



## ROLE OF FUNCTIONAL METABOLITES IN HEAT STRESS RESPONSES OF *SORGHUM BICOLOR* (L) MOENCH., AN AGROFORESTRY COMPATIBLE CROP

<sup>1</sup>Nwogwugwu, J.O., <sup>2</sup>Time, I. and <sup>3</sup>Adewale, B. D.

<sup>1</sup>Department of Forest Conservation and Protection,  
Forestry Research Institute of Nigeria, Ibadan, Nigeria

<sup>2</sup>Department of Crop and Environmental Protection,  
University of Agriculture, Makurdi, Nigeria

<sup>3</sup>Department of Crop Science and Horticulture,  
Federal University Oye-Ekiti, Ikole-Ekiti Campus, Ekiti State, Nigeria  
Corresponding Authors' email: [joyoluchi70@gmail.com](mailto:joyoluchi70@gmail.com)

### Abstract

Plant response to stress is critical and can be harnessed for greater yield increases in productive agriculture. Agriculture is susceptible to climate variability which impacts largely on food security by changing the balance among the key determinants of crops and yield. Enormous genetic diversity for stress tolerance such as Heat Shock Protein (HSP) genes exist in this plant species. Recent advances in sequencing technologies have rapidly unfolded the molecular genetic basis for heat stress responses in plants. In this regard, the role of heat-responsive HSP70 genes in *Sorghum bicolor* is investigated. The genome sequence of *S. bicolor*, retrieved from Heat Shock Protein Information Resource (HSPiR), database was studied using bioinformatics tools. Sequence analysis revealed 11 and 12 pairs of paralogous, for HSP70, indicating ancestral duplication and the crucial role in the expansion of *S. bicolor* HSP gene family. Hydropathy pattern of proteins sequence suggested that the protein might belong to hydrolase or transferase class and the prediction of disorder region of HSP70. The benefits derivable by integrating this species into Agri-silviculture due to its adaption tendencies under adverse climatic changes are also highlighted.

**Keywords:** Heat stress response, bioinformatics tools, agri-silviculture, gene family, and protein

### Introduction

*Sorghum* [*Sorghum bicolor* (L.) Moench] is a multipurpose cereal crop grown in drought-prone areas and offers benefits as food and feed, shelter, feedstock and bioenergy (Emendack *et al.*, 2018). It is ranked as the fifth most important cereal crop cultivated worldwide. In West Africa, Nigeria is the largest sorghum producer, accounting for about 71% of the total regional (West Africa) sorghum output (Ogbonna, 2011). In Nigeria, sorghum is grown mainly in the northern part of the country as primary food crop (USAID, 2011). Potential yield of this crop is constrained by various biotic and abiotic stresses at different crop developmental stages. Climate change, global warming and desertification are major environmental challenges affecting biodiversity and food security in sub-Saharan Africa (SSA). Therefore, with these obvious challenges; to understand and

identify cropping patterns in terms of species adaptability and management practices for optimizing resources that enhance climate change adaptation (FAO, 2010), requires advanced approaches and choice of species that minimize production risks under a given environment and season. Agroforestry systems offer great potential of increasing food production and conserving biodiversity; it plays a critical role in sustainable agriculture by multi land use systems (George *et al.*, 2012). Integrating drought resistant species such as *Sorghum* into agri-silviculture can sustainably enhance soil productivity, thereby ensuring both environmental protection and food security needs (ICRAF, 2000). Agroforestry land use management requires careful selection of tree-crop species that will meet both socioeconomic and environmental needs of the people. (German *et al.*, 2006). *Sorghum* is a crop attested for high yield and drought resistance in the face

of changing climate (FAOSTAT, 2017; Albuquerque *et al.*, 2011).

