Genome-wide association study of heat stress response in *Bos taurus*

Bartosz Czech[1, *], Yachun Wang[2], Joanna Szyda[1]

1 Biostatistics Group, Department of Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland
2 National Engineering Laboratory for Animal Breeding, Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA, College of Animal Science and Technology, China Agricultural University, Beijing, China

* bartosz.czech@upwr.edu.pl

Abstract

Heat stress is a major challenge in cattle production, affecting animal welfare, productivity, and economic viability of the industry. In this study, we conducted a genome-wide association study (GWAS) to identify genetic markers associated with tolerance to heat stress in Chinese Holstein cattle. We genotyped 68 cows using Illumina 150K Bovine BeadChip microarray and analysed 112,081 single nucleotide polymorphisms using a linear model-based GWAS approach. We identified 17 SNPs distributed on three chromosomes that showed statistically significant associations with tolerance to heat stress in Chinese Holstein cattle. Five of them were located in introns of two genes, *PDZRN4* and *PRKG1*. *PDZRN4* is involved in protein degradation pathways, while *PRKG1* encodes a protein kinase involved in smooth muscle relaxation and blood vessel dilation. Our findings highlight the potential importance of *PDZRN4* and *PRKG1* in heat stress tolerance in cattle and provide valuable genetic markers for further research and breeding programmes aimed at improving the tolerance to heat stress in Holstein cattle. However, more studies are needed to elucidate the exact mechanisms by which these SNPs contribute to tolerance to heat stress and their potential implications for practical cattle breeding strategies.

Author summary

Heat stress is a critical challenge in cattle production, leading to reduced productivity and increased mortality rates. In our study, we conducted a genome-wide association study (GWAS) to identify genetic markers associated with indicators of tolerance to heat stress in cattle. We found significant associations between indicators of heat stress tolerance and specific single nucleotide polymorphisms (SNPs) located in two genes, *PDZRN4* and *PRKG1*. These genes are known to play roles in protein degradation pathways and smooth muscle relaxation, respectively, and have previously been implicated in physiological responses to heat stress in other species. Our findings provide insight into the genetic mechanisms underlying heat stress tolerance in cattle and could potentially be used in genomic selection programmes aimed at improving heat stress tolerance in cattle populations. More research is needed to elucidate the functional importance of these SNPs and their potential applications in cattle breeding programmes.

Introduction

Heat stress in cattle occurs when the animal's body temperature rises above physiologically normal levels due to exposure to high temperatures, humidity and solar radiation [1]. This can occur in both dairy and beef cattle and is a significant problem for the livestock industry, particularly in regions with hot and humid climates [2].

Heat stress can affect cattle in several ways. First, it can cause a decrease in feed intake and, later, reduce weight gain or milk production [3]. Second, heat stress can result in respiratory distress, panting, and increased water consumption, which can put additional strain on the animal's cardiovascular system [4]. Finally, severe heat stress can lead to dehydration, electrolyte imbalances, and dramatic changes in animal physiology [5]. In addition to the direct impact on animal health, heat stress also results in economic losses due to reduced milk production and a lower reproduction rate [6].

An exposition of cows to prolonged periods of heat stress changes in gene expression and epigenetic modifications, which can ultimately affect the animal's health, productivity and even the genetics of their offspring [7].

Studies have shown that heat stress can lead to changes in the expression of genes related to immune function, metabolism, and reproduction. For example, heat stress can cause a decrease in the expression of genes involved in milk production and an increase in the expression of genes involved in stress responses [8].

In the case of heat stress in cattle, GWAS can be used to identify genetic variants that are associated with body temperature, drooling score, and respiratory score, for example. This will allow breeding strategies to be developed to select animals that are more tolerant to heat stress and maintain productivity under hot and humid conditions [9].

Moreover, from the scientific perspective, GWAS allows understanding of the genetic basis of heat stress, including the biological pathways and mechanisms involved in the response to heat stress. In general, the combination of NGS, genotyping microarrays, and GWAS can provide a powerful approach to the identification of genetic variants and even candidate genes associated with the response to heat stress in cattle. This knowledge can be used to develop new management practises, breeding strategies, and therapeutics to improve animal welfare and productivity in a changing environment.

