INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
21–0528/2021/26–4–555–560
DOI: 10.17957/IJAB/15.1867
http://www.fspublishers.org

Short Communication

Novel Polymorphisms in Complete Coding Region of Heat Shock Protein 70.1 Gene in Subtropically Adapted Red Sindhi Cattle Breed

Asma Saeed[1†], Abdul Wajid[1†], Kamran Abbas[2], Gohar Ayub[1], Ayesha Mohiud Din[1], Quratul Ain[1], Sana Zahoor[1], Akhtar Ali[2], Masroor Ellahi Babar[3] and Tanveer Hussain[2*]

[1]*Department of Biotechnology, Virtual University of Pakistan, Lahore 54000*
[2]*Department of Molecular Biology, Virtual University of Pakistan, Lahore 54000*
[3]*University of Agriculture, Dera Ismail Khan, Pakistan*
[*]For correspondence: tanveer.hussain@vu.edu.pk
[†]Contributed equally to this work and are co-first authors
Received 30 April 2021; Accepted 09 August 2021; Published 28 September 2021

Abstract

Thermal stress is a persistent challenge to farm animals including cattle under tropical and sub-tropical climatic conditions and causes severe effects on the productivity of animals. In this study, complete coding sequences (CDS) of heat shock protein 70.1 (*HSP70.1*) gene sequenced in 45 unrelated animals were subjected to determine the genetic variations. The complete coding region of the bovine *HSP70.1* gene was 1926 bp encoding 641 amino acids with a calculated molecular weight of 70.26 kDa. Out of 14 variations observed in CDS, 36% were non-synonymous (NS) and 74% were synonymous (Syn). Out of five non-synonymous variations, four were found to be novel. We observed a relatively higher ratio of non-synonymous to synonymous variations (0.56) it indicates the *HSP70.1* gene under selection owing to its association with cellular thermos-tolerance. Three microsatellite markers were detected within the bovine *HSP70.1* gene. The finding of this research will provide future directions on the identification of significant genetic variations in the *HSP70.1* gene in other Pakistani cattle breeds for the selection of animals with better climate resilience and superior performance.

Keywords: Thermal stress; *HSP70.1*; Sequencing; Genetic variations; Red Sindhi

Introduction

Thermal stress induces a series of conditions in animals such as a decrease in reproduction, milk production, growth, feed intake, and immunity. Environmental-induced thermal stress is a combination of heat and humidity that negatively impacts the farm animal's health and performance. In tropical, subtropical, and arid regions, climate change associated thermal stress directly interconnected with livestock production (Sodhi *et al.* 2013). According to Global Climate Risk Index, Pakistan is placed fifth most vulnerable country to climate changes. Climate change is causing an increase in ambient temperature that escalating heat stress (HS) with substantial consequences for livestock. In Pakistan, livestock is a crucial asset for over 70% of people in the rural areas depends on animals for food and income (Wajid *et al.* 2013). Livestock production makes up a significant contribution to food, nutrient scarcity, and poverty reduction worldwide (Hussain *et al.* 2018). Henceforth, climate change is expected to have a great impact on livestock production, especially in high climate change-affected countries.

High temperature coupled with relative humidity impedes various cellular functions of livestock animals. Respond of living organisms to various physiological and physical stresses at the cellular level rapidly increase the biosynthesis of different stress-proteins like heat shock proteins (HSPs). HSPs are highly conserved ubiquitous proteins found in animal and plant cells. They are responsible for cellular homeostasis under different kinds of environmental stresses such as extreme temperature, drought, salinity, and heavy metal (Yer *et al.* 2018). They are important molecular chaperones perform a dynamic role in the survival of prokaryotic and eukaryotic cells in response to elevated ambient temperatures. The HSPs are large protein families are named and organized into various classed according to their molecular-weight and amino-acid sequences homology. There are sex major HSP families including HSP100, HSP90, HSP70, HSP60, HSP40 and other small HSPs (Si *et al.* 2019; Tripathy *et al.* 2020). Among all HSPs, the HSP70 is highly abundant, largest and conserved protein family all through evolution (Tripathy *et al.* 2020). HSP70s are stress-related proteins strongly upregulated by a variety of stresses and play a critical role in cellular protection and thermos-tolerance by mounting the