Alterations in environmental conditions often lead to abiotic stress factors that constrain crop performance and sometimes favour plant pests like pathogens resulting to yield reductions. Most plants are known to possess genes that help them adapt to diverse stress factors in their environment; therefore, identifying genes in plants that are responsible for their adaptation to certain abiotic stress factors will guarantee good crop performance. Advancement in molecular biology has helped in identification and characterization of many high temperature responsive genes(). Sorghum just like other plants has been reported to respond to elevated temperatures and many other abiotic stresses (Soransen *et al.*, 2003). Expression of heat shock proteins (HSPs) is strongly induced in sorghum (Pavli *et al.*, 2011). Most HSPs are molecular chaperones that are involved in the stabilization and refolding of denatured proteins(). The evolutionary conservation of the heat shock response and the HSPs, along with the correlation of HSP expression with cellular resistance to high temperature, have led to the long standing hypothesis that, HSPs protect cells from the detrimental effects of high temperature().

HSP70 makes up a set of prominent cellular machinery that assists with a wide range of protein folding processes in almost all cellular compartments. It has vital functions in preventing aggregation and in assisting refolding of non-native proteins under both normal and stress conditions(). They are also involved in protein import, translocation processes, and in facilitating the proteolytic degradation of unstable proteins by targeting the proteins to lysosomes or proteasomes. In addition, some members of HSP70 control the biological activity of protein folding(), while small HSPs are synthesized ubiquitously in prokaryotic and eukaryotic cells in response to heat and other stresses, some small HSPs are expressed during certain developmental stages(). Based on sequence and structure similarities, computational analysis has been performed between small HSP and HSP70 in sorghum in this study. Sequence and structure of both proteins were compared using bioinformatics tools. The study was proposed to identify a functional correlation between these two proteins that could play significant role in protecting plants against heat stress by re-establishing normal protein conformation and establish cellular homeostasis, which will increase crop yield. Families of heat-inducible chaperones, their interaction with each other and how they function to protect organisms from stresses were equally investigated.

## **Methodology**

### **Source of sequence data, sequence processing and analysis**

The genome protein sequences of *Sorghum bicolor* HSP70 were retrieved from Heat Shock Protein Information Resource (HSPiR) Database (<http://pdslab.biochem.iisc.ernet.in/hspir/>). Thirty-four

sequences of HSP70 were identified. The SOBI2-HSP70\_1103 sequence which gave the highest identities in multiple sequence alignment using CLC Sequence Viewer 7.6 (<https://www.qiagenbioinformatics.com/.../clc-sequence>) was selected. Blast into the National Centre for Biotechnology Information (NCBI) was used to identify homologous sequences of HSPs, while percent similarities and identities were computed by NCBI blast2seq program.

### **Function prediction of sorghum for HSP70 genes**

To comprehend the functional relationship between small HSP and HSP70 Protein, functional family prediction tools (SVMProt) ([www.bidd2.nus.edu.sg/cgi-bin/svmprot/svmprot.cgi](http://www.bidd2.nus.edu.sg/cgi-bin/svmprot/svmprot.cgi)) were used(). The approach classifies a protein sequence into functional family mainly based on molecular function, biological process and cellular localization. A useful method for such classification used was the Support Vector Machine (SVM)–().

### **Prediction of disordered regions and Protein-protein interaction in sorghum HSP genes**

Disorder regions were searched in the amino acid sequence for HSP70 using DISO-clust server (<http://www.reading.ac.uk/bioinf/DISOclust>). Furthermore, queried protein interactions with other proteins in sorghum was analysed using STRING (<https://string-db.org>).

### **Hydropathy analysis of sorghum HSP70 protein sequences**

In protein sequences, the hydrophobic properties of the amino acids govern the protein structure and fold(;). Amino-acid sequences of HSP70 genes in sorghum were computed using the Prot-Scale Tool (<https://web.expasy.org/protscale/pscale/Hphob.Doolittle.html>).

### **Conserved Motif Analyses of sorghum for HSP70 genes and its phylogenetic Classification**

A comprehensive web-based tool for mining sequence motifs in proteins, DNA, and RNA, MEME Suite 4.11.0 (<http://meme-suite.org>) was used to analyse conserved motifs(). The maximum and minimal motif width, and the maximum number of motifs were set to 50, 6, and 10 respectively. The phylogenetic tree was constructed using 34 genes to the crop's HSP70 to identify the evolutionary distinction of domain structure in small HSP and HSP70 proteins through neighbour-joining (NJ) method.