The purpose of this study was to identify genetic variants and metabolic pathways associated with the response to heat stress in cattle that lead to a better understanding of the functional basis of tolerance to heat stress in cattle.

Materials and methods

The material consists of 68 cows representing Chinese Holstein cattle. These individuals were genotyped using the Illumina 150K Bovine BeadChip (Illumina Inc., San Diego, CA, USA), which consists of 123,268 single nucleotide polymorphisms (SNPs). For all animals, the responses to heat stress were expressed by rectal temperature (RT), drooling score (DS), and respiratory score (RS). These phenotypes were represented by deregressed proofs (DRP) of estimated breeding values (EBVs) predicted as previously described in Czech *et al.* [10] [11].

The preprocessing of genotype data consisted of retaining: i) individuals with call rate greater than 0.9, ii) SNPs with minor allele frequency (MAF) greater than 0.05, and iii) SNPs that were in the Hardy-Weinberg equilibrium (P-value > 0.05). The filtration process was performed using PLINK software (v1.90b6.21) [12]. Subsequently, GWAS was performed separately for each phenotype, using the following model:

$$y = \mu + \beta X + \epsilon \tag{1}$$

where y is a vector of DRPs, X contains SNP genotype coded as 0, 1, or 2, representing the number of reference alleles, β is the SNP additive effect, and ϵ represents residuals. The significance of a SNP effect was tested using the likelihood ratio test with the reduced model represented by model 1 without the SNP effect. The estimation of the model parameter and testing of the significance of the SNP effect were performed using the GEMMA software [13]. To control for multiple testing P-values were adjusted using the Bonferroni correction. Significant SNPs were considered based on the adjusted P-values lower than 0.05. All significant SNPs were annotated using the Variant Effect Predictor (VEP) implemented in the ensemblVEP R package with Ensembl Release 109 (Feb 2023) [14]. Additionally, the Animal QTL database was used (QTLdb) to explain the genetic basis of variation in heat stress phenotypes [15]. Next, the Gene-Set Enrichment Analysis (GSEA) was performed to detect potential functional pathways underlying the heat stress response by applying one-sided version of Fisher's exact test. The Gene Ontology (GO) [16] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [17] were considered in GSEA implemented in the clusterProfiler R package [18].

Ethics approval and consent to participate

The data collection process was carried out strictly according to the protocol approved by the Animal Welfare Committee of the China Agricultural University. All experimental protocols were approved by the Animal Welfare Committee of the China Agricultural University. All methods are reported in accordance with ARRIVE guidelines (https://arriveguidelines.org) for reporting animal experiments.

Results

Genome-wide association study

The filtration process retained 112 081 out of 123 268 SNPs (91%) for GWAS for all 68 individuals. As a result of rectal temperature, 17 significant SNPs were identified, while no significant hits were observed for drooling and respiratory scores. Significant SNPs associated with rectal temperature were located on chromosomes 5, 17, and 26. On the *Bos taurus* autosome (BTA) 5 there were three significant SNPs, on BTA17 there were 12 significant SNPs, while on BTA26 there were only two significant SNPs. Manhattan plots were presented in Figure 1 (for rectal temperature), in Figure 2 (for drooling score) and in Figure 3 (for respiratory score). Table 1 shows detailed information on the significantly associated SNPs with rectal temperature. All SNPs were further annotated and processed through GSEA.

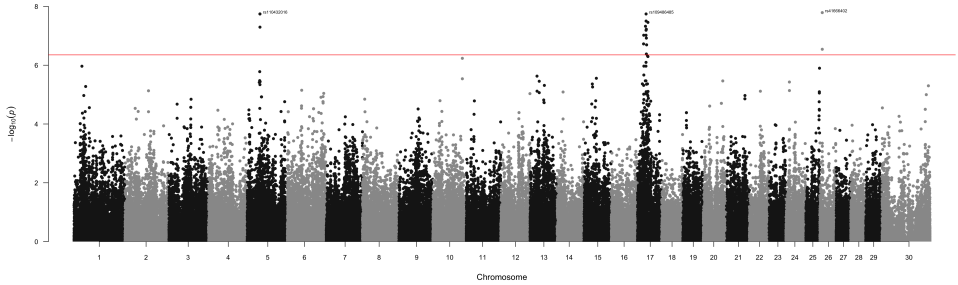

Fig 1. Manhattan plot for associations of SNPs with the rectal temperature. X-axis: SNPs positions on chromosomes, Y-axis: -log10 p-value. The red line indicates the 0.05 significance threshold corrected for multiple testing. Labels show the top significant SNPs.