To cite this paper: Saeed A, A Wajid, K Abbas, G Ayub, AM Din, Q Ain, S Zahoor, A Ali, ME Babar, T Hussain (2021). Novel Polymorphisms in Complete Coding Region of Heat Shock Protein 70.1 Gene in Subtropically Adapted Red Sindhi Cattle Breed. *Intl J Agric Biol* 26:555–560

chaperone activity in the cytosol of mammalian cells (Sodhi *et al.* 2013). There are 17 HSP70 genes in the bovine *HSP70* gene family distributed over 12 bovine chromosomes (Tripathy *et al.* 2020). Twelve genes are multiexonic and five are intronless. Phylogenetic analysis showed a total of eight evolutionary groups of the HSP70 gene family. The HSP family genes revealed wide variations in nucleotide (1911 to 54,017 base pair [bp] in *HSPA2* and *HSPA4* respectively) and amino acid (357 to 1001 in *HSPB1* and *HYOU1* respectively) sizes. *HSP70.1* (also known as *HSPA1A* [HS 70 kDa protein 1A]) is a key ATP-dependent protein to perform a wide variety of cellular processes including proper folding of newly synthesized under normal conditions and stabilization and/or refold misfolded proteins into biologically active states (Li *et al.* 2011). The cellular localization for *HSP70.1* is the nucleus as well as cytoplasm. The bovine *HSP70.1* gene is located on chromosome 23 (position: 27,520,317–27,522,790).

Among bovine breeds, the mechanism of adaptability to heat stress (HS) is different; some are well adapted to normalize the body temperature in response to HS than other breeds. Red Sindhi cattle are an important Zebu dairy breed distributed in subtropical regions of Sindh province. The breed is predominantly found in subtropical regions and well known for its survival in hot humid conditions and heat tolerance. This breed is selected in this study to investigate the gene structure and genetic variations in the *HSP70.1* gene conferring the trait of thermos-tolerance.

Materials and Methods

Samples collection

To investigate the existing variations in *HSP70.1* gene diversity, a total of 45 blood samples were collected from genetically unrelated subtropically adapted Red Sindhi cattle breed visiting different breeding tracts in the Sindh province of Pakistan including Red Sindhi Cattle farm at Tando Muhammad Khan, Sindh. A total of 5 mL blood was collected from the jugular vein puncture of the animal into EDTA containing vacutainer tubes. The samples were brought and processed at Animal Genomics Laboratory, Virtual University of Pakistan.

Genomic DNA isolation and gene amplification

The genomic DNA was isolated from the whole blood through the previously described standard protocol (Iqbal *et al.* 2020). The isolated DNA was quantified by Nanodrop spectrophotometer (Theromo-Scientific) and stored at -40ºC until further use. Direct sequencing approach was used for the detection of genetic variations in the coding sequences of the *HSP70.1* gene by previously reported amplification and sequencing primers (Sodhi *et al.* 2013). The amplification reaction mixture was 25 μL containing 2 μL genomic DNA (50 ng), 2 μL Mg^{++} (2 m*M*), 2.5 μL dNTPs (200 μM each), 3 μL Buffer (1X), 1 μL each forward and reverse primers (10 pmol), 0.6 μL Taq DNA polymerase enzyme (5 U/μL, Thermo Scientific), and 12 μL nucleases free water. The Thermal Cycler (Veriti, Thermo Fisher Scientific) profile was at 95ºC for 5 min, followed by 35 cycles at 95ºC for 45 s, 59ºC for 1 min, and 72ºC for 1 min with a final extension at 72ºC for 7 min. The PCR products were run on 1% agarose gel electrophoresis and purified using a Quick Clean DNA gel-extraction kit (Qiagen, Val, CA). The complete coding sequences (CDS) of the *HSP70.1* gene were sequenced using three forward primers by ABI 3130 automated sequencer (ABI, Inc, Foster City, CA).