### **Identification of Domain Conservation in sorghum HSP70**

To check reliability of the phylogeny, motif compositions were examined using MEME tools (<http://meme-suite.org/tools/meme>) and motifs of HSP70 proteins were determined.

## Results and Discussion

### *Sequence homology and hypothetical protein functions of sorghum HSP70*

The Sequence scanning of sorghum HSP70 genes revealed a few homologous heat shock proteins. Crystal structures of the DnaK Chaperone protein sequences from *Geobacillus kaustophilus* (HTA426) in Post-ATP hydrolysis state and those from *Mycobacterium tuberculosis* (H37Rv) with 62% and 58% identities respectively, were found for HSP70 (Table 1). The functional domain of sorghum HSP70 had a similar modular function with different value in terms of percentage. Analysis of the biological process and molecular function in figures 1 & 2, revealed approximately 21% protein folding, 20% metallo-sulfur cluster assembly and iron-sulfur cluster assembly, 15% response to stresses, 8% ATP metabolic process, 5% protein metabolic process, 4% nucleotide metabolic process gene, 3% response to topologically incorrect and 3% cell morphogenesis out of their whole biological response against the stresses faced by the crop. Protein function prediction (PFP) has been constructed for sorghum HSP70. In the present study, a significant coverage was found over cytoplasm followed by nucleus, membrane-bounded, microtubule and plasma membrane (Figure 3).

### *Determining the disordered regions and Protein-Protein Interaction in sorghum HSP70*

The confidence level for protein-protein interaction in Figure 4 and STRING results in Figure 5 showed the red node as the query protein having function with other ten proteins, but not necessarily physically bound with each other. The thickness of lines connecting nodes showed edge confidence, which is highest (0.900) in case of 18 edges and high (0.700) in case of 3 edges on right side. The seven proteins in this interaction including query protein are involved in ATP binding, unfolded protein binding and chaperone-mediated protein folding, hence, confirming the shared function of these proteins.

### *Hydropathy analysis of sorghum HSP70 protein sequences*

The hydropathy pattern of the proteins sequence gave 20 hits with score < 5. The lowest three hits suggested that the protein might belong to hydrolase or transferase class in which the side chains possess a spectrum of functional groups (Figure 6). However, most have at least one atom (nitrogen, oxygen or sulfur / sulphur – which is America?) with electron pairs available for hydrogen bonds to water and other molecules.

### *Phylogenetic Classification of sorghum HSP70 and identification of its domain conservation*

The phylogenetic tree constructed through neighbour-joining method, classified all HSP70 into three main clusters. Eleven and 12 pairs of paralogous genes were identified in *Sorghum bicolor* HSP70 (Figure 7). Motif compositions of HSP70 in sorghum revealed 10 distinct motifs. The length of motif searched was in window of 28 to 50 for HSP70 (Table 2).