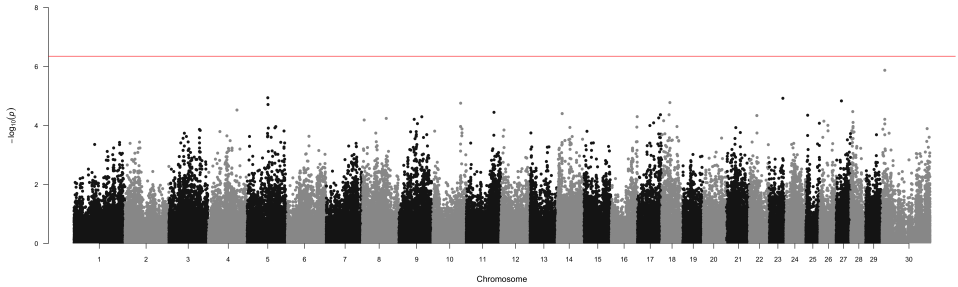

Fig 2. Manhattan plot for associations of SNPs with the drooling score. X-axis: SNPs positions on chromosomes, Y-axis: -log10 p-value. The red line indicates the 0.05 significance threshold corrected for multiple testing. Labels show the top significant SNPs.

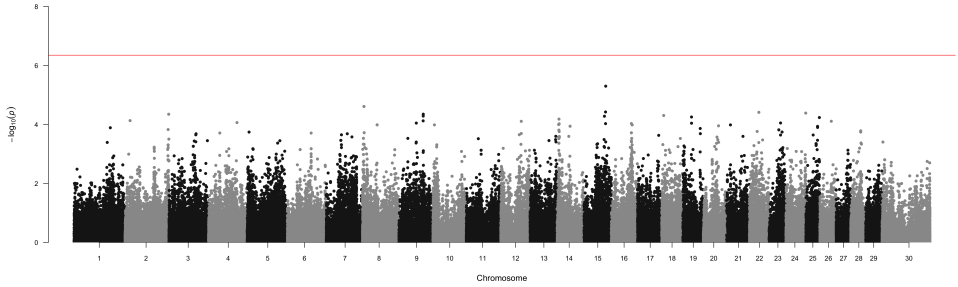

Fig 3. Manhattan plot for associations of SNPs with the respiratory score. X-axis: SNPs positions on chromosomes, Y-axis: -log10 p-value. The red line indicates the 0.05 significance threshold corrected for multiple testing. Labels show the top significant SNPs.

Table 1. The genomic annotation of associated SNPs with rectal temperature.

Chr.	Position	Alternative allele	MAF	Genomic annotation	Gene	p-adj
5	39 389 648	T	0.478	intron variant	*PDZRN4*	0.002
5	39 451 743	C	0.478	intron variant	*PDZRN4*	0.002
5	39 470 772	A	0.463	intron variant	*PDZRN4*	0.006
17	18 454 906	G	0.191	downstream gene variant	ENSBTAG00000015811	0.021
17	18 911 911	C	0.265	intergenic variant	-	0.011
17	24 328 037	C	0.309	intergenic variant	-	0.005
17	25 977 762	G	0.265	intergenic variant	-	0.011
17	26 111 051	T	0.235	intergenic variant	-	0.007
17	26 151 841	C	0.250	intergenic variant	-	0.002
17	26 207 829	C	0.235	intergenic variant	-	0.004
17	26 946 234	C	0.243	intergenic variant	-	0.007
17	27 068 312	C	0.243	intergenic variant	-	0.013
17	27 100 624	G	0.287	intergenic variant	-	0.047
17	27 837 317	T	0.221	intergenic variant	-	0.023
17	31 952 291	G	0.243	intergenic variant	-	0.004
26	8 148 418	G	0.353	intron variant	*PRKG1*	0.002
26	8 194 110	T	0.353	intron variant	*PRKG1*	0.032

