Whole genome mapping for identification of genes associated with production traits under heat stress conditions in Mehsana Buffalo

M. Darji[1], J.P. Gupta[2], B.P. Brahmkshtri[3], S. Saha[4], J. Chaudhari[1], A. Chaudhari[1] and S.K. Mohapatra[5]

[1]*Department of Animal Genetics and Breeding, CVSc and AH, KU, SKNagar, India*
[2]*Department of Animal Genetics and Breeding, BVC, Patna, India*
[3]*Department of Animal Genetics and Breeding, CVSc and AH, KU, Himmatagar, India*
[4]*Animal Breeding Group, NDDB, Anand,* [5]*Department of Veterinary Bitechnology, CVSc and AH, KU, SKNagar, India*

Abstract

A whole genome mapping analysis was conducted to identify genomic regions linked to heat stress tolerance in Mehsana buffaloes. The study focused on determining heat stress thresholds, analyze genetic components, and identifying candidate genes associated with first lactation test-day milk yield (FLTDMY), fat yield (FLTDFY), protein yield (FLTDPY), and SNF yield (FLTDSNFY) under heat stress conditions. The study utilized 130,475 test-day records from 13,887 buffaloes, obtained from the INAPH database of the National Dairy Development Board (NDDB), Anand, along with daily temperature and humidity data from the Deesa center of the India Meteorological Department (IMD), Pune. The statistical model included herd test-day, days-in-milk (DIM) classes, and the age of the buffalo as fixed effect and general and thermotolerance additive genetic and permanent environmental as random effects. The study revealed the THI threshold for FLTDMY, FLTDFY, FLTDPY and FLTDSNFY as 78, 79, 53, and 80, respectively. The results indicated that additive genetic variances for all milk production traits increased with rising temperature-humidity index (THI), indicating greater genetic selection potential under heat stress. Negative genetic correlations were found between general genetic effects and heat tolerance for FLTDMY (-0.40) and FLTDPY (-0.81), suggesting that high-producing buffaloes are more vulnerable to heat stress. In contrast, weaker correlations for FLTDFY (0.06) and FLTDSNFY (0.15) indicated a lower impact of heat stress on these traits. Genomic association analysis revealed five significant regions on chromosomes BBU2, BBU4, BBU5, BBU20, and BBU23 associated with milk production under heat stress conditions. These regions contained candidate genes, including KHDRBS2, FSTL4, ATP10A, ATG8, TBC1D15, RAB21, TPH2, SPG11, and FZD4, which play roles in heat stress response mechanisms. This study provided valuable insights into the genetic basis of heat tolerance in Mehsana buffaloes and highlighted potential of genomic selection to enhance resilience to heat stress.

Introduction

The Mehsana buffaloes are spread across the natural habitat within areas of North Gujarat, which is having a great influence on the dairy sector in India (DAHD and F, 2023). Mehsana buffaloes are also well known to be persistent milkers and regular breeders. The demand for milk and its products in India will soar to 266.5 million metric tonnes by 2030 as per study commissioned by the National Dairy Development Board (NDDB) (Economic Survey, 2021). Additionally, global demand for animal products, including meat, eggs, and milk, is expected to surge by 70% by 2050 (Berckman, 2017). Food and water security will be one of the biggest challenges for humankind in the latter half of the 21st century. While progeny testing, pedigree selection programs and high throughput genomic data offer solutions to improve the genetic potential of animals, the adverse effects of environmental changes can still hamper animal performance.

Climate change is one of the major threats to the survival of various species, ecosystems and sustainability of livestock production systems across the world, especially in tropical and temperate countries. Intergovernmental Panel on Climate Change (IPCC 2014) reported that the temperature of the earth is increasing by 0.2 °C per decade and also predicted that the global average surface temperature would increase to 1.4-5.8 °C by the year 2100. Livestock in a tropical country like India, experiences extreme weather conditions, where the temperature may reach up to 50°C in summer and decline to as low as -20°C during winter. Among many climatic factors, the most crucial is the thermal environment, and it is primarily accountable for changes in additive genetic variance (AGV) for particular traits in livestock populations. This thermal environment comprises ambient temperature, moisture content and air circulation. The best combination of these parameters creates an environmental condition in which an animal feels comfortable called a thermal comfort zone. When the conditions rise above this zone, the animal requires extra energy for thermoregulation and thus affects production.

When the environmental conditions deviate outside the conditions of the thermal comfort zone the animal falls under heat stress. Heat stress can be defined as “A condition that occurs when an animal cannot dissipate an adequate quantity of endogenous heat, whether it is produced or absorbed by the body, to maintain body thermal balance (Bernabucci *et al.,* 2014)”. It is broadly recognized that heat stress can cause several detrimental effects on bovines, such as reduced consumption of feed, reduced milk production, decreased reproductive efficiency, and increased disease susceptibility. Moreover, heat stress might cause physiological changes that include elevated heart and breathing rates, increased skin blood flow, and increased perspiration. These changes can induce dehydration, inflammation, and electrolyte loss. It is well documented that the stress of hot environment, lowers productive and reproductive efficiency in farm animals, especially among dairy animals. Likewise, research information is available to aid in the management of buffaloes in such adverse conditions. In response to these limitations, dairymen should plan the environmental and reproductive management of buffaloes during heat stress conditions (Ahmad, 2010).

The selection of animals under heat stress conditions is a critical aspect of livestock breeding programs, particularly in regions where high temperatures pose significant challenges to animal productivity and welfare. BLUP models account for genetic variation in animal populations, enabling the estimation of breeding values for traits of interest, even under heat stress conditions. This allows breeders to identify genetically superior animals with higher tolerance to heat stress and desirable production traits. Genomic BLUP models (ssGBLUP and GBLUP) and Bayesian regression models leverage genomic information, such as DNA markers, to enhance the accuracy of estimated breeding values. By considering genetic markers associated with heat tolerance and production traits, these models provide more precise predictions of an animal’s genetic merit, facilitating more effective selection decisions.