HSP70 are chaperones expressed in response to heat stress in diverse plants species (). These play a crucial role in maintaining the protein quality control in the plant, and share a conserved 90-amino-acid C-terminal domain called the  $\alpha$ -crystallin domain (). The small HSPs are not able to refold non-native proteins, and have a high capacity to bind non-native proteins, probably through hydrophobic interaction. They are also not able to stabilize and prevent non-native protein aggregation, thereby, facilitating their subsequent refolding by ATP-dependent chaperones such as the DnaK system(). Many plants respond differently to high temperatures at different stages of growth and some plants are more sensitive to high temperatures than others (). Lee and Vierling (2000) explained the cooperation along HSP70 and small HSP such as HSP18.1 in preventing aggregation of protein from *P. sativum*. HSP70 participates in the guidance complex import that binds to protein precursor to be transferred through the membranes into the organelles, like chloroplast. It also participates in photo protection and the repairing of photosystem II during and after the photo-inhibition. The protein has important role in differentiation of germinating seeds and its tolerance to heat. (). HSP70, plays a crucial role in the development of cross-adaptation to heat stress in plants by various processes(). Analyses in this study revealed that HSP70 halts aggregation and help refold native protein under heat stress condition (). The analysis of schematic pathway (Figure 6) indicated that the N-terminal domain ATP-binding play important role in the hydrolysis of ATP (; ). The substrate binding pocket of HSP70 is in the open conformation; hence, substrate affinity is low (Figure 7) (). Small HSP like HSP40 concomitant with ATP hydrolysis cause a conformational switch, whereby, the substrate binding pocket of HSP70 closes, facilitating the capture of non-native protein substrate (; ). Small HSP, together with HSP70, maintain cellular proteostasis by binding to misfolded polypeptides, and create a synergistic enhancement in the rate of HSP70-mediated ATP hydrolysis. Examination of the mechanism revealed also the synergistic action in the rate of HSP70-mediated ATP. However, after the addition of ATP to HSP70-substrate-ADP complexes, complete substrate dissociation is achieved before stoichiometric ATP hydrolysis can occur—(; ). It appeared that the transition of HSP70 from an ADP-bound to an ATP-bound form is derived by substrate release().

Hydropathy revealed that HSP70 binds to hydrophobic surfaces, preventing adventitious associations and stabilizing target proteins in a fully or partially unfolded state. Also, ATP promotes the release of the substrate, allowing it to continue folding, transport, or assembly(). Predicting a disorder in a protein sequence has been a heated topic in the molecular sciences because these regions are less sensitive to mutations and manipulations can easily be made due to the lack of structure (; ). The phylogenetic analysis classified HSP70 into three main clusters. Bootstrap analysis of 1000 replicates for statistical reliability used in



phylogenetic tree construction allows for identification of putative paralogous and orthologous HSP genes (). Paralogs are genes in the same genome created through gene duplication events (). Within each HSP protein class, 11 of paralogous genes were identified respectively in sorghum HSP70, indicating ancestral duplication, providing information regarding the evolutionary process of HSP genes family in sorghum(). Gene duplication is a major mechanism through which genomic rearrangement and expansion occurs. However, diversification of gene function is also generated during molecular evolution (). Studies have shown that duplication are generally involved in stress responses (Zou *et al.*, 2009; Kondrashov, 2012), suggesting that these duplicates are important for adaptive evolution to rapidly changing environments (; ). Therefore, segmental duplication events in sorghum species expansion could be related to the roles of these genes (). Protein similarities can be identified easily by comparing their three-dimensional structures –(). When such structures are not available, the alignment of related sequences can aid the recognition of distant similarities –(). The definition and construction of these patterns have been studied previously, and have been variously called motifs, profiles, position-specific score matrices, and Hidden Markov Models (). In essence, for each position in the derived pattern, every amino acid is assigned a score().

### Conclusion

*Sorghum bicolor* is adapted to a wide range of environmental conditions, particularly, drought. This study has further validated the adaptability of these species to adverse climatic changes and the need to integrate it into an agroforestry system. It has also demonstrated that under heat stress condition, HSP70 responded through biochemical or physiological adaptation by producing specific proteins that dealt with the issue of environmental stresses. Data generated provides a comprehensive insight into the cell metabolism regulated stress induced HSP70, and would enhance understanding of the mechanism by which HSP70 improves abiotic stress tolerance. Further research is necessary to determine the functions of the amino acids lining the active sites. Knowledge of the functions of these amino acids in the active site is critical in their genetic modifications for possible introduction to sorghum species for heat resistance.

### References

Alderson, T. R., Kim, J. H. and Markley, J. L. (2016). Dynamical structures of Hsp70 and Hsp70-Hsp40 complexes. *Structure*, 24 (7): 1014-1030.

Altschul, S. F. and Koonin, E. V. (1998). Iterated profile searches with PSI-BLAST—a tool for discovery in protein databases. *Trends in biochemical sciences*, 23(11): 444-447.