Sequences analysis

The obtained *HSP70.1* gene sequences were assembled, edited, and analyzed by BioEdit software v. 7 (Hall 1999). The CDS of *HSP70.1* gene was compared with the reference similar sequence of Bos indicus for detection of genetic variation by Ensemble Genome Browser and further confirmed by manual inspection. The phylogenetic analysis was performed based on the CDS of *HSP70.1* gene with other mammalian species by the Neighbor-joining method using MEGAv6 software (Tamura *et al.* 2013). ProtParam tool (https://web.expasy.org/ protparam/) was used for the characterization of chemical and physical parameters of HSP protein (Gasteiger *et al.* 2005). Repeat masker software was performed to find the repetition elements (http://repeatmasker.org/cgi-bin/).

Results

The complete coding sequence (CDS) of *the HSP70.1* gene in the Red Sindhi cattle breed was obtained using overlapping primer pairs. The *HSP70.1* gene sequences were submitted to GenBank, and the following accession numbers MW694838 to MW694842 were obtained. The complete CDS in bovine was found to be 1926 bp long. The intronless Open reading frame (ORF) of bovine *HSP70.1* encoding 641 amino acids [aa]. The analysis of the entire CDS region of the *HSP70.1* gene in Red Sindhi revealed 14 genetic variations at different positions (Table 1). Out of the 14 variations detected in CDS of *HSP70.1* gene, 64% (n = 9) were synonymous (Syn) and 36% (n = 5) were nonsynonymous (NS) variations. In comparison to the *HSP70.1* gene sequences available for *Bos indicus* (GU183097) and *Bos taurus* (NC_037350), no indel was observed. Of the 14 variations, 36% (n = 5) were GA, 22% (n = 3) were CT, 14% (n = 2) were GC or GT and 7% (n = 1) were CA or AT bases changed. Out of five non-synonymous variations, four were found to be novel, which has not been previously described. A total of two novel non-synonymous variations were assumed to have a damaging functional effect, while all other variations having a benign effect (Fig. 1). In the present study, three microsatellite makers were discovered within the coding sequences of *HSP70.1* gene in Red Sindhi cattle (Table 2).

Table 1: Description of genetic variations in the CDS of *HSP70.1* gene in Red Sindhi cattle breed

Sr .No.	Sites	Variation	Tran/Trans	Rep/Nov	ID	AA	Change	(dS/dN)	Protein domain
1	14	C > T	trans	Rep	rs385826597	5	Met/Thr	dN	N-terminal
2	15	G > A	trans	Rep	rs382492082	5	Met/Ile	dN	N-terminal
3	126	G > A	trans	Rep	rs135145204	42	Val/Val	dS	N-terminal
4	156	G > C	transv	Rep	rs110903970	52	Gly/Gly	dS	N-terminal
5	324	A > G	trans	Rep	rs109475441	108	Lys/Lys	dS	N-terminal
6	408	C > T	trans	Rep	rs134962783	136	Gly/Gly	dS	N-terminal
7	540	C > A	transv	Rep	rs133720614	180	Ala/Ala	dS	N-terminal
8	573	G > C	transv	Rep	rs110374561	191	Gly/Gly	dS	N-terminal
9	963	A > G	trans	Rep	rs136751944	321	Leu/Leu	dS	N-terminal
10	970	G > T	transv	Noval	-	324	Ala/Ser	dN	N-terminal
11	1469	C > T	trans	Noval	-	490	Ala/Val	dN	Peptide binding
12	1535	A > T	transv	Noval	-	512	Lys/Met	dN	Peptide binding
13	1560	G > T	transv	Noval	-	520	Glu/His	dN	Peptide binding
14	1632	G > A	trans	Rep	rs41257359	544	Ser/Ser	dS	C-terminal

