Kafkas Universitesi Veteriner Fakultesi Dergisi
ISSN: 1300-6045 e-ISSN: 1309-2251
Journal Home-Page: http://vetdergikafkas.org

Kafkas Univ Vet Fak Derg
28 (3): 391-399, 2022
DOI: 10.9775/kvfd.2022.27152

Research Article

In Silico Analysis of the Structural and Functional Consequences of Polymorphic Amino Acid Substitutions in the Cattle HSF1 Protein

Sertaç ATALAY [1,a (*)]

[1] Central Research Laboratory (NABILTEM), Tekirdağ Namık Kemal University, TR-59000 Tekirdağ - TÜRKİYE
[a] ORCID: 0000-0003-4942-7729

Article ID: KVFD-2022-27152 Received: 19.01.2021 Accepted: 18.05.2022 Published Online: 19.05.2022

Abstract: Heat stress causes a decrease in the productivity of livestock by negatively affecting some important economic features such as fertility, growth and milk production. The heat shock transcription factor 1 (HSF1) gene plays a key role in the regulation of the stress response. Therefore, the present study aimed to predict the most deleterious non-synonymous single nucleotide polymorphisms (nsSNP) on the cattle HSF1 gene via in silico analyses. Out of 170 nsSNPs in the HSF1 gene, 14 SNPs were predicted as deleterious by all the nine servers (PredictSNP, Mapp, PhDSNP, PolyPhen-1, PolyPhen-2, Sift, Snap, nsSNPAnalyzer, and Panther). Consurf analysis determined that the vast majority of SNPs predicted to be deleterious were evolutionary conserved. Protein structural analyses were performed I-Mutant, Mupro, Hope Project server, RaptorX and Swiss Model server. The 12 amino acid substitutions (V15G, F18L, L19R, K21M, I35T, V46E, V56G, F61L, A67D, Y76D, V81G, L112P) in the DNA binding region of the cattle HSF1 protein were predicted to be highly deleterious. The P112 variant was predicted to disrupt an α-helix structure. It was determined that the two amino acid changes (K21M, Y76D) on the surface of the protein were different in terms of hydrophobicity, charge, and size. These variants (M21, D76) might hamper the protein's interaction with the heat shock elements.

Keywords: *HSF1, Heat stress, Cattle, nsSNP, HSP*

Sığır HSF1 Proteinindeki Amino Asit Polimorfizmlerinin Yapısal ve Fonksiyonel Sonuçlarının *In Silico* Analizi

Öz: Sıcaklık stresi, hayvanların doğurganlık, büyüme, süt üretimi gibi bazı önemli ekonomik özelliklerini olumsuz etkileyerek verimde azalmaya neden olmaktadır. Isı şoku transkripsiyon faktörü 1 (HSF1) geni, stres yanıtının düzenlenmesinde önemli bir rol oynar. Bu çalışma, sığır HSF1 geni üzerindeki, en zararlı eş anlamlı olmayan tek nükleotid polimorfizmlerini (nsSNP) *in silico* analizler ile belirlemeyi amaçlamıştır. HSF1 geni üzerinde bulunan 170 nsSNP dokuz tahmin programı (PredictSNP, Mapp, PhDSNP, PolyPhen-1, PolyPhen-2, Sift, Snap, nsSNPAnalyzer ve Panther) ile değerlendirildi. 14 nsSNP tüm tahmin programları tarafından zararlı bulundu. Consurf analizi, zararlı olduğu tahmin edilen SNP'lerin büyük çoğunluğunun evrimsel olarak korunduğunu belirledi. Proteinin yapısal analizleri, I-Mutant, Mupro, Hope Project, RaptorX ve Swiss Model sunucuları kullanılarak gerçekleştirildi. Sonuç olarak, sığır HSF1 proteininin DNA bağlama bölgesindeki 12 amino asit ikamesinin (V15G, F18L, L19R, K21M, I35T, V46E, V56G, F61L, A67D, Y76D, V81G, L112P) oldukça zararlı olduğu tahmin edildi. P112 varyantının, bir α-sarmal yapısını bozduğu belirlendi. HSF1 proteininin yüzeyindeki iki amino asit değişiminin (K21M, Y76D), hidrofobiklik, yük ve boyut açısından farklılığa neden olduğu belirlendi. Proteinin, M21 ve D76 varyantlarının, ısı şoku elementleri ile etkileşimini engelleyerek, ısı şok proteinlerinin transkripsiyonunu azaltabileceği tahmin edildi.

Anahtar sözcükler: *HSF1, Sıcaklık stresi, Sığır, nsSNP, HSP*

Introduction

The most important abiotic stress factor for livestock is the ambient temperature [1]. High environmental temperatures cause organisms to absorb more heat than they can dissipate, resulting in heat stress [2,3]. Heat stress causes a decrease in the productivity of livestock by negatively affecting some important economic features such as fertility, growth and milk production [4,5]. It also causes adverse effects on the immune responses of livestock [6]. Therefore, developing cattle breeds that can both tolerate heat stress and maintain productivity has become an important goal of researchers and cattle breeders around the world [7].

How to cite this article?

Atalay S: *In silico* analysis of the structural and functional consequences of polymorphic amino acid substitutions in the cattle HSF1 protein. *Kafkas Univ Vet Fak Derg,* 28 (3): 391-399, 2022.
DOI: 10.9775/kvfd.2022.27152

(*) Corresponding Author

Tel: +90 282 250 1128 Fax: +90 282 250 9920
E-mail: satalay@nku.edu.tr

The heat shock response (HSR) begins to produce with the transcription of heat shock genes in response to stress signals triggered by proteotoxic conditions such as elevated temperatures, oxidative stress, heavy metals and infections [8]. The HSR activation allows the cells to increase the expression of heat shock protein (HSP) genes known as molecular chaperones. The HSPs whose essential role is in protein folding, maintain cellular proteome homeostasis by assisting in the refolding of misfolded proteins or driving them to degradation [9,10]. The HSR is regulated by the heat shock transcription factor family (HSF) consisting of five-member (HSF1, HSF2, HSF3, HSF4, HSF5) in mammalians [11]. The evolutionarily highly conserved HSF1 acts as a key regulator of HSPs expression [12,13].

