2014; 46:35. <https://doi.org/10.1186/1297-9686-46-35> PMID: 24898131.
37. Tiezzi F, Los Campos G de, Parker Gaddis KL, Maltecca C. Genotype by environment (climate) interaction improves genomic prediction for production traits in US Holstein cattle. *J Dairy Sci.* 2017; 100:2042–2056. <https://doi.org/10.3168/jds.2016-11543> PMID: 28109596.
38. Kühn C, Thaller G, Winter A, Bininda-Emonds ORP, Kaupé B, Erhardt G, et al. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. *Genetics.* 2004; 167:1873–1881. <https://doi.org/10.1534/genetics.103.022749> PMID: 15342525.
39. Cohen-Zinder M, Seroussi E, Larkin DM, Looor JJ, Everts-van der Wind A, Lee J-H, et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.* 2005; 15:936–944. <https://doi.org/10.1101/gr.3806705> PMID: 15998908.
40. Meredith BK, Kearney FJ, Finlay EK, Bradley DG, Fahey AG, Berry DP, et al. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genet.* 2012; 13:21. <https://doi.org/10.1186/1471-2156-13-21> PMID: 22449276.
41. Jiang J, Ma L, Prakapenka D, VanRaden PM, Cole JB, Da Y. A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle. *Front Genet.* 2019; 10:412. <https://doi.org/10.3389/fgene.2019.00412> PMID: 31139206.
42. Meredith BK, Berry DP, Kearney F, Finlay EK, Fahey AG, Bradley DG, et al. A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Front Genet.* 2013; 4:229. <https://doi.org/10.3389/fgene.2013.00229> PMID: 24223582.
43. Zhang X, Chu Q, Guo G, Dong G, Li X, Zhang Q, et al. Genome-wide association studies identified multiple genetic loci for body size at four growth stages in Chinese Holstein cattle. *PLOS ONE.* 2017; 12:e0175971. <https://doi.org/10.1371/journal.pone.0175971> PMID: 28426785.
44. Tenghe AMM, Bouwman AC, Berglund B, Strandberg E, Koning DJ de, Veerkamp RF. Genome-wide association study for endocrine fertility traits using single nucleotide polymorphism arrays and sequence variants in dairy cattle. *J Dairy Sci.* 2016; 99:5470–5485. <https://doi.org/10.3168/jds.2015-10533> PMID: 27157577.
45. Parker Gaddis KL, Null DJ, Cole JB. Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. *J Dairy Sci.* 2016; 99:6420–6435. <https://doi.org/10.3168/jds.2015-10444> PMID: 27209127.
46. Alam M, Lee Y-M, Park B-L, Kim J-H, Lee S-S, Shin H-D, et al. A Whole Genome Association Study to Detect Single Nucleotide Polymorphisms for Body Conformation Traits in a Hanwoo Population. *Asian-Australas J Anim Sci.* 2011; 24:322–329. <https://doi.org/10.5713/ajas.2011.11012>
47. Sollero BP, Junqueira VS, Gomes CCG, Caetano AR, Cardoso FF. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. *Genet Sel Evol.* 2017; 49:49. <https://doi.org/10.1186/s12711-017-0325-2> PMID: 28619006.
48. Bolormaa S, Neto LRP, Zhang YD, Bunch RJ, Harrison BE, Goddard ME, et al. A genome-wide association study of meat and carcass traits in Australian cattle. *J Anim Sci.* 2011; 89:2297–2309. <https://doi.org/10.2527/jas.2010-3138> PMID: 21421834.
49. Cole JB, Null DJ, VanRaden PM. Phenotypic and genetic effects of recessive haplotypes on yield, longevity, and fertility. *J Dairy Sci.* 2016; 99:7274–7288. <https://doi.org/10.3168/jds.2015-10777> PMID: 27394947.

50. Jiang L, Liu J, Sun D, Ma P, Ding X, Yu Y, et al. Genome wide association studies for milk production traits in Chinese Holstein population. *PLOS ONE*. 2010; 5:e13661. <https://doi.org/10.1371/journal.pone.0013661> PMID: 21048968.
51. Nayeri S, Sargolzaei M, Abo-Ismael MK, May N, Miller SP, Schenkel F, et al. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genet*. 2016; 17:75. <https://doi.org/10.1186/s12863-016-0386-1> PMID: 27287773.