Variants annotation and Gene-Set Enrichment Analysis

Results of the annotation process performed using VEP were summarised in Table 1 which shows that all significantly associated SNPs with rectal temperature on BTA5 and BTA26 were located in introns of *PDZRN4* (BTA5) and *PRKG1* (BTA25) genes. On BTA17, 11 out of 12 SNPs were located in the intergenic regions, while one SNP was located in the downstream part of the ENSBTAG00000015811 gene. The annotation of the phenotype based on QTLdb demonstrated that SNPs 17:27837317 (rs110432016) and 17:25977762 (rs109962820) were related to a maternal component of calving ease, dairy form, daughter pregnancy rate, foot angle, milk fat percentage, milk fat yield, net merit, length of productive life, milk protein percentage, milk protein yield, rear leg placement, and teat length. GSEA based on the GO indicated that genes related to the associated SNPs were enriched in the following ontologies: GO:0003682:chromatin binding, GO:0004672:protein kinase activity, GO:0004674:protein serine/threonine kinase activity, GO:0004692:cGMP-dependent protein kinase activity, GO:0005524:ATP binding, GO:0006468:protein phosphorylation, GO:0042802:identical protein binding, GO:0005515:protein binding, and GO:0046872:metal ion binding. GSEA for the KEGG pathways showed only enrichment of the cGMP-PKG signaling pathway (bta04022).

Discussion

Heat stress is an important environmental challenge for livestock production, including cattle, as it can negatively affect animal health, welfare, and productivity. In this study, we performed a GWAS to identify genetic markers associated with heat stress in cattle. Our findings revealed significant associations between heat stress and single nucleotide polymorphisms (SNPs) located in the *PDZRN4* and *PRKG1* genes, shedding light on the mechanisms underlying the response to heat stress in cattle. This GWAS study serves as a follow-up to previous analyses of differential gene expression and differential abundance of the microbiota in the context of heat stress, providing further insight into the complex interactions between genetics, gene expression, microbiota, and response to heat stress. Previous studies have already demonstrated the importance of rectal temperature as a main indicator of heat stress in cattle. It has been shown that all of the three phenotypes (rectal temperature, drooling score, and respiratory score), rectal temperature showed a major association with gene expression and abundance of microbiota in cattle under heat stress conditions [10] [11]. The heat stress phenotype is difficult to quantify and out of the three measurements that were available in this study, only rectal temperature appeared to be the most representative of heat stress.

There are many publications on GWAS related to heat stress in cattle, however, almost all of them focused on the standard case-control experimental design in which one cannot identify potential candidate genes responsible for the heat stress response. It is due to the complex nature of heat stress and the involvement of multiple genes and environmental factors. However, unlike previous studies that have focused on controlled experimental environments, our study examined animals in their production environment, providing valuable information on the genetic factors that influence the tolerance of heat stress in cattle under production conditions. Although this approach allows for capturing the genetic variation present in the population, it also has limitations, including potential confounding factors and the lack of control over environmental variables that may interact with the heat stress phenotype in real production systems.

The identification of SNPs in *PDZRN4* and *PRKG1* associated with heat stress in cattle suggests that these genes may play a role in the physiological response of cattle to heat stress. *PDZRN4* gene (PDZ domain containing the ring finger 4) also known as

LNX4 (Ligand of Numb Protein-X 4) plays a potential role as a tumour suppressor gene and may have an antiproliferative effect on hepatocellular carcinoma cell proliferation [19]. Another study showed that *PDZRN4* is a functional suppressor of prostate cancer growth [20]. However, there are no studies in which *PDZRN4* was indicated as a candidate gene related to the response to heat stress. Studies related to other livestock species showed that this gene could affect fat metabolism in pigs [21]. Furthermore, *PDDRN4* was found to be a significant gene associated with poor sperm motility in Holstein-Friesian bulls [22]. However, another gene identified in this study was *PRKG1* that encodes a protein called cGMP-dependent protein kinase 1 [23]. *PRKG1* was found as the gene associated with tick resistance in South African Nguni cattle [24]. Another study showed the importance of this gene in the local adaptation of indigenous Ugandan cattle to East Coast Fever [25]. However, the most interesting is that the gene *PRKG1* has already been found as a gene with a key role in body thermoregulation. In the study that focused on the adaptation to cold of indigenous Siberian populations, *PRKG1* has been shown to be the gene involved in cold acclimatisation [26]. Another study showed that this gene was the key to minimising heat loss by regulating blood vessel constriction in Yakutian horses [27]. There is also a study confirming the important role of *PRKG1* in temperature regulation in a cold environment in the Amur tiger [28]. Regarding the phenomenon of heat stress, it has been shown that *PRKG1* is associated with adaptation to heat stress in Egyptian sheep breeds [29].