The HSF1 gene (ENSBTAG00000020751) is located on chromosome 14 in cattle and consists of 13 exons that encode 525 amino acids [14]. The bovine HSF1 protein (UniProtKB ID: Q08DJ8) contains several functional domains according to UniProtKB; DNA binding domain (position; 15-120 amino acids), N-terminal oligomerization domains-hydrophobic repeat (HR)-A/B (position; 130-203 amino acids), D domain (position; 203-224 amino acids), regulatory domain (position; 221-310 amino acids), oligomerization domain HR-C (position: 380-405 amino acids), transactivation domain and disordered region (position; 367-525 amino acids). The best-conserved region in the HSF family is the DNA binding domain (DBD) [15]. The HSF1 binds to DNA as a trimer. The trimer recognizes a specific nGAAn sequence in DNA known as heat shock elements (HSE) [16]. In the absence of stress, the majority of HSF1 exists in a monomeric conformation with a little affinity for the HSE. Spontaneous trimerization of the HSF1 is suppressed by the HR-C [17,18]. In stress conditions, the HSF1 is converted into a trimeric form that is a transcriptionally active form. The trimerization is regulated by the HR-A/B. The D domain interacts with JNK1 and MAPK3 and is involved in the translocation of HSF1 to the nucleus. The HSF1 is transported to the nucleus and post-transcriptional modifications occur. The regulatory domain is necessary for transcriptional activation through its phosphorylation. The transactivation domain is involved in directing HSF1 to specific target genes and regulating the extent of its activation [16,17].

It has been reported in previous studies that the HSF1 gene variants are associated with tolerance to heat stress [19], meat quality traits [20], reproductive traits [21] and milk yield [19]. The present study was designed to identify the deleterious nsSNPs in the bovine HSF1 gene using *in silico* analyses. These nsSNPs that are detrimental to the structure and function of the protein may be associated with susceptibility to heat stress in cattle.

Material and Methods

SNP Dataset

The SNPs in the HSF1 gene (ENSBTAG00000020751) were obtained from the Ensembl genome browser (www.ensemble.org). The amino acid sequence of HSF1 protein (Transcript ID: ENSBTAT00000083220.1) in the fasta format was retrieved from the Ensembl.

Prediction of Functional Effect

PredictSNP [22] was used to predict the effect of missense mutations in the bovine HSF1 gene. This tool is a consensus classifier that comprises scores from different predictors (MAPP, PhDSNP, PolyPhen-1, PolyPhen-2, SIFT, SNAP, nsSNPAnalyzer, and PANTHER) and classifies the variants as "Deleterious" and "Neutral". In addition, it converts the individual confidence scores of each estimator into a comparable scale that represents the expected percentage of accuracy, ranging from 0-100%. MAPP, PhDSNP, PolyPhen-1, PolyPhen-2, SIFT, SNAP, nsSNPAnalyzer, and PANTHER tools use different classification methods; physicochemical properties and alignment score, support vector machine, expert set of empirical rules, Bayesian classification, alignment score, neural network, random forest, and alignment score respectively [23]. SNPs that were found to be predicted as neutral by one tool have been excluded from the study.

Prediction of Protein Stability and Amino Acid Conservancy

The protein stability changes resulting from missense variants were predicted using I-Mutant 3.0 [24] and Mupro [25] servers. I-Mutant, a support vector machine-based automated web server, estimates the effect of an amino acid substitution effect on protein stability by calculating the free energy change value (DDG) of the native and mutant protein. The range of $-0.5 \leq DDG \leq 0.5$ values is classified as neutral mutation, while <-0.5 as a large decrease and $0.5<$ as a large increase. The Mupro is a server based on support vector machines and neural networks machine learning methods, which predict how single amino acid substitution affects protein stability. A negative DDG score indicates that amino acid substitution decreases protein stability, while a positive score indicates that protein stability increases. The bigger the DDG score, the more confident the result.

Conservation analysis of the HSF1 protein was performed using the Consurf webserver [26]. Utilized the phylogenetic relationships between homologous sequences, Consurf calculates the conservation scores of amino acid positions and determines functional regions. These conservations scores (1-3 is variable, 4-6 is average and 7-9 is highly conserved) are organized in color-coded regions depicted in the structure of the protein for representation.

Prediction of Protein 3-D Structure

The 3-D structural analysis and modeling studies of HSF1 protein were performed using the HOPE server [27], RaptorX [28], Swiss-PdbViewer v4.1 [29], and PyMOL v2.5 (The PyMOL Molecular Graphics System, Version 2.5 Schrödinger, LLC). The Hope server was used to investigate the effects of missense variants on protein structure. Hope server utilizes Protein Data Bank (PDB), UniProt databases, What IF web services, Reprof, Yasara programs, and Distributed Annotation System servers to estimate the effect of substitution amino acid on protein structure. It predicts hydrophobicity, charge, size change, and modeling of 3-D structure for natural and mutant amino acids.

The full-length 3-D structure of bovine HSF1 protein is not available in PDB. Hence, RaptorX was used to make a 3-D structural model for HSF1. It is a web server predicting the structure property of a protein sequence without using any template. Furthermore, the Swiss-Model server was used to predict the 3-D structure of the DNA binding domain of the HSF1 gene. The quality of models was evaluated according to Ramachandran Plot and MolProbity score by Swiss-Model Structure Assessment. PyMOL v2.5 and Swiss PDB Viewer v4.1 were used to visualize the HSF1 protein 3-D structures.

Results

Retrieval of nsSNPs and Function Prediction

The nsSNPs (n: 170) of the HSF1 gene were retrieved from the Ensembl SNP database. The retrieved nsSNPs were analyzed using PredictSNP, Mapp, PhD-SNP, PolyPhen-1, PolyPhen-2, Sift, Snap, and Panther. PredictSNP classified 51 out of 170 missense mutations to be deleterious, Mapp predicted 46 out of 170 missense mutations to be deleterious, PhD-SNP predicted 40 out of 170 missense mutations to be deleterious, PolyPhen-1 predicted 57 out of 170 missense mutations to be deleterious, PolyPhen-2 predicted 80 out of 170 missense mutations to be deleterious, Sift predicted 85 missense mutations to be deleterious, Snap predicted 44 out of 170 missense mutations to be deleterious, and Panther predicted 22 out of 170 missense mutations to be deleterious. Forty-four missense mutations were classified as "unknown" by the Panther tool. The effect of 14 SNPs *(Table 1)* categorized as "deleterious" in all tools on the protein structure and stability was investigated.