Arrigo, A.P. and Müller, W.E.G. (2012). Small Stress Proteins. *Progress in Molecular and Subcellular Biology*, Springer-Verlag New York Inc., New York. 265pp.,

Ashcroft, F. M. (1999). Ion channels and disease:

Academic press. San Diego, California, USA. 481pp.

Brylinski, M. and Feinstein, W. P. (2013). eFindSite: improved prediction of ligand binding sites in protein models using meta-threading, machine learning and auxiliary ligands. *Journal of computer-aided molecular design*, 27(6):551-567.

Cerros-Tlatilpa, R., Columbus, J. T. and Barker, N. P. (2011). Phylogenetic relationships of *Aristida* and relatives (Poaceae, Aristidoideae) based on noncoding chloroplast (trnL-F, rpl16) and nuclear (ITS) DNA sequences. *American Journal of Botany*, 98(11): 1868-1886.

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics (Oxford, England)*, 14(9): 755-763.

Emendack, Y., Burke, J., Haydee, L., Sanchez, J., and Chad, H. (2018). Abiotic Stress Effects on Sorghum Leaf Dhurrin and Soluble Sugar Contents throughout Plant development. *Crop Sci.*, 58:1706–1716.

Esser, C., Alberti, S. and Höhfeld, J. (2004). Cooperation of molecular chaperones with the ubiquitin/proteasome system. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1695(1): 171-188.

FAO (2010). Food and Agriculture Organization. Agriculture Based Livelihood System in Drylands in the context of Climate Change: Inventory of Adaptation Practices and Technologies of Ethiopia: Food and Agriculture Organization, Addis Ababa, Ethiopia.

FAOSTAT (2017). Food and Agricultural Organization ( F A O ) . Retrieved from : <http://www.fao.org/faostat/en/#data/QC>. Accessed on Sat May 04 14:9:42 EAT 2019.

George, S.J., Harper, R.J., Hobbs, R.J., Tibbett, M. (2012). A sustainable agricultural landscape for Australia: A review of interlacing carbon sequestration, biodiversity and salinity management in agroforestry systems. *Agriculture, Ecosystems and Environment*, 163:28-36.

German, L. A., Kidane, B. and Shemdoe, R. (2006). Social and environmental trade-offs in tree species selection: a methodology for identifying niche incompatibilities in agroforestry. *Environment, Development and Sustainability*, 8(4): 535–552.

Goloubinoff, P. and De Los Rios, P. (2007). The mechanism of Hsp70 chaperones:(entropic) pulling the models together. *Trends in biochemical sciences*, 32(8): 372-380.

Hanada, K., Zou, C., Lehti-Shiu, M. D., Shinozaki, K. and Shiu, S.H. (2008). Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant physiology*, 148(2): 993-1003.