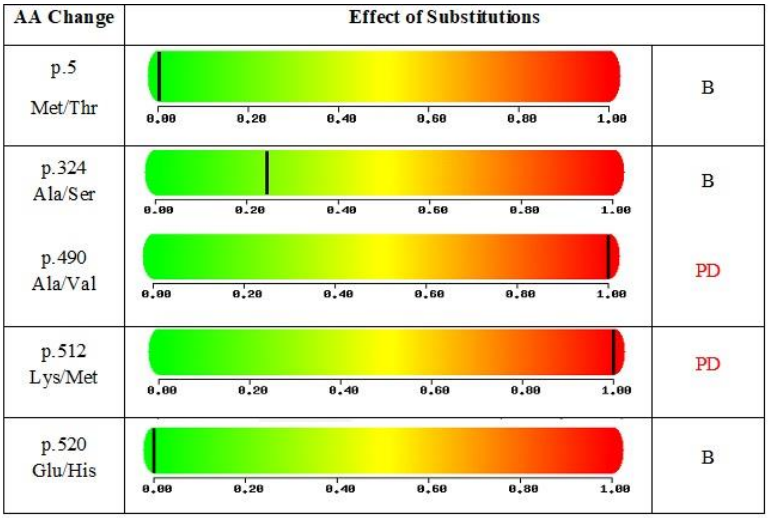

Fig. 1: Effect of amino acid polymorphisms of *HSP70.1*: B=Benign, PD=possibly damaging

At polymorphic sites of the *HSP70.1* gene, the ratio of dS/dN (ω) substitutions was found < 1 indicating purifying selection. Phylogenetic analysis revealed clustering of Red Sindhi cattle breed with *Bos indicus* as the nearest neighbor (Fig. 2). Cattle is closely related to yak (*Bos grunniens*), followed by buffalo (*Bubalus bubalis*), sheep (*Ovis aries*), and goat (*Capra hircus*). Phylogenetic and comparative sequence analysis for bovine *HSP70.1* indicated a high similarity with yak (99.9%), while *Sus scrofa*, *Mus musculus*, *Macaca mulatta*, Homo sapiens, and *Panthera pardus* showed maximum divergence (97.22 to 82.94%) and formed a distinct cluster.

To gain further insight into *HSP70.1* protein in bovine, the structure, conserved motifs, and physiochemical properties were predicted. The *HSP70.1* protein is highly conserved across many species that contained three putative structural domains, Actin-like ATPase domain (4–187, 190–382) that hydrolyzes ATP, peptide/substrate-binding domain (386–543) that binds substrate and C-terminal subdomain (538–619) provide a lid for the substrate domain were observed. The comparative analysis of *HSP70.1* gene in Red Sindhi cattle with other sequences revealed three signature patterns HSP70_1 (aa position 9–16; IDLGTTYS), HSP70_2 (aa position 197–210; IFDLGGGTFDVSIL) and HSP70_3 (aa position 334-348 LVLVGGSTRIPKVQK) were identified. Predicted *HSP70.1* protein of bovine possesses a molecular weight of 70258.51 Da with 641 aa of which 92 are negatively charged residues (Asp + Glu) and 83 are positively charged residues (Arg + Lys). Isoelectric point (pI) of bovine *HSP70.1* protein is acidic (5.67). The AI value of the *HSP70.1* protein was predicted 85.07 indicated that this protein is thermostable as well as contains a high amount of hydrophobic aa. *HSP70.1* had a negative GRAVY value (-0.398), indicating a soluble protein.