Tolerance to heat stress is a complex trait that involves the interplay of multiple genetic and environmental factors [30]. SNPs located in *PDZRN4* and *PRKG1* provide valuable markers for selecting heat-stress-tolerant animals in breeding programmes. This may lead to the development of genomic selection programmes to improve heat stress resistance in cattle and improve animal welfare and productivity in hot climates.

It is important to note that our study has some limitations. First, the sample size may affect the statistical power to detect all SNPs associated with heat stress. However, having low power implies that the significant associations observed in our study may represent genes with an especially high impact on resistance to heat stress. Especially that *PRKG1* has already been confirmed as a heat stress-associated gene in other species, including humans. However, functional validation of SNPs located in both genes is warranted to further elucidate the underlying physiological mechanisms. Furthermore, more studies with larger sample sizes are needed to verify our findings and eventually identify additional SNPs and candidate genes with lower effects on heat stress.

Conclusion

In this study, we identified significant associations between SNPs located in the *PDZRN4* and *PRKG1* genes and heat stress in cattle, providing important information on the genetic basis of the tolerance to heat stress. Our findings contribute to the understanding of the physiological mechanisms underlying the response to heat stress and provide potential genetic markers for selecting animals tolerant to heat stress in breeding programmes. Due to the limited sample size, further research is needed to validate our findings and identify genes with low to moderate effects.

References

1. Idris M, Uddin J, Sullivan M, McNeill DM, Phillips CJC. Non-Invasive Physiological Indicators of Heat Stress in Cattle. Animals. 2021;11(1):71. doi:10.3390/ani11010071.

2. Summer A, Lora I, Formaggioni P, Gottardo F. Impact of heat stress on milk and meat production. Animal Frontiers. 2018;9(1):39–46. doi:10.1093/af/vfy026.

3. Liu J, Li L, Chen X, Lu Y, Wang D. Effects of heat stress on body temperature, milk production, and reproduction in dairy cows: a novel idea for monitoring and evaluation of heat stress — A review. Asian-Australasian Journal of Animal Sciences. 2019;32(9):1332–1339. doi:10.5713/ajas.18.0743.

4. Kitajima K, Oishi K, Miwa M, Anzai H, Setoguchi A, Yasunaka Y, et al. Effects of Heat Stress on Heart Rate Variability in Free-Moving Sheep and Goats Assessed With Correction for Physical Activity. Frontiers in Veterinary Science. 2021;8. doi:10.3389/fvets.2021.658763.

5. Parker AJ, Hamlin GP, Coleman CJ, Fitzpatrick LA. Dehydration in stressed ruminants may be the result of acortisol-induced diuresis1. Journal of Animal Science. 2003;81(2):512–519. doi:10.2527/2003.812512x.

6. Wankar AK, Rindhe SN, Doijad NS. Heat stress in dairy animals and current milk production trends, economics, and future perspectives: the global scenario. Tropical Animal Health and Production. 2021;53(1). doi:10.1007/s11250-020-02541-x.

7. Corvo MD, Lazzari B, Capra E, Zavarez L, Milanesi M, Utsunomiya YT, et al. Methylome Patterns of Cattle Adaptation to Heat Stress. Frontiers in Genetics. 2021;12. doi:10.3389/fgene.2021.633132.

8. Corazzin M, Saccà E, Lippe G, Romanzin A, Foletto V, Borso FD, et al. Effect of Heat Stress on Dairy Cow Performance and on Expression of Protein Metabolism Genes in Mammary Cells. Animals. 2020;10(11):2124. doi:10.3390/ani10112124.

9. Gallagher MD, Chen-Plotkin AS. The Post-GWAS Era: From Association to Function. The American Journal of Human Genetics. 2018;102(5):717–730. doi:10.1016/j.ajhg.2018.04.002.