GWAS is crucial for obtaining reliable estimates of genomic breeding values, which in turn are central to the success of genomic selection for improving various quantitative traits in cattle (Zhang *et al.,* 2010; Lee *et al.,* 2017). A powerful approach to annotate phenotypic effects related to various economic traits could be employing SNPs as genetic markers in GWAS. One of the most significant gains from a genome-wide association study could be the identification of the novel quantitative trait nucleotides associated with various traits, which could be instrumental in making the selection procedure highly economic, thus providing a better cost-benefit ratio. Mehsana buffaloes are considered to have a lineage of both Murrah and Surti buffaloes. These buffaloes are distributed in the Mehsana, Banaskantha and Sabarkantha districts of the northern part of Gujarat state. Here, the temperature in summer goes up to 40°C and in winter, it falls to as low as 9-10°C. Thus, studying the breeding value of Mehsana buffalo under heat stress in the North Gujarat region provides valuable insights on improving productivity, with implications for both local farmers and broader livestock management practices. Taking into consideration all the above facts, the proposed study is to be undertaken with the objectives; to estimate heat stress threshold for the First lactation test day milk yield (FLTDMY), First lactation test day fat yield (FLTDFY), First lactation SNF yield (FLTDSNFY) and First lactation protein yield (FLTDPY) and conduct GWAS for identification of SNPs association with production traits during heat stress.

Materials and methods

Data

Data on 130,475 FLTDMY, FLTDFY, FLTDSNFY and FLTDPY records of 13,887 Mehsana buffaloes calved from 1994 to 2022 in 274 villages of 3 districts (Mehsana, Patan and Gandhinagar) of Gujarat state, India were gathered from the Information Network for Animal Productivity and Health (INAPH) database at National Dairy Development Board (NDDB), Anand. After calving, the 5th day was considered a first test day, and the milk yield data was recorded on the subsequent monthly test day. Data regarding milking was gathered by summing the morning and evening milk yields on the test day. The total milk yield for each day was then taken into account for the study. On most farms, hand milking was used. At least 5 test-day lactation records, from the 5th to 305th days in milk (DIM), were considered for the study.

Moreover, all buffaloes considered for the study were healthy and dewormed following the regular deworming schedule (1st dose at 3 days of age followed by monthly intervals up to 6 months and then thrice a year). Buffaloes were kept in stall feeding with a shaded area (Figure 1). The rolling herd average of the population over the years is given in Table 1. There was no need for ethical authorization for the current study because we only worked on datasets. Pedigree files were created by tracing back to the base population involving 27,340 animals, including 342 sires.

Weather information was acquired from one of the centers of the India Meteorological Department, Pune, located at Deesa station in Gujarat State, India. The minimum and maximum distances from the weather station to the farm were 50 km and 159 km, averaging 100 km (Figure 2). Daily THI values were calculated using the THI formula given by Bianca (1962) and validated by Purohit *et al* (2021) for the first lactation milk yield trait in Mehsana buffaloes of North Gujarat region,

$$THI=(0.15\times T_{db}+0.85\times T_{wb})\times 1.8+32$$

Where T_{db} is dry bulb temperature, and T_{wb} is wet bulb temperature in Celsius degrees. As recommended by Bohmanova *et al* (2007), the average daily THI over the three days preceding the test day was allocated to each test-day record. The range of THI in the present study was 42 to 83.

Figure 1. Buffalo housing system of operations in the study area.

Year of calving	MN_TY (Kg)	Year of calving	MN_TY (Kg)	Year of calving	MN_TY (Kg)
1994	1164.7	2004	1862.8	2014	2037.4
1995	1930.3	2005	1870.6	2015	2016.6
1996	2009.53	2006	1871.3	2016	1962.96
1997	1937.13	2007	1900.51	2017	2047.38
1998	1902.72	2008	1952.05	2018	1977.38
1999	1925.06	2009	2110.34	2019	1948.48
2000	1868.17	2010	2160.44	2020	1897.5
2001	1908.05	2011	2103.13	2021	1898.23
2002	1896.31	2012	2138.44	2022	1844.5
2003	1865.69	2013	2052.99		

MN_TY is Average of total milk yield of first lactation over the population.

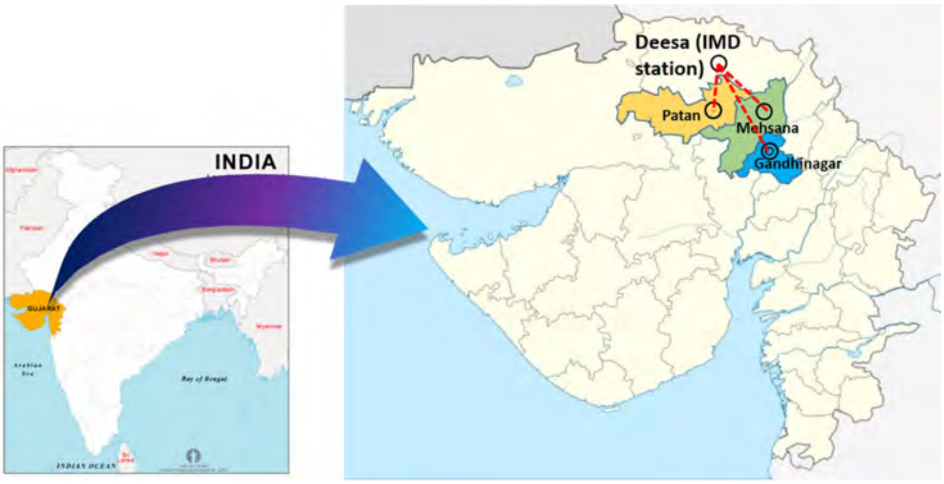

Figure 2. IMD station and farm area of 3 districts of Gujarat.

Estimation of THI threshold

The effect of each THI value was estimated on the FLTDMY using the following mixed model:

$$y_{ijklm} = a_i + pe_i + age_j + hy_k + THI_l + dims_m + e_{ijklm}$$

where a_i and pe_i are random additive and permanent environmental effects of ith animal, respectively; age_j is a fixed effect of jth calving age class where classes are arranged at every 6 months; h$\square_{\square}$ (herd x year of calving) is a random effect of the kth class of 1,429 total classes; THI_l is a fixed effect of lth THI; $dims_m$ (dim x season of calving) is a fixed effect of mth dims class of 978 total classes; e_{ijklm} is a residual. In this study, the village effect was taken as a herd effect.

The package Segmented, Mass, and nlme (Muggeo, 2003, 2008) developed by the R Core Team (2023) was used to compute the threshold THI through segmented-regression analysis, assuming that the influence of THI on the milk yield is linear.

To ascertain the reduction in milk yield trait in circumstances of thermal stress, a THI function termed f (THI) was developed as follows:

$$f(THI)= \begin{cases} 0 & \text{if } THI \le THI_{thr} \\ THI - THI_{thr} & \text{if } THI > THI_{thr} \end{cases}$$

Where the *f (THI)* was fell into the range of 0 to *THI - THI*$_{thr}$, the different heat stress levels developed after the thermoneutral zone (Ravagnolo and Misztal, 2000).