52. Suchocki T, Szyda J. Genome-wide association study for semen production traits in Holstein-Friesian bulls. *J Dairy Sci*. 2015; 98:5774–5780. <https://doi.org/10.3168/jds.2014-8951> PMID: 26051317.
53. Ali A, Al-Tobasei R, Lourenco D, Leeds T, Kenney B, Salem M. Genome-wide scan for common variants associated with intramuscular fat and moisture content in rainbow trout. *BMC Genomics*. 2020; 21:529. <https://doi.org/10.1186/s12864-020-06932-0> PMID: 32736521.
54. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S. AmiGO: online access to ontology and annotation data. *Bioinformatics*. 2009; 25:288–289. Epub 2008/11/25. <https://doi.org/10.1093/bioinformatics/btn615> PMID: 19033274.
55. Gobeil F, Fortier A, Zhu T, Bossolasco M, Leduc M, Grandbois M, et al. G-protein-coupled receptors signalling at the cell nucleus: an emerging paradigm. *Can J Physiol Pharmacol*. 2006; 84:287–297. <https://doi.org/10.1139/y05-127> PMID: 16902576.
56. Kaminski WE, Wenzel JJ, Piehler A, Langmann T, Schmitz G. ABCA6, a novel a subclass ABC transporter. *Biochem Biophys Res Commun*. 2001; 285:1295–1301. <https://doi.org/10.1006/bbrc.2001.5326> PMID: 11478798.
57. Sakumoto R, Hayashi K-G, Fujii S, Kanahara H, Hosoe M, Furusawa T, et al. Possible Roles of CC- and CXC-Chemokines in Regulating Bovine Endometrial Function during Early Pregnancy. *Int J Mol Sci*. 2017; 18. <https://doi.org/10.3390/ijms18040742> PMID: 28362325.
58. Hussen J, Schubert H-J. Heterogeneity of Bovine Peripheral Blood Monocytes. *Front Immunol*. 2017; 8:1875. <https://doi.org/10.3389/fimmu.2017.01875> PMID: 29312348.
59. Stairiker CJ, van Meurs M, Leon LG, Brouwers-Haspels AA, Rijsbergen L, Mueller YM, et al. Heat9 is an infection responsive gene that affects cytokine production in alveolar epithelial cells. *PLOS ONE*. 2020; 15:e0236195. <https://doi.org/10.1371/journal.pone.0236195> PMID: 32678841.
60. Rodgers UR, Kevorkian L, SurrIDGE AK, Waters JG, Swingler TE, Culley K, et al. Expression and function of matrix metalloproteinase (MMP)-28. *Matrix Biol*. 2009; 28:263–272. <https://doi.org/10.1016/j.matbio.2009.04.006> PMID: 19375502.
61. Schwefel D, Arasu BS, Marino SF, Lamprecht B, Köchert K, Rosenbaum E, et al. Structural insights into the mechanism of GTPase activation in the GIMAP family. *Structure*. 2013; 21:550–559. <https://doi.org/10.1016/j.str.2013.01.014> PMID: 23454188.
62. Alitalo K. The lymphatic vasculature in disease. *Nat Med*. 2011; 17:1371–1380. <https://doi.org/10.1038/nm.2545> PMID: 22064427.
63. NCBI Gene Database. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information: 2004 [Cited 2021 Feb 26]. Available from: <https://www.ncbi.nlm.nih.gov/gene/246176>.
64. Cassar-Malek I, Boby C, Picard B, Reverter A, Hudson NJ. Molecular regulation of high muscle mass in developing Blonde d'Aquitaine cattle fetuses. *Biol Open*. 2017; 6:1483–1492. <https://doi.org/10.1242/bio.024950> PMID: 28838967.
65. Kwong AK-Y, Fung C-W, Chan S-Y, Wong VC-N. Identification of SCN1A and PCDH19 mutations in Chinese children with Dravet syndrome. *PLOS ONE*. 2012; 7:e41802. <https://doi.org/10.1371/journal.pone.0041802> PMID: 22848613.
66. Cong EH, Bizet AA, Boyer O, Woerner S, Gribouval O, Filhol E, et al. A homozygous missense mutation in the ciliary gene TTC21B causes familial FSGS. *J Am Soc Nephrol*. 2014; 25:2435–2443. <https://doi.org/10.1681/ASN.2013101126> PMID: 24876116.