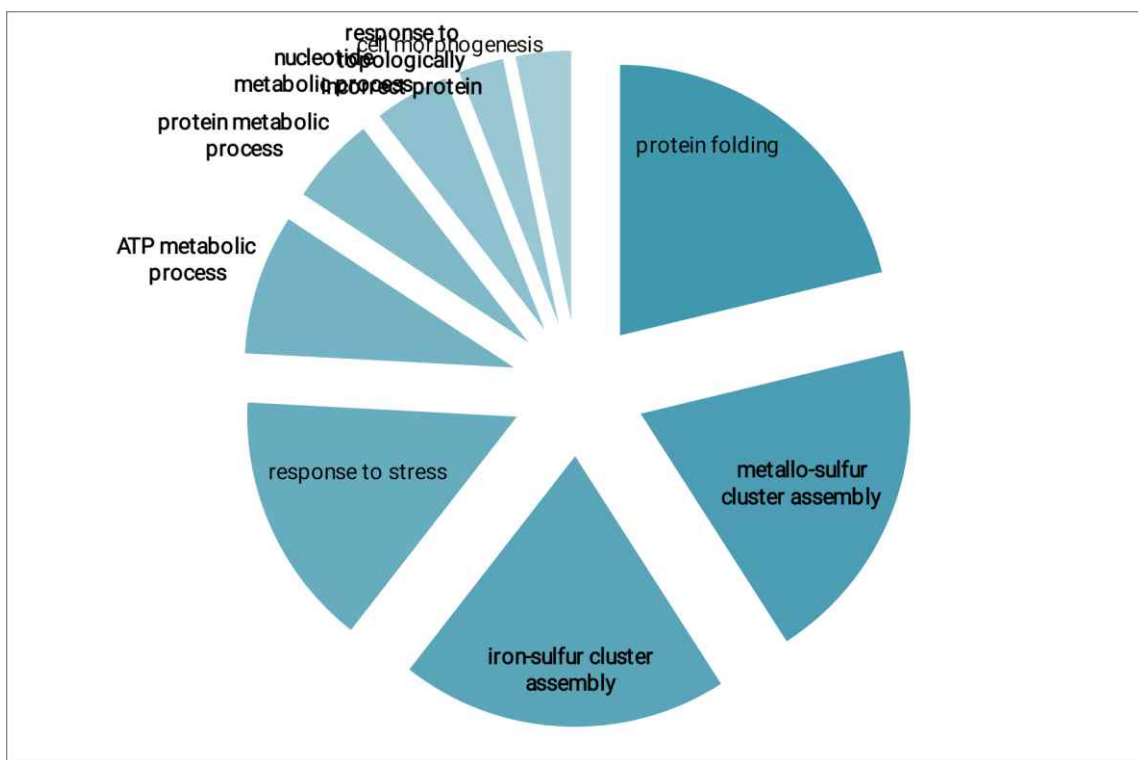
Hartl, F. U., Bracher, A. and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, 475(7356): 324-332.

ICRAF(2000). International Centre for Research in Agroforestry. Tree-Soil-Crop Interaction. In Asian Regional Research Programme. P. 2.

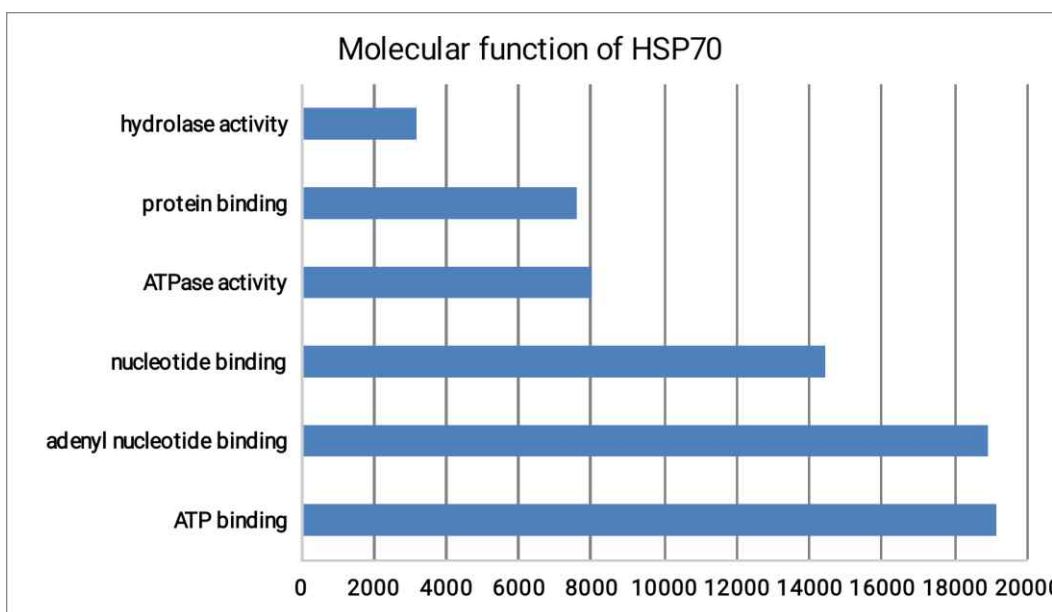
Huang, B. and Xu, C. (2008). Identification and