10. Czech B, Wang Y, Wang K, Luo H, Hu L, Szyda J. Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition. Frontiers in Microbiology. 2022;13. doi:10.3389/fmicb.2022.998093.

11. Czech B, Szyda J, Wang K, Luo H, Wang Y. Fecal microbiota and their association with heat stress in Bos taurus. BMC Microbiology. 2022;22(1). doi:10.1186/s12866-022-02576-0.

12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics. 2007;81(3):559–575. doi:10.1086/519795.

13. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics. 2012;44(7):821–824. doi:10.1038/ng.2310.

14. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016;17(1). doi:10.1186/s13059-016-0974-4.

15. Hu ZL, Park CA, Wu XL, Reecy JM. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Research. 2012;41(D1):D871–D879. doi:10.1093/nar/gks1150.

16. 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 Research. 2018;47(D1):D419–D426. doi:10.1093/nar/gky1038.

17. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Research. 2015;44(D1):D457–D462. doi:10.1093/nar/gkv1070.

18. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation. 2021;2(3):100141. doi:10.1016/j.xinn.2021.100141.

19. Hu T, Yang H, Han ZG. PDZRN4 acts as a suppressor of cell proliferation in human liver cancer cell lines. Cell Biochemistry and Function. 2015;33(7):443–449. doi:10.1002/cbf.3130.

20. Jin P, Wu L, Zhang G, Yang B, Zhu B. PDZRN4 suppresses tumorigenesis and androgen therapy-resistance in prostate cancer. Journal of Cancer. 2022;13(7):2293–2300. doi:10.7150/jca.69269.

21. Wang BB, Hou LM, Zhou WD, Liu H, Tao W, Wu WJ, et al. Genome-wide association study reveals a quantitative trait locus and two candidate genes on Sus scrofa chromosome 5 affecting intramuscular fat content in Suhuai pigs. Animal. 2021;15(9):100341. doi:10.1016/j.animal.2021.100341.

22. Hering DM, Olenski K, Kaminski S. Genome-wide association study for poor sperm motility in Holstein-Friesian bulls. Animal Reproduction Science. 2014;146(3-4):89–97. doi:10.1016/j.anireprosci.2014.01.012.

23. Coudert E, Gehant S, de Castro E, Pozzato M, Baratin D, Neto T, et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics. 2022;39(1). doi:10.1093/bioinformatics/btac793.

24. Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks and Tick-borne Diseases. 2016;7(3):487–497. doi:10.1016/j.ttbdis.2016.02.005.

25. Vajana E, Barbato M, Colli L, Milanesi M, Rochat E, Fabrizi E, et al. Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever. Frontiers in Genetics. 2018;9. doi:10.3389/fgene.2018.00385.

26. Cardona A, Pagani L, Antao T, Lawson DJ, Eichstaedt CA, Yngvadottir B, et al. Genome-Wide Analysis of Cold Adaptation in Indigenous Siberian Populations. PLoS ONE. 2014;9(5):e98076. doi:10.1371/journal.pone.0098076.

27. Librado P, Sarkissian CD, Ermini L, Schubert M, Jónsson H, Albrechtsen A, et al. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. Proceedings of the National Academy of Sciences. 2015;112(50). doi:10.1073/pnas.1513696112.

28. Du H, Yu J, Li Q, Zhang M. New Evidence of Tiger Subspecies Differentiation and Environmental Adaptation: Comparison of the Whole Genomes of the Amur Tiger and the South China Tiger. Animals. 2022;12(14):1817. doi:10.3390/ani12141817.

29. Aboul-Naga AM, Alsamman AM, Allali AE, Elshafie MH, Abdelal ES, Abdelkhalek TM, et al. Genome-wide analysis identified candidate variants and genes associated with heat stress adaptation in Egyptian sheep breeds. Frontiers in Genetics. 2022;13. doi:10.3389/fgene.2022.898522.

30. Cheruiyot EK, Haile-Mariam M, Cocks BG, Pryce JE. Improving Genomic Selection for Heat Tolerance in Dairy Cattle: Current Opportunities and Future Directions. Frontiers in Genetics. 2022;13. doi:10.3389/fgene.2022.894067.