Statistical model

Univariate test day model was employed to calculate the variance components by applying random regression on THI function for FLTDMY, FLTDFY, FLTDSNFY and FLTDPY traits.

$$y_{ijklm} = HTD_i + DIM_{jk} + age_l + a_m + v_m[f(THI)] + p_m + q_m[f(THI)] + e_{ijklm}$$

where, y_{ijklm} is record for the yield trait under consideration; HTD_i is a fixed effect of ith herd test day; DIM_{jk} is a fixed effect of j^{th} DIM class, where classes are arranged every 20 days in the *k*th season; age_l is a fixed effect of l^{th} calving age class where classes are arranged at every 6 month; a_m is a general random additive genetic effect (AGE) for the buffalo m(intercept); v_m is a heat tolerant random AGE of the buffalo m(slope); *f(THI)* is a THI function for herd test day *l*; p_m is a random permanent environmental effect for the buffalo m in general conditions; $q_m$is a random permanent environmental effect of the buffalo m in thermal stress conditions; and e_{ijklm} is the random residual effect.

The variances are:

$$Var\begin{bmatrix} a \\ v \\ p \\ q \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & A\sigma_{av} & 0 & 0 & 0 \\ A\sigma_{av} & A\sigma_v^2 & 0 & 0 & 0 \\ 0 & 0 & I\sigma_p^2 & I\sigma_{pq} & 0 \\ 0 & 0 & I\sigma_{pq} & I\sigma_q^2 & 0 \\ 0 & 0 & 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

where A is numerator relationship matrix; I is identity matrix of the permanent environmental effect; $\sigma_{\boldsymbol{a}}^2$ is the variance of general AGE; $\sigma_{\boldsymbol{v}}^2$ is the variance of heat tolerant AGE; $\sigma_{\boldsymbol{av}}$ is the covariance between general and heat tolerant AGE; $\sigma_{\boldsymbol{p}}^2$ is the variance of general permanent environmental effect; $\sigma_{\boldsymbol{q}}^2$ is the variance of heat tolerant permanent environmental effects; σ_{pq} is the covariance between general and heat tolerant permanent environmental effects and σ_e^2 is the residual variance.

Estimation of variance component

Using univariate repeatability models, variance components for the FLTDMY, FLTDFY, FLTDSNFY and FLTDPY traits of Mehsana buffaloes were estimated by the REML method in the AIREMLF90 (version 1.93) tool of the BLUPF90 programme.

$$h^2 = \frac{\sigma_a^2}{\sigma_{total}^2}$$

where,

$$\sigma_{total}^2 = \sigma_a^2 + \sigma_p^2 + \sigma_e^2$$

Further, the estimate of heritability for production at heat stress level f(i) was calculated as,

$$h^2 = \frac{\sigma_a^2 + f(i)^2\sigma_v^2 + 2f(i)\sigma_{av}}{\sigma_a^2 + f(i)^2\sigma_v^2 + 2f(i)\sigma_{av} + \sigma_p^2 + f(i)^2\sigma_q^2 + 2f(i)\sigma_{pq} + \sigma_e^2}$$

The genetic correlation (r_{av}^G) of general AGE with heat tolerant AGE was calculated as,

$$corr\ [a, f(i)v] = \frac{f(i)\sigma_{av}}{\sqrt{\sigma_a^2 f(i)^2 \sigma_v^2}}$$

Genome-wide association study

Whole-genome association mapping was conducted using the ssGBLUP for production traits. The ssGBLUP model resembles the traditional BLUP model, but it substitutes the inverse of the pedigree relationship matrix (A^{-1}) with the inverse of the realized relationship matrix (H^{-1}) which integrates both pedigree and genomic information (Aguilar *et al.,* 2010). The combined pedigree and genomic relationship matrix H^{-1} was computed as part of the analysis.

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

where, G^{-1} = The inverse of the genomic relationship matrix
A_{22}^{-1} = the inverse of the pedigree relationship matrix of the animals with genotype information

Here, G^{-1} has a different dimension for different production and reproduction traits (Table 2).

Identification of genes

The UOA_WB_1 water buffalo genome sequence assembly served as the reference for assigning SNPs, employing the Bioconductor R package biodbNcbi (version 1.6.0, available at https://bioconductor.org/packages/biodbNcbi). SNPs were linked to genes if they fell within the genomic sequence of the gene or were located within 15 kb upstream or downstream of the gene. SNPs contributing to more than 0.5% of the genetic variance in a specific trait were deemed significantly associated with that trait.

Table 2. Dimension of G^{-1} and A matrix for production and reproduction traits.

Sr No.	Trait	Dimension of G^{-1}	Dimension of A matrix	Buffaloes with records	Sire
1	FLTDMY	2458 × 2458	27340 × 27340	2117	105
2	FLTDFY	2458 × 2458	27340 × 27340	2117	105
3	FLTDSNFY	1424 × 1424	27340 × 27340	1083	105
4	FLTDPY	1424 × 1424	27340 × 27340	1083	105
5	AFC	2202 × 2202	27340 × 27340	1861	105

Table 3. Segmented regression analysis of corrected yield of different traits on THI of the day for identification of heat stress threshold.

Traits	Parameter	Estimate	Std. Error	t value	Pr>|t|	Multiple R^2	Threshold
FLTDMY	Intercept	6.6607	0.038883	171.299	<0.001	0.76	78
	THI	-0.0048	0.000644	-7.441	<0.001		
	U1.THI	-0.0141	0.009613	-1.466	0.018		
FLTDFY	Intercept	0.5710	0.0032	178.47	<0.001	0.45	79
	THI	-0.0002	5.25E-05	-3.351	<0.001		
	U1.THI	-0.0023	0.001526	-1.523	0.025		
FLTDSNFY	Intercept	0.8666	6.28E-03	138.071	<0.001	0.86	80
	THI	-0.0012	9.91E-05	-12.216	<0.001		
	U1.THI	-0.0025	4.37E-03	-0.581	0.025		
FLTDPY	Intercept	0.3114	0.0193	16.133	<0.001	0.85	53
	THI	0.0003	0.000402	0.682	<0.001		
	U1.THI	-0.0010	0.000406	-2.52	0.021		
AFC	Intercept	1958.4180	31.259	62.651	<0.001	0.68	76
	THI	-3.0160	0.493	-6.118	<0.001		
	U1.THI	15.7420	2.787	5.649	0.023		