- characterization of proteins associated with plant tolerance to heat stress. *Journal of integrative plant biology*, 50(10): 1230-1237.
- Huang, X., Tao, P., Li, B., Wang, W., Yue, Z., Lei, J. and Zhong, X. (2015). Genome-wide identification, classification, and analysis of heat shock transcription factor family in Chinese cabbage (*Brassica rapa pekinensis*). *Genet. Mol. Res*, 14(1): 2189-2204.
- Kaim, W., Schwederski, B. and Klein, A. (2013). Bioinorganic Chemistry--*Inorganic Elements in the Chemistry of Life: An Introduction and Guide*: John Wiley & Sons.
- Kapoor, M. and Roy, S. (2015). 1 Heat-Shock Proteins and Molecular Chaperones: Role in Regulation of Cellular Proteostasis and Stress Management. *Abiotic Stresses in Crop Plants*, 1.
- Kernytsky, A. (2008). *Predicting protein function using sequence derived features selected by genetic algorithms*: Columbia University.
- Kityk, R., Kopp, J., Sinning, I. and Mayer, M. P. (2012). Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. *Molecular cell*, 48(6): 863-874.
- Kondrashov, A.F. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. Review. Proceedings of the Royal Society B 279 ; 5048 – 5057 doi:10.1098/rspb.2012.1108
- Li, H., Liu, S. S., Yi, C. Y., Wang, F., Zhou, J., Xia, X. J. and Yu, J. Q. (2014). Hydrogen peroxide mediates abscisic acid induced HSP70 accumulation and heat tolerance in grafted cucumber plants. *Plant, cell & environment*, 37(12), 2768-2780.
- Liao, Z., Ju, Y. and Zou, Q. (2016). Prediction of G protein-coupled receptors with SVM-prot features and random forest. *Scientifica* 1-10. Available on: <http://dx.doi.org/10.1155/2016/8309253>
- Lu, G., Westbrooks, J. M., Davidson, A. L. and Chen, J. (2005). ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(50): 17969-17974.
- Nollen, E. A. and Morimoto, R. I. (2002). Chaperoning signaling pathways: molecular chaperones as stress-sensing heat shock proteins. *Journal of cell science*, 115(14): 2809-2816.
- Orengo, C. A., Michie, A., Jones, S., Jones, D. T., Swindells, M. and Thornton, J. M. (1997). CATH—a hierarchic classification of protein domain structures. *Structure*, 5(8): 1093-1109.
- Ogbonna A. C. (2011). Current Developments in Malting and Brewing Trials with Sorghum in Nigeria: A Review. *Journal of the Institute of Brewing* 117(3): 394–400. Available on: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2050-0416.2011.tb00485.x>
- Pavli, O.I., Ghikas, D.V., Katsiotis, A. and Skaracis, G. (2011). Differential expression of heat shock protein genes in sorghum (*Sorghum bicolor* L.) genotypes under heat stress. *Australian Journal of Crop Science* (AJCS) 5(5):511-515.
- Proost, S., Van Bel, M., Sterck, L., Billiau, K., Van Parys, T., Van de Peer, Y. and Vandepoole, K. (2009). PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *The Plant Cell*, 21(12): 3718-3731.
- Remm, M., Storm, C. E. and Sonnhammer, E. L. (2001). Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *Journal of molecular biology*, 314(5): 1041-1052.
- Rentzsch, R. and Orengo, C. A. (2009). Protein function prediction—the power of multiplicity. *Trends in biotechnology*, 27(4): 210-219.
- Saier Jr, M. H., Tran, C. V. and Barabote, R. D. (2006). TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic acids research*, 34(suppl\_1), D181-D186.
- Schlecht, R., Scholz, S. R., Dahmen, H., Wegener, A., Sirrenberg, C., Musil, D. and Bukau, B. (2013). Functional analysis of Hsp70 inhibitors. *PloS one*, 8(11): e78443.
- Shi, L. (2011). Application of molecular modeling techniques to study the structure, dynamics, and interactions of membrane proteins. A Dissertation submitted to the Faculty of the Graduate School of the University of Minnesota. 285pp.
- Sørensen, J. G., Kristensen, T. N. and Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecology Letters*, 6(11): 1025-1037.
- Sun, W., Van Montagu, M. and Verbruggen, N. (2002). Small heat shock proteins and stress tolerance in plants. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1577(1): 1-9.
- Teixeira, E. I., Fischer, G., van Velthuisen, H., Walter, C. and Ewert, F. (2013). Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*, 170: 206-215.
- Timperio, A. M., Egidi, M. G. and Zolla, L. (2008). Proteomics applied on plant abiotic stresses: role of heat shock proteins (HSP). *Journal of proteomics*, 71(4): 391-411.
- Tusnady, G. E. and Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *Journal of molecular biology*, 283(2): 489-506.
- Tyedmers, J., Mogk, A. and Bukau, B. (2010). Cellular strategies for controlling protein aggregation. *Nature reviews Molecular cell biology*, 11(11): 777-788.
- Valliyodan, B. and Nguyen, H. T. (2006). Understanding regulatory networks and engineering for enhanced drought tolerance in plants. *Current opinion in plant biology*, 9(2): 189-195.
- Vaseva, I., Sabotič, J., Šuštar-Vozlič, J., Meglič, V., Kidrič, M., Demirevska, K. and Simova-Stoilova, L. (2012). The response of plants to drought stress: the role of dehydrins, chaperones, proteases and protease inhibitors in maintaining cellular protein function. *Droughts: New Research*, 1-45.