Prediction of Protein Stability and Amino Acid Conservancy

I-Mutant 3.0 and Mupro were used to predict changes in protein stability caused by functionally damaging 14 nsSNPs. The DDG value and binary classification estimate in I-Mutant 3.0 showed that thirteen amino acid substitutions (V15G, F18L, L19R, I35T, V46E, V56G, F61L, A67D, Y76D, V81G, L112P, V172E, F496C) were able to cause largely decreased stability (<-0.5 DDG) of the protein. The K12M mutation was classified as neutral (DDG= -0.14) by I-Mutant 3.0 *(Table 2)*. The Mupro made similar predictions for amino acid substitutions other than K21M. It predicted that the K21M mutation could

Table 1. *Results of nsSNPs predicted with I-Mutant, Mu-Pro and Consurf servers*

rsIDs	Allele	Residue Change	I-Mutant DDG Value (kcal/mol)	Mu-Pro DDG Value (kcal/mol)	Consurf Results		
					Functional/ Structural	Score	Domain
rs481839785	A/C	V15G	-2.43	-2.12	Structural	9	DNA binding
rs474682793	G/T	F18L	-1.10	-0.97	Structural	9	DNA binding
rs459488723	A/C	L19R	-1.75	-1.43	Structural	9	DNA binding
rs451876582	T/A	K21M	-0.14	0.18	Functional	9	DNA binding
rs445864779	A/G	I35T	-2.49	-1.97	Structural	9	DNA binding
rs454788300	A/T	V46E	-0.95	-1.22	Structural	9	DNA binding
rs442669059	A/C	V56G	-2.19	-2.25	Structural	9	DNA binding
rs432002899 rs453627348	G/C A/G	F61L	-1.21	-1.05	Structural	9	DNA binding
rs456086940	G/T	A67D	-0.76	-0.77	Structural	9	DNA binding
rs470395900	A/C	Y76D	-1.17	-1.20	Functional	9	DNA binding
rs477899807	A/C	V81G	-2.59	-1.98	Structural	9	DNA binding
rs443947366	A/G	L112P	-1.67	-1.87	Structural	9	DNA binding
rs481574682	A/T	V172E	-0.50	-0.81	-	8	N-terminal oligomerization domains
rs439067006	A/C	F496C	-1.34	-1.06	-	6	Disordered

Table 2. *The HOPE server analysis results for deleterious nsSNPs*

Amino Acid Substitutions	Size of Mutant Type a.a.[a] Relative to Wild Type a.a.[a]	Hydrophobicity of Wild Type a.a.[a] Relative to Mutant Type a.a.[a]	Wild Type a.a.[a] Charge	Mutant Type a.a. Charge
V15G	Smaller	More hydrophobic	-	-
F18L	Smaller	-	-	-
L19R	Bigger	More hydrophobic	Neutral	Positive
K21M	Smaller	Less hydrophobic	Positive	Neutral
I35T	Smaller	More hydrophobic	-	-
V46E	Bigger	More hydrophobic	Neutral	Negative
V56G	Smaller	More hydrophobic	-	-
F61L	Smaller	-	-	-
A67D	Bigger	More hydrophobic	Neutral	Negative
Y76D	Smaller	More hydrophobic	Neutral	Negative
V81G	Smaller	More hydrophobic	-	-
L112P	Smaller	-	-	-
V172E	Bigger	More hydrophobic	Neutral	Negative
F496C	Smaller	-	-	-

[a] *Amino acid*

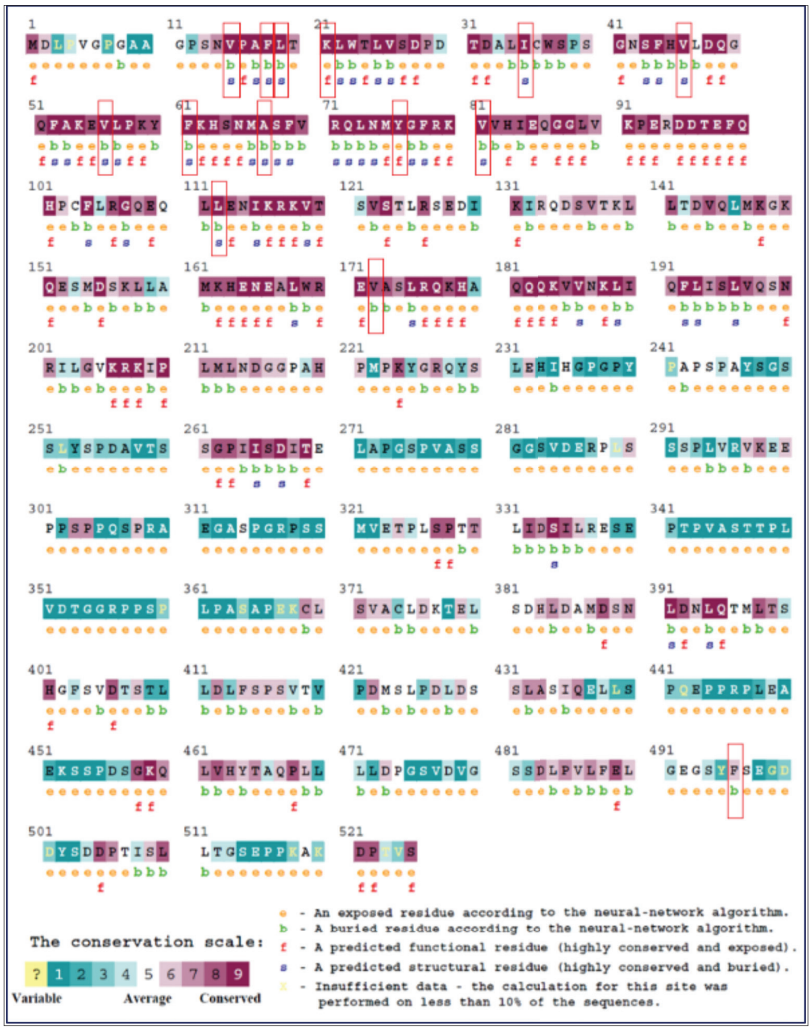

Fig 1. The Consurf analysis of the bovine HSF1 protein. The positions of the amino acid substitutions predicted to be most deleterious are marked with a red rectangle