U1.THI is difference-in-slope parameter of the different trait

Results and discussion

Estimation of THI threshold

It was thought that heat stress would have a linear effect, the THI threshold estimated by segmented regression analysis was found to be 78 for the FLTDMY, 79 for the FLTDFY, 80 for the FLTDSNFY and 53 for the FLTDPY (Table 3). The FLTDMY, FLTDFY and FLTDSNFY observed 6.47 Kg, 0.57 Kg and 0.80 Kg at THI 42 which were fluctuated at many levels and reached 6.31 Kg, 0.55 Kg and 0.77 Kg at the threshold level, respectively. When the THI values crossed the threshold, the FLTDMY, FLTDFY and FLTDSNFY decreased gradually due to heat stress circumstances. The FLTDPY observed 0.32 Kg at THI 42 and then it was slightly increased up to 0.33 Kg at threshold level, but after crossing the threshold level it was decreased and found 0.30 Kg at extreme level of heat stress (Figure 3).

Stefani (2020) found a 77 THI threshold for TDMY records in dairy buffaloes using piecewise linear regression. Nonetheless, the THI threshold for FLTDMY, FLTDFY and FLSNFY in these investigation were determined to be greater than what Ravagnolo and Misztal (2000) had discovered. They have reported a threshold of THI 72 for the milk yield, fat yield and SNF yield trait in Holstein cattle. The range of threshold of THI for milk yield trait in cattle was 60 to 76 (Hagiya *et al.,* 2019; Ekine-Dzivenu *et al.,* 2022) and in buffalo, it was 72 to 77 (Purohit *et al.,* 2021; Stefani, G. 2020). This disparity in threshold values may be consequence of the methods used for threshold estimation or herd cooling systems e.g. fans or sprinklers. The climatic regions and milk production levels may also affect the threshold value (Carabaño *et al.,* 2017). The present findings extend the range of threshold values earlier obtained for buffalo for FLTDMY, FLTDFY and FLTDSNFY. Moreover, the results of higher heat tolerant capacity of buffaloes compared to cattle support the hypothesis that as compared to cattle, the buffalo is a more suitable animal for the milk production trait under heat stress conditions. The greater threshold of the THI level in the current investigation may be because of the acclimatization of Mehsana buffaloes against the heat stress conditions of the semi-arid region of North Gujarat.

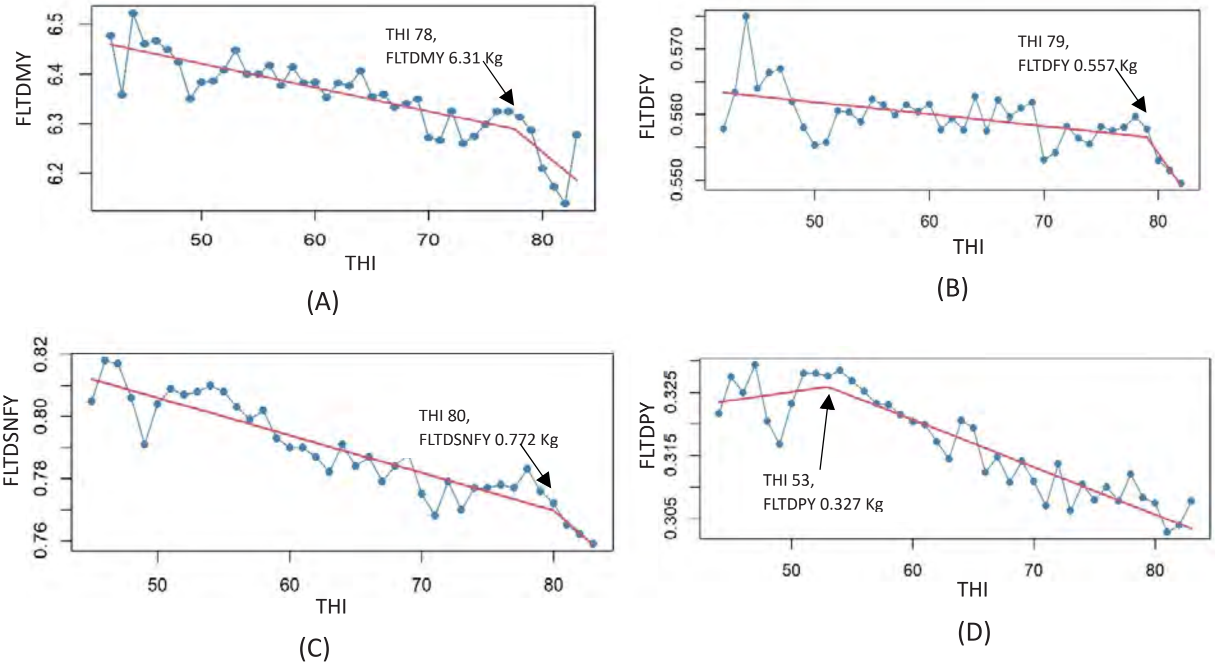

Figure 3. THI threshold was estimated by segmented regression in R version 4.3.0 for FLTDMY (A), FLTDFY (B), FLTDSNFY (C) and FLTDPY (D).

Estimation of variance components

Variance components for all the traits in the general condition (intercept) and the heat stress condition (slope) are displayed in table 4. The variance components for thermal stress conditions are quantitatively small because they represent variance for one THI unit over the threshold. Due to the lack of available literature on estimated variance components in buffalo, the same are compared with the genetic parameters of cattle. The additive genetic variances for general and thermal stress conditions were determined to be 0.813 and 0.013 for FLTDMY, 0.0033 and 0.0001 for FLTDFY, 0.0055 and 0.0002 for FLTDSNFY, and, 0.0017 and 0.0 for FLTDPY, respectively (Table 3). The variance components were obtained to be lower in the current investigation than Ravagnolo and Misztal (2000), who estimated it as 4.3 and 0.021 for the FLTDMY, 0.604 and 0.002 for FLTDFY, 0.37 and 0.001 for FLTDPY in Holstein cattle, respectively. Similarly, Aguilar *et al* (2009) also mentioned additive genetic variances for general and heat stress conditions as 5.5 and 0.028 for FLTDMY, 0.74 and 0.21 for FLTDFY, 0.42 and 0.17 for FLTDPY in Holstein cattle, respectively. Sigdel *et al* (2019) estimated it as 9.26 and 0.009 for FLTDMY, 1.19 and 0.18 for FLTDFY, 0.55 and 0.08 for FLTDPY variance components in Holstein cattle, respectively. The lower additive genetic variances found in Mehsana buffalo compared to Holstein cattle indicates that still there is a scope of genetic improvement in Mehsana buffalo for production and reproduction traits. The covariance between the additive genetic effect of general and heat stress conditions was determined to be negative for FLTDMY and FLTDPY, which confirms the results of Ravagnolo and Misztal (2000), who also estimated a negative covariance. Moreover, Ekine-Dzivenu *et al* (2022) also estimated a negative covariance in cattle for similar parameters. The covariance for FLTDFY and FLTDSNFY were found 0.0 and 0.0002, respectively.