Discussion

There are contrary environments in Pakistan, physio-graphically divided into three climatic zones, the western highlands, northern mountains, and the Indus plain. As for south Pakistan, it is humid and hot. There are predominantly *Bos indicus* in the south of Sindh province. The major problem in native cattle breeds is low productivity due to

Table 2: Microsatellite in CDS of *HSP70.1* gene in Red sindhi cattle breed

Start	End	Unit	Repeats
532	540	GCC	3
1699	1707	AAG	3
1804	1809	GT	3

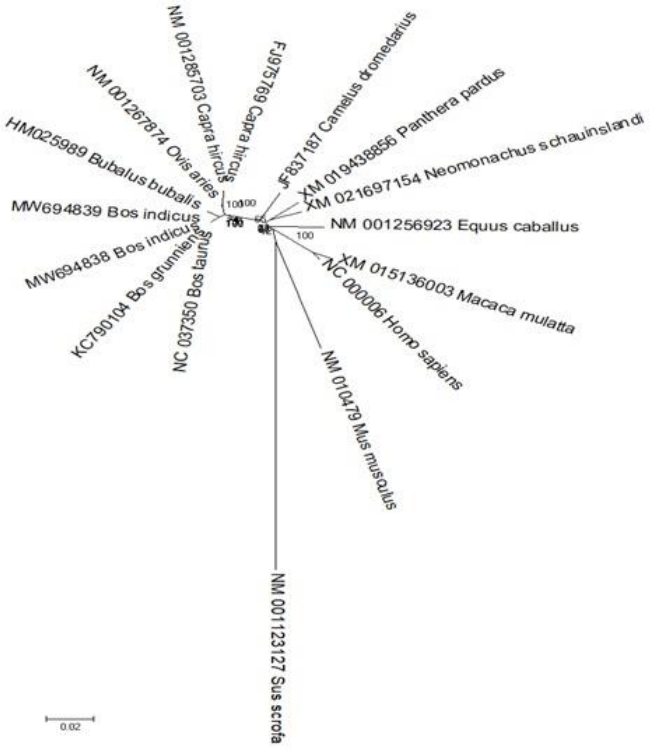

Fig. 2: Phylogenetic analysis of CDS of bovine *HSP70.1* gene with other mammalian species

persistent HS. Thermal stress causes severe effects on the health and performance of dairy animals globally resulting in huge economic losses. Knowledge of the genetic background of the animal may help in understanding the basis of heat tolerance and disease resistance, which are the key adaptable traits of cattle breeds (Saravanana *et al.* 2021). Identifying and selecting animals that are thermo-tolerant is an attractive alternative for reducing the negative effects of HS on animal's performance., The study was conducted to sequence the CDS of *HSP70.1* gene in subtropically adapted Red Sindhi cattle breed of Pakistan. *HSP70.1* gene is a key component of the HSP70 genes family and plays pivotal role incorporation with other molecular chaperone in cell survival and acquisition of thermos-tolerance (Sodhi *et al.* 2013). The *HSP70.1* plays an indispensable physiological role in cellular and systematic stresses, especially in livestock animals. The expression of *HSP70.1* in response to various physical and physiological stresses can help to improve the thermotolerance of animals. Previous studies in *Bos taurus* suggested that the genetic variations identified in coding sequence of *HSP70.1* gene were associated with stress-tolerance and disease susceptibility (Li *et al.* 2011). The Red Sindhi cattle breed distributed in different regions of Sindh province is known for its adaptability to elevated temperature with relative humidity. The thermos-tolerant animals could be used to investigate the polymorphisms in HSP genes conferring the thermos-tolerant trait (Mkize and Zishiri 2019). In this study, the complete *HSP70.1* gene (1926 bp encoding 641 aa) was sequenced in Red Sindhi cattle breed revealed 14 genetic variations with 64% (n = 9) were synonymous (Syn) and 36% (n = 5) were nonsynonymous (NS) variations. All the polymorphisms were previously studied in animals of high-temperature regions, except four novel polymorphisms identified in this study which have not been previously described. These polymorphisms could probably explicate the relatively high thermotolerance exhibited by the Red Sindhi cattle breed. The genetic variations in the coding sequence are important as these may alter the protein interactions and hence the animal's response to HS (Hassan *et al.* 2019). The SNPs in the promoter, untranslated regions, and coding sequence of the *HSP70.1* gene have been reported to be associated with heat tolerance, stress resilience, production performance and disease vulnerability of animals (Hassan *et al.* 2019). Moreover, previous studies revealed the association