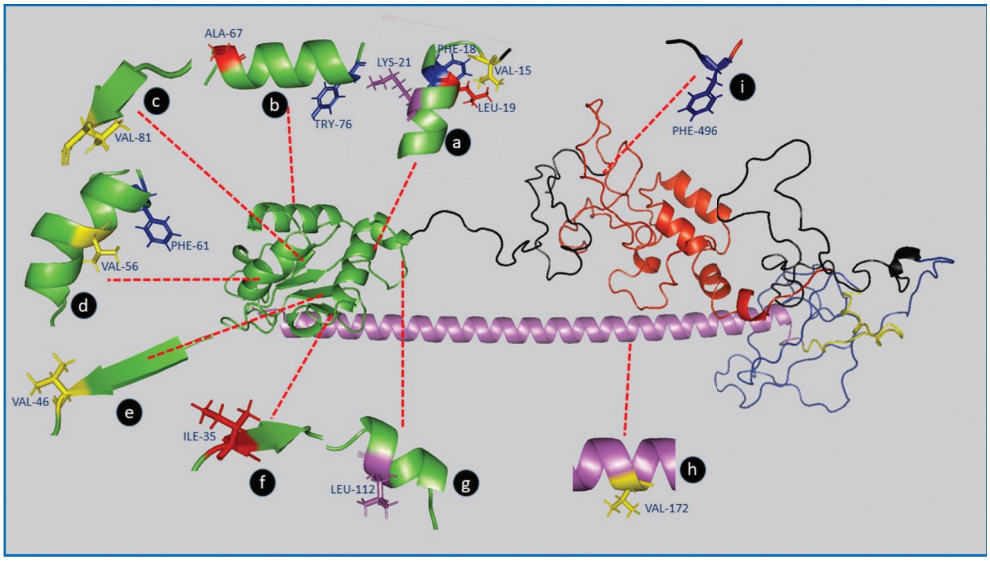

Fig 2. The 3-D structure of the bovine HSF1 protein generated by RaptorX; DNA binding domain (green), N-terminal oligomerization domain HR-A/B *(magenta),* D domain (yellow), transactivation domain *(red),* regulatory domain *(blue),* disordered regions *(black).* The locations of deleterious amino acid substitutions are shown on the structure; **a-** The three mutations (F18L, L19R, K21M) on the first α-helix of the DBD and a mutation (V15G) on the coil, **b-** The two mutations (A67D, Y76D) on the fourth α-helix of the DBD, **c-** A mutation (V81G) on the third beta-sheet of the DBD, **d-** Two mutations (V56G, F61L) in the third α-helix of the DBD, **e-** A mutation (V46E) on the second beta-sheet of the DBD, **f-** The mutation (I35T) on the first beta-sheet of the DBD, **g-** The mutation (L112P) on the last α-helix of the DBD, **h-** The mutation (V172E) on the N-terminal oligomerization domain, **i-** The mutation (F496C) on the disordered region

result in increased protein stability (DDG= 0.18). Since the DDG values predicted by the I-Mutant (DDG= -0.14) and Mupro (DDG= 0.18) programs for the K21M were less than 0.5, it was accepted as neutral *(Table 2).*

The fasta sequence of the HSF1 protein was analyzed on the Consurf server. Of 14 amino acid residues, twelve (V15G, F18L, L19R, I35T, V46E, V56G, F61L, A67D, V81G, L112P, V172E, F496C) were categorized as structural residues, and two (K21M, Y76D) as functional residues *(Fig. 1).* It was determined that 13 (V15G, F18L, L19R, K21M, I35T, V46E, V56G, F61L, A67D, Y76D, V81G, L112P, V172E) of them had a high conservation score and one (F496C) of them had an average conservation score. According to the Consurf analysis results, the most conserved region of the HSF1 protein is the DBD region. Out of 106 residues in the DBD region, 93 residues (87.74%) are in the range of 7-9 conserved scores *(Fig. 1).*

Hope Result and 3-D Structure Prediction

The fourteen amino acid substations were also submitted for the HOPE project analysis. According to the hope results it was found that, out of the 14 mutations, 10 have altered hydrophobicity, 6 differed in charge, and all mutated amino acids differed in size *(Table 2).* Hope server determined that the L112P mutation will disrupt the last α-helix of the DBD *(Fig. 3).* The structural effect of this mutation was examined using the Swiss PDB viewer. It was observed that the H-bond formed by L122 with E109 was not formed by P122. The proline residue can disrupt the α-helix in the absence of the hydrogen bond (P122 with E109), and this can have serious effects on the structure and function of the protein.

The HSF1 protein fasta sequence (525 amino acids) from the Ensemble genome browser was the input for the RaptorX server. The five 3-D models were generated by RaptorX. The obtained structures were assessed according to Ramachandran Plot and MolProbity score in the Swiss-Model server. The 3-D model of the HSF1 protein with the best score (MolProbity Score: 3.26, Ramachandran Favored: 90.06) was visualized using the PyMOL *(Fig. 2).* The positions of the amino acid mutations predicted to be deleterious were shown in this model *(Fig. 2).*

Fasta sequence of the DBD (between 15-200 amino acids) was submitted as the input file for Swiss-Model. The human HSF1 protein (PDB ID: 5D5U.1.b) having sequence coverage of 80% and sequence identity of 98.11% was selected as a template for bovine HSF1 protein and the 3-D model was constructed automatically in Swiss-Model. The amino acid substitutions in DBD were evaluated using this model in Swiss PDB viewer *(Fig. 3, Fig. 4).*

Discussion

The HSF1, the main regulator of the HSPs expression, plays an important role in cell survival under stress [30]. The HSPs are the major molecular chaperones that modulate