Table 4. Variance components estimated with a repeatability test-day model for production traits.

Parameters	FLTDMY	FLTDFY	FLTDSNFY	FLTDPY
σ_a^2	0.816	0.0033	0.0055	0.00171
σ_v^2	0.013	0.0001	0.0002	0.00000
σ_{av}	-0.041	0.0000	0.0002	-0.00004
σ_p^2	0.988	0.0060	0.0117	0.00284
σ_q^2	0.067	0.0004	0.0005	0.00000
σ_{pq}	-0.075	-0.0005	-0.0015	-0.00004
σ_e^2	1.019	0.0082	0.0113	0.00295
h^2 *	0.29	0.19	0.19	0.23
r_{av}^G	-0.40	0.06	0.15	-0.81

* h^2 estimates at below THI threshold

GWAS for Production Traits

The identified genes associated with production traits were mentioned in the table 5. Figure 4 (A) display Manhattan plots showing the genomic regions associated with the FLTDMY trait under heat stress conditions. The results are presented in terms of the proportion of genetic variance explained by 2.0 Mb SNP windows. As expected, the plots show a clear peak on BBU2 at 51.5–53.5 Mb; this region harbors *KHDRBS2* (KH RNA binding domain containing, signal transduction associated 2), a gene that affects milk yield of buffalo under heat stress conditions. The two SNPs near or within the *KHDRBS2* gene were reported to be associated with teat number in Large White pig breed (Verardo *et al.* 2016). Additionally, the SNP near to *KHDRBS2* gene was also associated with pregnancy status in Brahman beef cattle (Reverter *et al.* 2016). This region on BBU2 that harbors *KHDRBS2* explained about 1.4% of genetic variance for the FLTDMY trait. Thus, this gene has a role in multiple biological functions in different species.

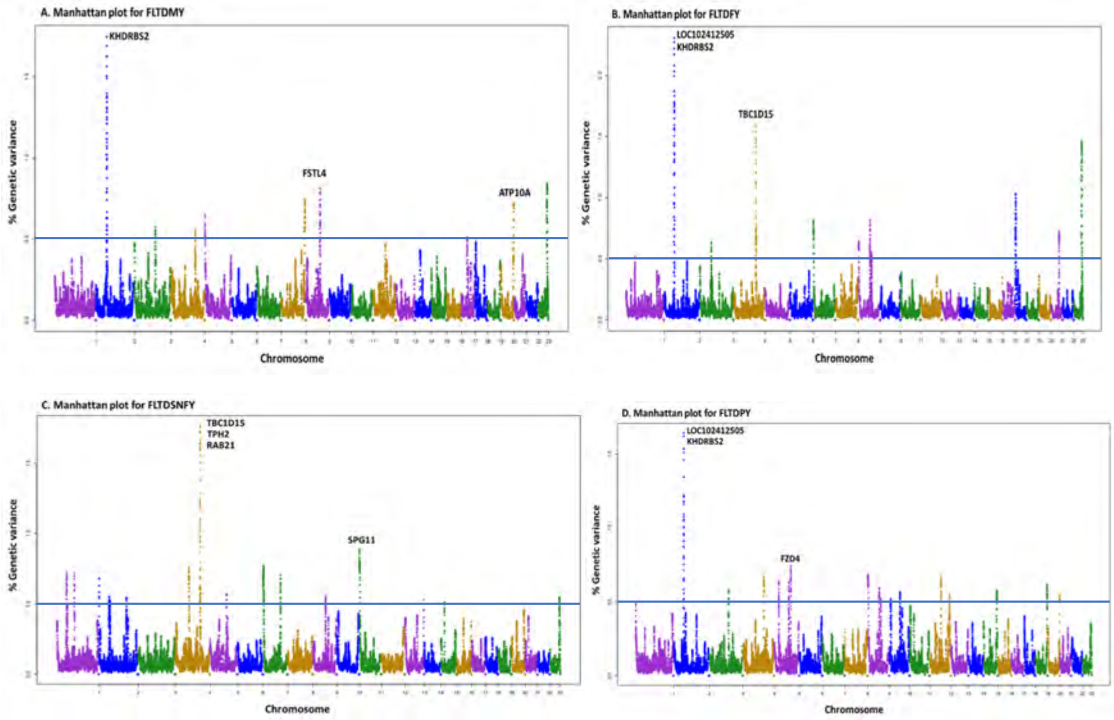

Figure 4. Manhattan plot for FLTDMY (A), FLTDFY (B), FLTDSNFY (C) and FLTDPY (D).

Another 2.0 Mb SNP window on BBU9 (66.1–68.1 Mb) explained about 0.81% of the additive genetic variance for the FLTDMY trait (Figure 4 (A)). Notably, this region harbors the gene *FSTL4*, the gene known for the expression of follicle-stimulating hormone (FSH). Lu *et al* (2021) reported the association of this gene with Rump angle (RA). Follicle-stimulating hormone (FSH) affects the ovarian follicular and corpus luteum dynamics, reproductive hormone secretion, and oestrus behaviour of dairy cows (Guarini *et al.,* 2019). However, according to current investigation, this gene plays a major role in an additive genetic variance of milk production in buffaloes.

Figure 4 (B) display Manhattan plots showing the genomic regions associated with the FLTDFY trait under heat stress conditions. The plots (Figure 4 (B)) show a clear peak on BBU2 at 50.5–52.5 Mb; this region harbors *LOC102412505* (alkylated DNA repair protein alkB homolog 8-like) suggesting a role in DNA repair processes and *KHDRBS2*, a gene that affects the fat yield of first lactation in Mehsana buffalo under heat stress conditions. It is involved in RNA metabolism and may influence lipid synthesis and secretion in mammary tissue. The region on BBU2 that harbors *LOC102412505* and *KHDRBS2* explained about 2.27% and 1.83% of genetic variance for the FLTDFY trait. Specifically, certain alleles or haplotypes of these genes were correlated with higher fat yield in Mehsana buffalo under heat stress conditions. These associations highlight the potential importance of *LOC102412505* and *KHDRBS2* in modulating lipid metabolism and fat synthesis in the mammary gland.