between *HSP70.1* gene expression and thermos-tolerance mechanisms that revealed the increased expression under HS conditions (Abdelnour *et al.* 2018). *HSP70.1* gene expression in dermal fibroblast and skin was up-regulated during the summer season (40 and 44ºC) in Tharparkar and Karan-Fries cattle breeds (Singh *et al.* 2014; Maibam *et al.* 2017). Furthermore, the SNP (C/- & G/T) in the 5´UTR region of *HSP70.1* gene in Italian Holstein dairy cows revealed a significant association with tolerance to heat and stress response in peripheral blood mononuclear cells (PBMC) (Basirico *et al.* 2011). Sodhi *et al.* (2013) identified a total of 4 and 16 polymorphisms in 5´UTR and coding regions of *HSP70.1* gene respectively in 14 cattle breeds, while the 3´UTR was monomorphic. Li *et al.* (2011) identified four and one genetic variations in the 3´UTR and coding regions of *HSP70.1* gene respectively. Three thermo-tolerant genotypes (DD, FF and AB) among a total of 11 different genotypes showed higher milk fat and yield, and potassium content in erythrocytes respectively in Chinese Holstein cows. We observed a relatively higher ratio of dN/dS variations in *the HSP70.1* gene (0.56), it indicates the *HSP70.1* gene under selection owing to its association with cellular thermos-tolerance (Hassan *et al.* 2019).

Phylogenetic and evolutionary analysis showed that the HSP70 is the most conserved protein among the mammalian species. The DNA sequence of complete *HSP70.1* gene was 1926 bp in cattle encoded 641 aa. The aliphatic index (AI) is associated with the thermal stability of a protein (Ikai 1980). It is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine) of a protein. The increase in AI value indicates the more thermally stable protein (Ikai 1980). The AI value of the *HSP70.1* protein was predicted 85.07 indicated that this protein is thermostable as well as contains a high amount of hydrophobic aa. GRAVY (grand average of hydropathy) score indicates the hydrophilicity and hydrophobicity of a protein. A negative value of a protein indicates hydrophilic and a positive score indicates hydrophobic (Gasteiger *et al.* 2005). In the present study, *HSP70.1* protein had a negative GRAVY value (-0.398), indicating a soluble protein. In the present study, the findings indicate a high AI and negative GRAVY values showed thermal stability and hydrophilic nature respectively of *HSP70.1* protein that certifies the consistency of its chaperoning role in protein protection against specific stresses.

In brief, the genetic pattern of the *HSP70.1* gene in Red Sindhi cattle breed distributed in high temperature and persistent humidity zones of Sindh province was determined in this study. A few newly and several previously identified variants in the CDS of the *HSP70.1* gene were discovered.

Conclusion

In this study, we identified unique genetic variations in the Red Sindhi breed that may be associated to modulating gene expression in response to heat stress. Moreover, further research is needed, with a large sample size, to better understand the role and functions of the *HSP70.1* gene in cattle.

Acknowledgments

This work was supported by the Higher Education Commission of Pakistan (Grand No. NRPU-4885) provided to Dr. Tanveer Hussain. The authors would like to thanks Syed Salman Ahmad Deputy Director Head Quarter, L&DD Sindh and Red Sindhi Cattle farm at Tando Muhammad Khan, Sindh province.

Author Contributions

TH, AW, and MEB design and perceived the experiment, TH, AW and KA collected the blood samples, AS, AM, GA, and QA execute the experiment, AW, AA, SZ, AS and QA analyzed the data, AW wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest

Ethics Approval

This research work was carried out in accordance with the guidelines issued by the Ethical Review Committee of Virtual University of Pakistan (#VUP_01/2019).