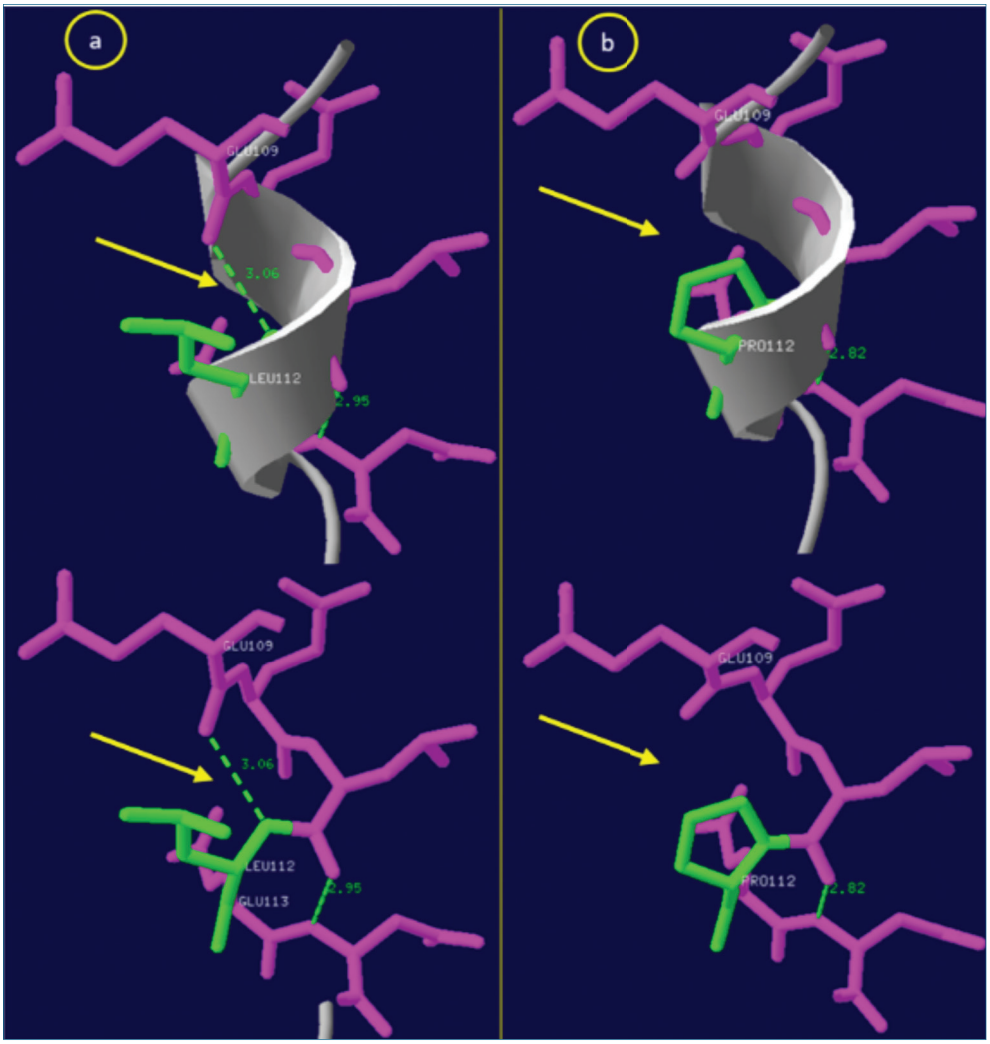

Fig 3. Structural effect of the L112P mutation on the native structure of HSF1 protein visulated by Swiss PDB Viewer. The backbone, sidechain and ribbon presentation structures are presented in the upper pictures, and the backbone and sidechain structures are presented in the lower pictures. H-bonds are indicated by green discontinuous lines. **a**- In the native HSF1 protein, L112 forms two H-bonds with E109 (3.06 Å) and E113 (2.95 Å), **b**- In the mutant HSF1 protein, P122 forms one H-bond with E113 (2.82 Å)

protein folding. They maintain protein homeostasis under stress factors such as increasing temperatures, infectious diseases, and heavy metals [31,32]. The expression of the HSF1 gene is increased in cattle under heat stress [30]. It has been suggested that the HSF1 gene expression level may be an indicator of thermotolerance in cattle [4]. Moreover, SNPs in the HSF1 gene are associated with susceptibility to heat stress in cattle [5,33].

SNPs cause alteration of a single base pair in both coding and non-coding regions. SNPs in non-coding regions can affect gene expression at the transcriptional and posttranscriptional levels. nsSNPs alter amino acid sequences, affect protein structure and function, and have potentially deleterious effects [34]. Today, thanks to rapidly developing sequencing technologies, there are many nsSNPs in variation databases. However, determining the effects of nsSNPs by molecular genetic experiments is a time-consuming and laborious task [35]. The fastest and cheapest way to predict the potential consequences of an nsSNP is to perform bioinformatics analysis [36,37]. Using multiple bioinformatics tools that evaluate different parameters helps to identify results with higher confidence levels [38].

In this study, 13 bioinformatics tools (PredictSNP, MAPP, PhDSNP, PolyPhen-1, PolyPhen-2, SIFT, SNAP, nsSNPAnalyzer, PANTHER, Consurf and Hope project, I-Mutant, MUpro) were used for the prediction of deleterious nsSNPs in the HSF1 gene. Fourteen nsSNPs were predicted to be deleterious SNPs by these tools. Of these nsSNPs, twelve are in the DBD, one in the N-terminal oligomerization domain, and one in the disordered region *(Table 1)*. After activation by stress, HSF1 trimerizes and binds to HSEs in the promoter regions of HSPs. DBD, the most conserved region of HSF1,

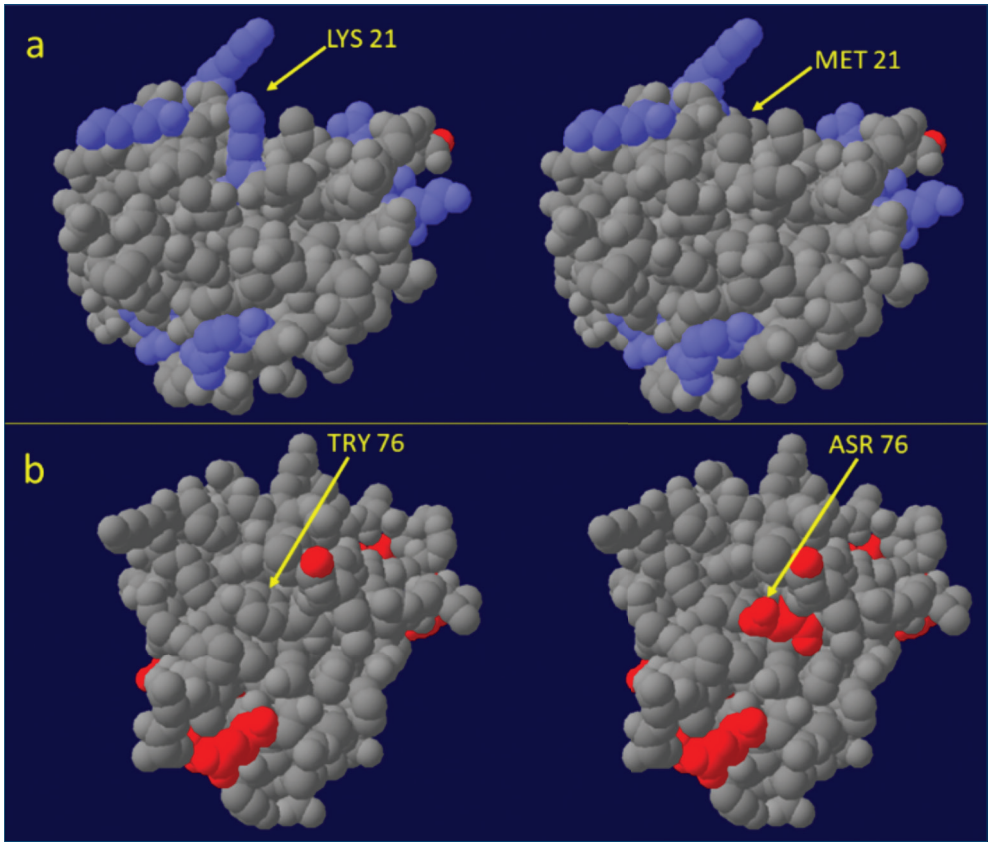

Fig 4. The 3-D surface structure of the DNA binding domain of the bovine HSF1 protein. K21M and Y76D mutations cause charge changes on the protein surface. **a**- Positively charged amino-acid residues are shown in blue, **b**- Negatively charged amino-acid residues are shown in red

plays a major role in recognizing and binding to HSEs [12]. The N-terminal oligomerization domain is responsible for the trimerization of HFS1. The binding affinity of the HSF1 trimer to HSE is significantly increased relative to the monomer HSF1 [39]. The disordered regions that remain unstructured have important roles in determining the function and structure of the protein [40]. Thus, the mutations in these three domains may reduce the affinity of HSF1 to HSE and result in decreased stability of HSF1 structure [41].