The genomic region on BBU4 (118.5-120.5 Mb) was associated with the FLTDFY trait in Mehsana buffalo under heat-stress conditions (Figure 4 (B)). This region possesses the putative gene *TBC1D15* (TBC1 Domain Family Member 15). This gene interacts with *ATG8* proteins on damaged lysosomes and provides a scaffold to assemble and stabilize the autophagic lysosomal reformation machinery. This potentiates the formation of lysosomal tubules and subsequent dynamin-2-dependent scission (Bhattacharya *et al.,* 2023). These findings suggest that specific alleles or haplotypes of TBC1D15 may be correlated with altered fat synthesis and secretion in mammary tissue, ultimately impacting fat yield in lactating animals. The region that harbors the *TBC1D15* gene explained about 1.56% genetic variance for the FLTDFY trait.

Figure 4 (C) display Manhattan plots showing the genomic regions associated with the FLTDSNFY trait under heat stress conditions. The plots (Figure 4 (C)) show a peak on BBU4 at 118.5–120.5 Mb; this region harbors *TBC1D15* (TBC1 Domain Family Member 15) and *RAB21* (RAB21, Member RAS Oncogene Family), a gene that affects the SNF yield of first lactation in Mehsana buffalo under heat stress conditions. The discovery of the *RAB21* gene as a locus associated with SNF yield in buffalo offers intriguing insights into the genetic determinants of milk production traits in this economically important livestock species, particularly under heat stress conditions. Interestingly, previous research has linked *RAB21* to oncogenesis and therapeutic targets in human glioma cells, suggesting a multifaceted role for this gene across different biological contexts (Ge *et al.,* 2017). Pei *et al* (2023) demonstrated an essential role for the small GTPase *RAB21* in regulating retromer-mediated recycling of the glucose transporter SLC2A1/GLUT1 and macroautophagy/autophagy. The involvement of *RAB21* in glucose transporter recycling and autophagy regulation suggests that it could influence the metabolic processes critical for milk production and composition. Proper functioning of the retromer complex and efficient recycling of glucose transporters are likely crucial for maintaining the metabolic homeostasis necessary for optimal lactation performance. The region on BBU4 that harbors *TBC1D15* and *RAB21* explained about 1.76% and 1.66% of genetic variance for the FLTDSNFY trait.

The genomic region on BBU11 (0.2-2.2 Mb) was associated with the FLSNFY trait in Mehsana buffalo under heat stress conditions (Figure 4 (C)), which harbors the candidate gene *SPG11* (SPG11 Vesicle Trafficking Associated, Spatacsin). The region that harbors this gene explained about 0.8 % of total genetic variance for the FLTDSNFY.

Figure 4 (D) display Manhattan plots showing the genomic regions associated with the FLTDPY trait under general and heat stress conditions. The plots show a peak on BBU2 at 50.5–52.5 Mb (Figure 4 (D)); this region harbors *LOC102412505* (alkylated DNA repair protein alkB homolog 8-like) and *KHDRBS2*, a gene that affects the protein yield of first lactation in Mehsana buffalo under heat stress conditions. The region on BBU2 that harbors *LOC102412505* and *KHDRBS2* explained about 1.62% and 1.05% of genetic variance for the FLTDPY trait.

Another 2.0 Mb SNP window on BBU5 (84.31–86.3 Mb) explained about 0.71% of the additive genetic variance for the FLTDPY trait (Figure 4 (D)). Notably, this region harbors the gene *FZD4*, the gene known as Frizzled Class Receptor 4. The identification of the *FZD4* gene associated with FLTDPY in buffalo under heat stress conditions represents a notable discovery in the field of livestock genetics. While previous research had primarily focused on *FZD4* in the context of plumage coloration in chickens (Davoodi *et al.,* 2022), this study unveils a novel role for *FZD4* in regulating milk protein synthesis and lactation performance in buffalo. *FZD4*, encoding the Frizzled-4 protein, is a member of the Frizzled family of transmembrane receptors involved in Wnt signaling pathways. In chickens, *FZD4* has been implicated in the regulation of pigmentation patterns, particularly in the development of plumage coloration. However, the association of *FZD4* with the FLTDPY in buffalo highlights its functional versatility across different species and physiological processes.

The association of *RNF14* with FLTDPY in our study suggests that this gene might be involved in optimizing protein synthesis and secretion in the mammary glands during heat stress. The ability of *RNF14* to function as a transcriptional regulator may allow it to fine-tune the expression of genes involved in protein metabolism and stress responses, thereby supporting higher protein yield during lactation.

Table 5. Identification of genes, chromosome and position based on the genetic variance explained by windows of 2 Mb for production traits.