References

Abdelnour SA, ME El-Hack, AF Khafaga, M Arif, AE Taha, AE Noreldin (2018). Stress biomarkers and proteomics alteration to thermal stress in ruminants: A review. *J Therm Biol* 79:120–134

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

Gasteiger EHC, A Gattiker, S Duvaud, MR Wilkins, RD Appel, A Bairoch (2005). *Protein Identification and Analysis Tools on the ExPASy Server, in The Proteomics Protocols Handbook*, pp:571–607. Walker JM (Ed). Humana Press, New York City, USA

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *In: Nucleic Acids Symposium Series*, pp:95–98. Jan 41 Information Retrieval Ltd., c1979–c2000, London

Hassan F, A Nawaz, MS Rehman, MA Ali, SMR Dilshad, C Yang (2019). Prospects of HSP70 as a genetic marker for thermo-tolerance and immuno-modulation in animals under climate change scenario. *Anim Nut* 5:340–350

Hussain T, ME Babar, MD Donato, A Wajid, A Nadeem, Z Ahmad, WA Khan, SO Peters, IG Imumorin (2018). Phylogeny of Pakistani Cattle Breeds using Mitochondrial Cytochrome b Gene. *Pak J Zool* 50:2029–2035

Ikai A (1980). Thermostability and aliphatic index of globular Proteins. *J Biochem* 88:1895–1898

Iqbal LN, Q Ain, A Wajid, A Mohiuddin, TM Yousaf, N Hussain, S Sherzada, F Manzoor, T Hussain, ME Babar (2020). Detection of genetic variations in pattern recognition receptors (PRRs) gene of Tharparkar cattle breed. *Pak Vet J* 40:400–402

Li Q, J Han, F Du, Z Ju, J Huang, J Wang, R Li, C Wang, J Zhong (2011). Novel SNPs in HSP70A1A gene and the association of polymorphisms with thermo tolerance traits and tissue specific expression in Chinese Holstein cattle. *Mol Biol Rep* 38:2657–2663

Maibam U, OK Hooda, PS Sharma, AK Mohanty, SV Singh, RC Upadhyay (2017). Expression of HSP70 genes in skin of zebu (Tharparkar) and crossbred (Karan Fries) cattle during different seasons under tropical climatic conditions. *J Therm Biol* 63:58–64

Mkize LS, OT Zishiri (2019). Novel single nucleotide polymorphisms in the heat shock protein 70.1 gene in South African Nguni crossbred cattle. *Trop Anim Health Prod* 52:893–901

Saravanana A, M Panigrahia, H Kumara, S Paridab, B Bhushana, GK Gaura, T Duttc, BP Mishrad, RK Singh (2021). Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. *Genomics* 113:955–963

Si FL, L Qiao, QY He, Y Zhou, ZT Yan, B Chen (2019). HSP superfamily of genes in the malaria vector *Anopheles sinensis*: diversity, phylogenetics and association with pyrethroid resistance. *Malar J* 18; Article 132

Singh U, R Deb, RR Alyethodi (2014). Molecular markers and their applications in cattle genetic research: A review. *Biomark Genom Med* 6:49–58

Sodhi M, M Mukesh, A Kishore (2013). Novel polymorphisms in UTR and coding region of inducible heat shock protein 70.1 gene in tropically adapted Indian zebu cattle (*Bos indicus*) and riverine buffalo (*Bubalus bubalis*). *Gene 527*:606–615

Tamura K, G Stecher, D Peterson, A Filipski, S Kumar (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729

Tripathy K, M Sodhi, RS Kataria, M Chopra, M Mukesh (2020). *In silico* analysis of HSP70 gene family in bovine genome. *Biochem Genet* 59:134–158

Wajid A, T Hussain, M Wasim, ME Babar, AA Anjum, SA Shah, K Abbas, MM Manzoor, N Badshah (2013). The future prospective of genomic biotechnology in animal Breeding: their potential for livestock production in Pakistan. *J Anim Plant Sci* 23:944–955

Yer EN, MC Baloglu, S Ayan (2018). Identification and expression profiling of all HSP family member genes under salinity stress in different poplar clones. *Gene* 678:324–336