Evolutionarily conserved protein residues are the most important parts of protein folding, function, and structure. Mutations in conserved amino acids can cause changes in the 3-D structure of the protein and its interaction with other molecules [42]. Therefore, evolutionary conservation analysis was done with the Consurf server. All twelve nsSNPs predicted by the bioinformatic tools to be potentially deleterious in the DBD were found to have high conservation (scores=9) *(Table 1)*. Two (K21 and Y76) of them were categorized as functional residues (exposed) and ten (V15, F18, L19, I35, V46, V56G, F61, A67, V81, L112, V172, F496) as structural residues (buried). Conservation scores for the residues in the N-terminal oligomerization domain (V172) and disordered region (F496) residues were estimated as 8 and 6, respectively. These results are similar to previous studies [42,43], which suggest that nsSNPs that cause substitution in evolutionarily conserved residues may change the function of the protein.

For a protein to perform its functions efficiently, it must be folded correctly. Interactions between hydrophobic amino acids have an important role in accurately folding a protein chain [44]. The Hope Project server determined that wild type 9 amino acids (V15G, L19R, I35T, V46E, V56G, A67D, Y76D, V81G, V172E) in the HSF1 protein are more hydrophobic than the mutant type. Seven of these residues are buried and one (Y76D) is on the surface of the protein. These mutations will cause a loss of hydrophobic interactions either in the core of the protein or on the surface. It is known that hydrophobic residues on the surface and core of the protein support protein stability [45].

Prolines are known to have a very rigid structure, sometimes forcing the backbone in a specific conformation [46] and it may disturb the α-helix structure [47]. L112P mutation in the last α-helix of the DBD in HSF1 may cause a significant deterioration in the function and structure of the protein by disrupting this α-helix structure *(Fig.*

3). It is also predicted that mutation results in loss of interaction in the core of the protein because proline is smaller than leucine. When the mutant residue is smaller than the wild type residue, it causes an empty space in the core of the protein [48].

For the exposed Y76D variation, tyrosine is more hydrophobic than aspartic acid. In addition, tyrosine (Y76-mutant type) is neutral and aspartic acid (D76-wild type) is negatively charged *(Fig. 4)*. This mutation can both decrease protein stability and significantly hamper interaction with other molecules. Another amino acid on the surface, lysine (K21-wild type) is positively charged, while methionine (M21-mutant type) is neutral *(Fig. 4)*. The difference in charge disrupts the ionic attractions that are important for protein structure and function [49]. For these two exposed variations (K21M, Y76D), mutant residues are smaller than native residues. Smaller residues may cause a loss of external interactions [50]. Considering these physicochemical changes, K21M and Y76D mutations can cause a significant decrease in binding HSF1 to HSEs.

The results of this study suggest that the 14 nsSNPs in the HSF1 in cattle might represent associated with heat stress susceptibility. The deleterious nsSNPs on the biological function of the HSF1 protein provide a starting point for genetic marker-assisted selection against heat stress of cattle. These findings need to be validated by performing wet-lab experiments.

Availability of Data and Materials

Datasets analyzed during the current study are available in the author on reasonable request.

Acknowledgments

The author would like to thank Assoc. Prof. Dr. Süleyman Kök for his helpful advice on various technical issues examined in this paper.

Funding Support

There is no specific grant funding source

Competing Interest

The author declare that have no conflict of interest.

References

1. Mishra SR: Significance of molecular chaperones and micro RNAs in acquisition of thermo-tolerance in dairy cattle. *Anim Biotechnol*, 1-11, 2020. DOI: 10.1080/10495398.2020.1830788

2. Strandén I, Kantanen J, Russo IRM, Orozco-terWengel P, Bruford MW, the Climgen Consortium: Genomic selection strategies for breeding adaptation and production in dairy cattle under climate change. *Heredity,* 123 (3): 307-317, 2019. DOI: 10.1038/s41437-019-0207-1

3. Tesema Z, Taye M, Ayichew D: The role of phenotypic and genetic basis of livestock selection for climate change adaptation and mitigation: A review. *J Appl Adv Res,* 4 (2): 66-77, 2019. DOI: 10.21839/jaar.2019.v4i2.251

4. Rocha RFB, Baena MM, de Cássia Estopa A, Gervásio IC, Ibelli AMG, Gionbelli TRS, Gionbelli MP, de Freitas RTF, Meirelles SLC: Differential expression of HSF1 and HSPA6 genes and physiological responses in Angus and Simmental cattle breeds. *J Therm Biol,* 84, 92-98, 2019. DOI: 10.1016/j.jtherbio.2019.06.002

5. Rong Y, Zeng M, Guan X, Qu K, Liu J, Zhang J, Chen H, Huang B, Lei C: Association of HSF1 genetic variation with heat tolerance in Chinese cattle. *Animals,* 9 (12): 1027, 2019. DOI: 10.3390/ani9121027

6. Bagath M, Krishnan G, Devaraj C, Rashamol VP, Pragna P, Lees AM, Sejian V: The impact of heat stress on the immune system in dairy cattle: A review. *Res Vet Sci,* 126, 94-102, 2019. DOI: 10.1016/j.rvsc.2019.08.011

7. Scharf B, Carroll JA, Riley DG, Chase Jr CC, Coleman SW, Keisler DH, Weaber RL, Spiers DE: Evaluation of physiological and blood serum differences in heat-tolerant (Romosinuano) and heat-susceptible (Angus) *Bos taurus* cattle during controlled heat challenge. *J Anim Sci,* 88 (7): 2321-2336, 2010. DOI: 10.2527/jas.2009-2551

8. Balasubramaniam B, Balamurugan K: Role of Heat Shock Factors in Diseases and Immunity. **In,** Asea AAA, Kaur P (Eds): Heat Shock Proteins in Human Diseases. 197-210, Springer, Cham, 2020. DOI: 10.1007/7515_2020_21

9. Zeng L, Qu K, Zhang J, Huang B, Lei C: Genes related to heat tolerance in cattle-A review. *Anim Biotechnol*, 1-9, 2022. DOI: 10.1080/10495398.2022.2047995

10. Onasanya GO, Thiruvenkadan AK, Sreekumar C, Okpeku M, Tirumurugaan GK, Msalya GM, Sanni MT, Olowofeso O, Fafiolu AO, Decampos JS, Obadire FO, Abdullahi AA, Ikeobi CON: Molecular characterization of hsp 70 gene using single nucleotide polymorphism in nigerian breeds of zebu cattle. *Fuw Trends Sci Technol J,* 4, 714-720, 2019.