Traits	Chr	Position (Mb)	Genetic variance	Genes identified
FLTDMY	2	49.6-52.9	0.52-1.69	*LOC102412505, LOC112578417, KHDRBS2, PRIM2, LGSN*
	3	101.3-101.7	0.51-0.57	*IL33, ERMP1, RIC1, KIAA2026, MLANA, TPD52L3*
	4	118.4-118.6	0.50-0.56	*TBC1D15, RAB21, TPH2, TMEM19*
	5	0.01-0.49	0.50-0.65	*TNNT2, LAD1, TNNI1, PHLDA3, CSRP1, NAV1, RNPEP, ELF3, TIMM17A, IPO9, SHISA4, LMOD1*
	8	116.5-117.4	0.55-0.73	*SHH, RBM33, CNPY1, EN2, INSIG1, HTR5A, DPP6, PAXIP1*
	9	66.00-67.10	0.50-0.81	*FSTL4, HSPA4, ZCCHC10, AFF4, LEAP2, GDF9, SHROOM1, SOWAHA, SCAMP4, ADAT3, ABHD17A, KLF16, REXO1*
	17	28.3-28.5	0.50-0.51	*FBRSL1, LRCOL1, P2RX2, POLE, PXMP2, PGAM5, ANKLE2, GOLGA3*
	20	66.3-66.7	0.69-0.72	*ATP10A, TRNAE-UUC*
	23	46.8-48.4	0.54-0.84	*MKI67, PTPRE, MGMT, CLRN3*
FLTDFY	1	43.3	0.51	*LOC112586137*
	2	49.3-52.9	0.51-2.27	*KHDRBS2, PRIM2, TRNAE-UUC, LOC112578779*
	3	55.8-56.1	0.52-0.57	*DGKE, NOG*
	4	117.4-119.9	0.57-1.59	*TBC1D15, TPH2, RAB21, TMEM19, THAP2, ZFC3H1, LGR5, TRHDE, TSPAN8, RHOU*
	7	0.01-1.47	0.57-0.82	*IDUA, SLC26A1, TMEM175, DGKQ, GAK, RNF212, CPLX1, FGFRL1, PCGF3, PDE6B, ATP5ME, SLC49A3, SPON2, CTBP1, MAEA, UVSSA, NKX1-1, LETM1, FAM53A, FGFR3, SLBP, POLN, NSD2, TMEM129, TACC3, NAT8L, PIGG, NELFA, HAUS3, MXD4, ZFYVE28*
	9	0.04-66.71	0.50-0.78	*ARAP3, DIAPH1, PCDH1, PCDH12, DELE1, RNF14, GNPDA1, NDFIP1, PCDHGC4, SLC25A46, STARD4, WDR36, FGF1, TMEM232, SPRY4, FSTL4, SLC35A4, APBB3, SRA1, EIF4EBP3, HSPA4, ARHGAP26, TAF7, ZCCHC10*
	18	0.98-2.77	0.66-1.03	*ZNRF1, WDR59, LDHD, ZFP1, BCAR1, FA2H, GABARAPL2, ADAT1, KARS, TMEM170A, TMEM231, ST3GAL2, TRNAG-UCC, PDPR, GLG1, SF3B3, COG4, FUK, AARS, EXOSC6, TRNAG-GCC, VAC14, MTSS1L, IL34, MLKL, RFWD3, VSTM2B*
	21	43.2-44.7	0.51-0.72	*ARHGEF3, FAM208A, DNAH12, CCDC66, IL17RD, DENND6A, ERC2, HESX1, APPL1, SLMAP, FLNB, TRNAE-UUC*
	23	46.6-48.5	0.55-1.46	*MKI67, PTPRE, MGMT, CLRN3, FOXI2*
FLTDSNFY	1	42.0-83.7	0.51-0.71	*CSMD1, MYOM2, LOC112587415*

Traits	Chr	Position (Mb)	Genetic variance	Genes identified
	2	0.09-51.7	0.50-0.55	*LRRC1, DUSP22, KHDRBS2, MLIP, KLHL31, LOC112581120*
	4	62.5-119.9	0.52-1.77	*TBC1D15, TPH2, RAB21, TMEM19, THAP2, ZFC3H1, TRHDE, LGR5, TSPAN8, METTL7B, ITGA7, BLOC1S1, RDH5, CD63, GDF11, SARNP, PYM1, DGKA, PMEL, CDK2, RAB5B, IKZF4, RPS26, ESYT1, MYL6B, MYL6, SMARCC2, ERBB3*
	5	76.2-76.4	0.53-0.57	*PANX1*
	7	0.01-84.4	0.51-0.77	*FAM53A, SLBP, TMEM129, TACC3, NAT8L, POLN, FGFR3, LETM1, NSD2, NELFA, UVSSA, NKX1-1, PIGG*
	9	57.6-58.3	0.51-0.54	*PCDH1, PCDH12, DELE1, RNF14*
	11	0.02-1.22	0.51-0.89	*SPG11, PATL2, EIF3J, TRIM69, RPS6KA5, TERB2, TTC7B, CTDSPL2, CASC4, CALM1*
	15	16.01-16.9	0.50-0.51	*VPS13B, KCNS2, STK3, RIDA, POP1, NIPAL2, MATN2, RPL30, OSR2*
	23	49.3-49.7	0.53-0.55	*TRNAC-ACA*
FLTDPY	2	49.5-52.1	0.51-1.62	*KHDRBS2, PRIM2, LOC112578779*
	3	101.1-101.5	0.52-0.58	*GLDC, UHRF2, IL33, UHRF2, TPD52L3, IL33, GLDC*
	4	109.0-110.3	0.54-0.67	*PPFIA2, ACSS3, TRNAC-ACA, LIN7A, PTPRQ, OTOGL, MYF5, MYF6*
	5	17.5-85.5	0.51-0.74	*FZD4, PRSS23, ME3, TMEM135, CCDC81, EED, HIKESHI, CCDC83, CACNA1E, PICALM, SYTL2, ZNF648, TEDDM1, GLUL*
	9	0.04-66.7	0.50-0.68	*WDR36, STARD4, CAMK4, SLC25A46, PCDH1, TMEM232, DELE1, RNF14, PCDH12, ARAP3, FGF1, GNPDA1, NDFIP1, SPRY4, HSPA4, DIAPH1, AFF4*
	10	3.25-53.7	0.50-0.57	*FBXL4, POU3F2, PDE10A*
	12	56.3-102.1	0.50-0.66	*CFAP77, BARHL1, DDX31*
	15	77.5-79.1	0.50-0.58	*TRAPPC9, KCNK9, FAM135B, COL22A1, TRNAC-ACA, CHRAC1, AGO2*
	19	69.12-69.19	0.53-0.54	*IRX1, TRNAC-GCA*
	20	66.3-66.7	0.51-0.54	*TRNAE-UUC, ATP10A*

Conclusions

The higher additive genetic variance under heat stress conditions for production traits gives the opportunity to breeders for the selection of animals under heat stress conditions. The negative genetic correlation between additive genetic effects in general and in heat stress conditions indicates the selection of animals for specific traits without considering heat stress, increases the susceptibility of animals to heat stress. Thermotolerance is a quantitative trait that was affected by several regions across the genome, with some prominent peaks on BBU2, BBU4, BBU5, BBU20, and BBU23.

List of references

Aguilar, I., Misztal, I., and Tsuruta, S. (2009). Genetic components of heat stress for dairy cattle with multiple lactations. Journal of Dairy Science, 92(11), 5702-5711.

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., and Lawlor, T. J. (2010). Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. Journal of Dairy Science, 93(2), 743-752.

Ahmad, S., and Tariq, M. (2010). Heat stress management in water buffaloes, a review. The Reverend Vet, 21(1), 297-310.