67. Sanglard LP, Nascimento M, Moriel P, Sommer J, Ashwell M, Poore MH, et al. Impact of energy restriction during late gestation on the muscle and blood transcriptome of beef calves after preconditioning. *BMC Genomics*. 2018; 19:702. <https://doi.org/10.1186/s12864-018-5089-8> PMID: 30253751.
68. Heimer L, van Hoesen GW. The limbic lobe and its output channels: implications for emotional functions and adaptive behavior. *Neurosci Biobehav Rev*. 2006; 30:126–147. <https://doi.org/10.1016/j.neubiorev.2005.06.006> PMID: 16183121.
69. O'Rourke T, Boeckx C. Converging roles of glutamate receptors in domestication and prosociality. *bioRxiv*. 2019. <https://doi.org/10.1101/439869>
70. Litwack ED, Babey R, Buser R, Gesemann M, O'Leary DDM. Identification and characterization of two novel brain-derived immunoglobulin superfamily members with a unique structural organization. *Mol Cell Neurosci*. 2004; 25:263–274. <https://doi.org/10.1016/j.mcn.2003.10.016> PMID: 15019943.

71. Naderi S, Moradi MH, Farhadian M, Yin T, Jaeger M, Scheper C, et al. Assessing selection signatures within and between selected lines of dual-purpose black and white and German Holstein cattle. *Anim Genet.* 2020; 51:391–408. <https://doi.org/10.1111/age.12925> PMID: 32100321.
72. Alshawi A, Essa A, Al-Bayatti S, Hanotte O. Genome Analysis Reveals Genetic Admixture and Signature of Selection for Productivity and Environmental Traits in Iraqi Cattle. *Front Genet.* 2019; 10:609. <https://doi.org/10.3389/fgene.2019.00609> PMID: 31379916.
73. Du C, Deng T, Zhou Y, Ye T, Zhou Z, Zhang S, et al. Systematic analyses for candidate genes of milk production traits in water buffalo (*Bubalus Bubalis*). *Anim Genet.* 2019; 50:207–216. <https://doi.org/10.1111/age.12739> PMID: 30937948.
74. Marete A, Sahana G, Fritz S, Lefebvre R, Barbat A, Lund MS, et al. Genome-wide association study for milking speed in French Holstein cows. *J Dairy Sci.* 2018; 101:6205–6219. <https://doi.org/10.3168/jds.2017-14067> PMID: 29705414.
75. Barger CJ, Branick C, Chee L, Albahrani M, Klinkebiel D, Drapkin R, et al. Co-regulation and functional cooperativity of FOXM1 and RHNO1 bidirectional genes in ovarian cancer. *bioRxiv.* 2019:630442. <https://doi.org/10.1101/630442>
76. Demetriou C, Chanudet E, Joseph A, Topf M, Thomas AC, Bitner-Glindzicz M, et al. Exome sequencing identifies variants in FKBP4 that are associated with recurrent fetal loss in humans. *Hum Mol Genet.* 2019; 28:3466–3474. <https://doi.org/10.1093/hmg/ddz203> PMID: 31504499.
77. Bourneuf E, Otz P, Pausch H, Jagannathan V, Michot P, Grohs C, et al. Rapid Discovery of De Novo Deleterious Mutations in Cattle Enhances the Value of Livestock as Model Species. *Sci Rep.* 2017; 7:11466. <https://doi.org/10.1038/s41598-017-11523-3> PMID: 28904385.
78. Cassandri M, Smirnov A, Novelli F, Pitolli C, Agostini M, Malewicz M, et al. Zinc-finger proteins in health and disease. *Cell Death Discov.* 2017; 3:17071. <https://doi.org/10.1038/cddiscovery.2017.71> PMID: 29152378.
79. Schmidlin F, Amadesi S, Dabbagh K, Lewis DE, Knott P, Bunnett NW, et al. Protease-activated receptor 2 mediates eosinophil infiltration and hyperreactivity in allergic inflammation of the airway. *J Immunol.* 2002; 169:5315–5321. <https://doi.org/10.4049/jimmunol.169.9.5315> PMID: 12391252.
80. Imhof BA, Aurrand-Lions M. Adhesion mechanisms regulating the migration of monocytes. *Nat Rev Immunol.* 2004; 4:432–444. <https://doi.org/10.1038/nri1375> PMID: 15173832.