11. Jaiswal L, De S, Singh RK, Baithalu RK: Molecular characterization and protein structure prediction of heat shock transcriptional factors in goat *(Capra hircus)* and sheep *(Ovis aries). Anim Biotechnol,* 31 (5): 432-439, 2020. DOI: 10.1080/10495398.2019.1615497

12. Barna J, Csermely P, Vellai T: Roles of heat shock factor 1 beyond the heat shock response. *Cell Mol Life Sci,* 75 (16): 2897-2916, 2018. DOI: 10.1007/s00018-018-2836-6

13. Kovács D, Sigmond T, Hotzi B, Bohár B, Fazekas D, Deák V, Vellai T, Barna J: HSF1Base: A comprehensive database of HSF1 (heat shock factor 1) target genes. *Int J Mol Sci,* 20 (22): 5815, 2019. DOI: 10.3390/ijms20225815

14. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Fioretto LDR, Davidson C, Dodiya K, El Houdaigui B, Fatima R, Gall A, Giron CG, Grego T, Guijarro-Clarke C, Haggerty L, Hemrom A, Hourlier T, Izuogu OG, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Martinez JG, Marugán JC, Maurel T, McMahon AC, Mohanan S, Moore B, Muffato M, Oheh DN, Paraschas D, Parker A, Parton A, Prosovetskaia I, Sakthivel MP, Abdul Salam AI, Schmitt BM, Schuilenburg H, Sheppard D, Steed E, Szpak M, Szuba M, Taylor K, Thormann A, Threadgold G, Walts B, Winterbottom A, Chakiachvili M, Chaubal A, De Silva N, Flint B, Frankish A, Hunt SE, IIsley GR, Langridge N, Loveland JE, Martin FJ, Mudge JM, Morales J, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Cunningham F, Yates AD, Zerbino DR, Flicek P: Ensembl 2021. *Nucleic Acids Res,* 49 (D1): D884-D891, 2021. DOI: 10.1093/nar/gkaa942

15. Anckar J, Sistonen L: Regulation of HSF1 function in the heat stress response: Implications in aging and disease. *Annu Rev Biochem,* 80, 1089-1115, 2011. DOI: 10.1146/annurev-biochem-060809-095203

16. Sharma C, Seo YH: Small molecule inhibitors of HSF1-activated pathways as potential next-generation anticancer therapeutics. *Molecules,* 23 (11): 2757, 2018. DOI: 10.3390/molecules23112757

17. Gomez-Pastor R, Burchfiel ET, Thiele DJ: Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat Rev Mol Cell Biol,* 19 (1): 4-19, 2018. DOI: 10.1038/nrm.2017.73

18. Hentze N, Le Breton L, Wiesner J, Kempf G, Mayer MP: Molecular mechanism of thermosensory function of human heat shock transcription factor Hsf1. *eLife*, 5:e11576, 2016. DOI: 10.7554/eLife.11576

19. Badri TM, Chen KL, Alsiddig MA, Li L, Cai Y, Wang GL: Genetic polymorphism in *Hsp90AA1* gene is associated with the thermotolerance in Chinese Holstein cows. *Cell Stress Chaperones,* 23 (4): 639-651, 2018. DOI:

10.1007/s12192-017-0873-y

20. Malheiros JM, Enríquez-Valencia CE, da Silva Duran BO, de Paula TG, Curi RA, de Vasconcelos Silva JAII, Dal-Pai-Silva M, de Oliveira HN, Chardulo LAL: Association of *CAST2*, *HSP90AA1*, *DNAJA1* and *HSPB1* genes with meat tenderness in Nellore cattle. *Meat Sci,* 138, 49-52, 2018. DOI: 10.1016/j.meatsci.2018.01.003

21. Shergojry SA, Ganayi BA, Ramesha KP, Rengarajan K, Srihari GV, Das DN, Kataktalware MA: Association of single nucleotide polymorphisms (SNPS) of HSP90AA1 gene with reproductive traits in Deoni cattle. *Int J Livest Res,* 1 (1): 17-29, 2011. DOI: 10.5455/ijlr.20120204082821

22. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, Brezovsky J, Damborsky J: PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *Plos Comp Biol,* 10 (1):e1003440, 2014. DOI: 10.1371/journal.pcbi.1003440

23. Mueller SC, Backes C, Haas J, the INHERITANCE Study Group, Katus HA, Meder B, Meese E, Keller A: Pathogenicity prediction of non-synonymous single nucleotide variants in dilated cardiomyopathy. *Brief Bioinform,* 16 (5): 769-779, 2015. DOI: 10.1093/bib/bbu054

24. Capriotti E, Fariselli P, Casadio R: I-Mutant2. 0: Predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res,* 33 (Suppl. 2): W306-W310, 2005. DOI: 10.1093/nar/gki375

25. Cheng J, Randall A, Baldi P: Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins Struct Funct Bioinf,* 62 (4): 1125-1132, 2006. DOI: 10.1002/prot.20810

26. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N: ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res,* 38 (Suppl. 2): W529-W533, 2010. DOI: 10.1093/nar/gkq399

27. Venselaar H, Te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G: Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinform,* 11 (1): 1-10, 2010. DOI: 10.1186/1471-2105-11-548

28. Wang S, Li W, Liu S, Xu J: RaptorX-Property: A web server for protein structure property prediction. *Nucleic Acids Res,* 44 (W1): W430-W435, 2016. DOI: 10.1093/nar/gkw306

29. Guex N, Peitsch MC: SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis,* 18 (15): 2714-2723, 1997. DOI: 10.1002/elps.1150181505