Berckmans, D. (2017). General introduction to precision livestock farming. Animal Frontiers, 7(1), 6-11. https://doi.org/10.2527/af.2017.0102

Bernabucci, U., Biffani, S., Buggiotti, L., Vitali, A., Lacetera, N., and Nardone, A. (2014). The effects of heat stress in Italian Holstein dairy cattle. Journal of Dairy Science, 97(1), 471-486.

Bhattacharya, A., Mukherjee, R., Kuncha, S. K., Brunstein, M. E., Rathore, R., Junek, S., Münch, C., and Dikic, I. (2023). A lysosome membrane regeneration pathway depends on TBC1D15 and autophagic lysosomal reformation proteins. Nature Cell Biology, 25(5), 685-698.

Bohmanova, J., Misztal, I., and Cole, J. B. (2007). Temperature-humidity indices as indicators of milk production losses due to heat stress. Journal of Dairy Science, 90(4), 1947-1956. https://doi.org/10.3168/jds.2006-513

Carabaño, M. J., Ramón, M., Díaz, C., Molina, A., Pérez-Guzmán, M. D., and Serradilla, J. M. (2017). Breeding and genetics symposium: Breeding for resilience to heat stress effects in dairy ruminants. A comprehensive review. Journal of Animal Science, 95(4), 1813-1826.

Davoodi, P., Ehsani, A., Vaez Torshizi, R., and Masoudi, A.A. (2022). New insights into genetics underlying of plumage color. Animal Genetics, 53(1), 80-93.

Department of Animal Husbandry and Dairying, Ministry of Fisheries, Animal Husbandry and Dairying Government of India. DAHD (2023). Annual Report 2021-22. AnnualEnglish.pdf (dahd.nic.in)

Economic survey. (2021). Page 613 - ES 2020-21_Volume-1-2 [28-01-21] (indiabudget.gov.in)

Ekine-Dzivenu, C. C., Mrode1, R., Oloo, R. D., Komwihangilo, D., Lyatuu, E., Msuta, G., Ojango, J.M., and Okeyo, A. M. (2022, December). Genomic analysis of milk yield and heat tolerance in small holder dairy system of sub-Saharan Africa. In Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges (pp. 1709-1712). Wageningen Academic Publishers.

Ge, J., Chen, Q., Liu, B., Wang, L., Zhang, S., and Ji, B. (2017). Knockdown of Rab21 inhibits proliferation and induces apoptosis in human glioma cells. Cellular and Molecular Biology Letters, 22, 1-11.

Guarini, A. R., Lourenco, D. A. L., Brito, L. F., Sargolzaei, M., Baes, C. F., Miglior, F., Misztal, I., and Schenkel, F. S. (2019). Genetics and genomics of reproductive disorders in Canadian Holstein cattle. Journal of Dairy Science, 102(2), 1341-1353.

Hagiya, K., Bamba, I., Osawa, T., Atagi, Y., Takusari, N., Itoh, F., and Yamazaki, T. (2019). Length of lags in responses of milk yield and somatic cell score on test day to heat stress in Holsteins. Animal Science Journal, 90(5), 613-618.

IPCC Fifth Assessment Synthesis Report-Summary for Policymakers-an Assessment of InterGovernmental Panel on Climate Change. IPCC, C. C. (2014). https,//www.ipcc.ch/site/assets/uploads/2018/02/SYR AR5 FINAL full.pdf.

Lee, J., Cheng, H., Garrick, D., Golden, B., Dekkers, J., Park, K., Lee, D., and Fernando, R. (2017). Comparison of alternative approaches to single-trait genomic prediction using genotyped and non-genotyped Hanwoo beef cattle. Genetics Selection Evolution. 49(1), 1-9. https://doi.org/10.1186/s12711-016-0279-9

Lu, X., Abdalla, I. M., Nazar, M., Fan, Y., Zhang, Z., Wu, X., Xu, T., and Yang, Z. (2021). Genome-wide association study on reproduction-related body-shape traits of Chinese Holstein cows. Animals, 11(7), 1927-1943.

Muggeo, V. M. (2003). Estimating regression models with unknown break-points. Statistics in Medicine, 22(19), 3055-3071. https://doi. org/10.1002/(ISSN)1097-0258

Muggeo, V. M. (2008). Segmented: an R package to fit regression models with broken-line relationships. R News, 8(1), 20-25. http://cran.r-project.org/doc/Rnews/

Purohit, P.B., Gupta, J.P., Chaudhari, J.D., Bhatt, T.M., Pawar, M.M., Srivastava, A.K., Patel, M.P., Patel, P.A., and Prajapati, M.N. (2021). Effect of heat stress on age at first calving in Mehsana buffaloes under field progeny testing programme. Indian Journal of Dairy Science, 74(4), 316-322.

R Core Team. (2023). R: A language and environment for statistical computing, reference index version 4.3.2. R Foundation for Statistical Computing, Vienna, Austria. Retrieved on 31 October 2023 from http://www.R-project.org/

Ravagnolo, O., and Misztal, I. (2000). Genetic component of heat stress in dairy cattle, parameter estimation. Journal of Dairy Science, 83(9), 2126-2130.

Reverter, A., Porto-Neto, L. R., Fortes, M. R. S., McCulloch, R., Lyons, R. E., Moore, S., Nicol, D., Henshall, J. and Lehnert, S. A. (2016). Genomic analyses of tropical beef cattle fertility based on genotyping pools of Brahman cows with unknown pedigree. Journal of Animal Science, 94(10), 4096-4108.

Sigdel, A., Abdollahi-Arpanahi, R., Aguilar, I., and Peñagaricano, F. (2019). Whole genome mapping reveals novel genes and pathways involved in milk production under heat stress in US Holstein cows. Frontiers in Genetics, 10, 1-10.

Stefani, G. (2020). Genetics of tolerance to heat stress in milk yield of dairy buffaloes assessed by a random regression model. [Dissertation, UNIVERSIDADE ESTADUAL PAULISTA – UNESP]. Google scholar.

Verardo, L. L., Silva, F. F., Lopes, M. S., Madsen, O., Bastiaansen, J.W., Knol, E. F., Kelly, M., Varona, L., Lopes, P.S., and Guimarães, S. E. (2016). Revealing new candidate genes for reproductive traits in pigs: combining Bayesian GWAS and functional pathways. Genetics Selection Evolution, 48, 1-13.

Zhang, Z., Liu, J., Ding, X., Bijma, P., de Koning, D. J., and Zhang, Q. (2010). Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. PloS One, 5(9), e12648-12654.