30. Mohapatra S, Kundu AK, Mishra SR, Senapati S, Jyotiranjan T, Panda G: HSF1 and GM-CSF expression, its association with cardiac health, and assessment of organ function during heat stress in crossbred Jersey cattle. *Res Vet Sci,* 139, 200-210, 2021. DOI: 10.1016/j.rvsc.2021.07.018

31. Wankhede NL, Kale MB, Upaganlawar AB, Taksande BG, Umekar MJ, Behl T, Abdellatif AAH, Bhaskaran PM, Dachani SR, Sehgal A, Singh S, Sharma N, Makeen HA, Albratty M, Dailah HG, Bhatia S, Al-Harrasi A, Bungau S: Involvement of molecular chaperone in protein-misfolding brain diseases. *Biomed Pharmacother,* 147:112647, 2022. DOI: 10.1016/j.biopha.2022.112647

32. Höhfeld J, Benzing T, Bloch W, Fürst DO, Gehlert S, Hesse M, Hoffmann B, Hoppe T, Huesgen PF, Köhn M, Kolanus W, Merkel R, Niessen CM, Pokrzywa W, Rinschen MM, Wachten D, Warscheid B: Maintaining proteostasis under mechanical stress. *Embo Rep,* 22 (8):e52507, 2021. DOI: 10.15252/embr.202152507

33. Li QL, Ju ZH, Huang JM, Li JB, Li RL, Hou MH, Wang CF, Zhong JF: Two novel SNPs in *HSF1* gene are associated with thermal tolerance traits in Chinese Holstein cattle. *DNA Cell Biol,* 30 (4): 247-254, 2011. DOI: 10.1089/dna.2010.1133

34. Lee YS, Won K, Shin D, Oh JD: Risk prediction and marker selection in nonsynonymous single nucleotide polymorphisms using whole genome sequencing data. *Anim Cells Syst,* 24 (6): 321-328, 2020. DOI: 10.1080/19768354.2020.1860125

35. Kumar A, Rajendran V, Sethumadhavan R, Shukla P, Tiwari S, Purohit R: Computational SNP analysis: Current approaches and future prospects. *Cell Biochem Biophys,* 68 (2): 233-239, 2014. DOI: 10.1007/s12013-013-9705-6

36. Domené S, Scaglia PA, Gutiérrez ML, Domené HM: Applying bioinformatic platforms, *in vitro,* and *in vivo* functional assays in the characterization of genetic variants in the GH/IGF pathway affecting growth and development. *Cells,* 10 (8): 2063, 2021. DOI: 10.3390/cells10082063

37. Oliver GR, Zimmerman MT, Klee EW, Urrutia RA: "The molecule's the thing:" The promise of molecular modeling and dynamic simulations in aiding the prioritization and interpretation of genomic testing results. *F1000Res,* 5:766, 2016. DOI: 10.12688/f1000research.8600.3

38. Gassoum A, Abdelraheem NE, Elsadig N: Comprehensive analysis of rsSNPs associated with hypertension using *in-silico* bioinformatics tools. *Open Access Lib J,* 3 (7): 1, 2016. DOI: 10.4236/oalib.1102839

39. Feng N, Feng H, Wang S, Punekar AS, Ladenstein R, Wang DC, Zhang Q, Ding J, Liu W: Structures of heat shock factor trimers bound to DNA. *iScience,* 24 (9):102951, 2021. DOI: 10.1016/j.isci.2021.102951

40. Alshehri MA, Manee MM, Al-Fageeh MB, Al-Shomrani BM: Genomic analysis of intrinsically disordered proteins in the genus *Camelus*. *Int J Mol Sci,* 21 (11): 4010, 2020. DOI: 10.3390/ijms21114010

41. Bridges TM, Scheraga RG, Tulapurkar ME, Suffredini D, Liggett SB, Ramarathnam A, Potla R, Singh IS, Hasday JD: Polymorphisms in human heat shock factor-1 and analysis of potential biological consequences. *Cell Stress Chaperones,* 20 (1): 47-59, 2015. DOI: 10.1007/s12192-014-0524-5

42. Wang Z, Huang C, Lv H, Zhang M, Li X: *In silico* analysis and high-risk pathogenic phenotype predictions of non-synonymous single nucleotide polymorphisms in human *Crystallin beta A4* gene associated with congenital cataract. *Plos One,* 15 (1):e0227859, 2020. DOI: 10.1371/journal.pone.0227859

43. Ghani U, Ullah R, Anjum T, Ali Y, Hussain S, Shah S, Zulfiqar I, Sultan T, Tanveer MA, Sabar MF: Identification of most lethal non-synonymous SNPs in tollip gene-An *in-silico* analysis. 2021. DOI: 10.21203/rs.3.rs-488110/v1

44. Mirzaie M: Hydrophobic residues can identify native protein structures. *Proteins Struct Funct Genet,* 86 (4): 467-474, 2018. DOI: 10.1002/prot.25466

45. Kumar A, Biswas P: Effect of site-directed point mutations on protein misfolding: A simulation study. *Proteins Struct Funct Genet,* 87 (9): 760-773, 2019. DOI: 10.1002/prot.25702

46. Sudarshan S, Kumar M, Kaur P, Kumar A, Sethuraman G, Sapra S, Gulati S, Gupta N, Kabra M, Roy Chowdhury M: Decoding of novel missense TSC2 gene variants using *in-silico* methods. *BMC Med Genet,* 20 (1): 1-9, 2019. DOI: 10.1186/s12881-019-0891-y

47. Agrahari AK, Kumar A, Siva R, Zayed H, Doss CGP: Substitution impact of highly conserved arginine residue at position 75 in *GJB1* gene in association with X-linked Charcot–Marie-tooth disease: A computational study. *J Theor Biol,* 437, 305-317, 2018. DOI: 10.1016/j.jtbi.2017.10.028

48. Elnasri HA, Al Bkrye AM, Khaier MAM: *In silico* analysis of non synonymous SNPs in *DHCR7* gene. *Am J Bioinform Res*, 8 (1): 12-18, 2018.

49. Sedaghat A, Zamani M, Jahanshahi A, Ghaderian SB, Shariati G, Saberi A, Hamid M, Aminzadeh M, Galehdari H: Frequent novel mutations are causative for maple syrup urine disease from Southwest Iran. *Meta Gene,* 16, 96-104, 2018. DOI: 10.1016/j.mgene.2018.01.010

50. Abdelhalim DM, Albkrye AM, Salih MA, Abodlaa AEE, Elnasri HA, Khaier MAM: *In silico* analysis of a single nucleotide polymorphism (SNPS) in human GSTM1 gene associated with cancer development. *European J Biomed Pharm Sci,* 7 (1): 514-521, 2020.