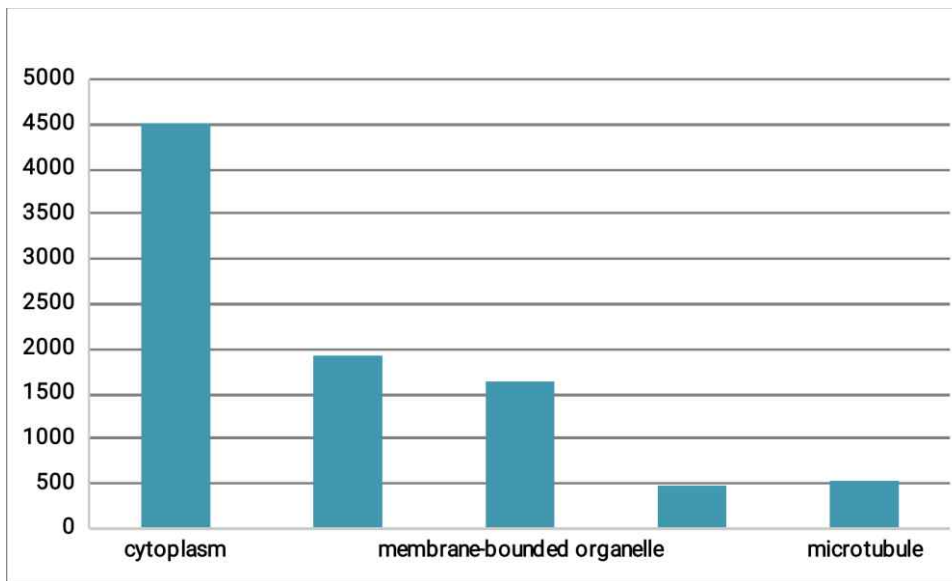
- Ventura, S. (2005). Sequence determinants of protein aggregation: tools to increase protein solubility. *Microbial Cell Factories*, 4(1):11.
- Wahid, A., Gelani, S., Ashraf, M. and Foolad, M. R. (2007). Heat tolerance in plants: an overview. *Environmental and experimental botany*, 61(3): 199-223.
- Wan'g, W., Vinocur, B., Shoseyov, O. and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in plant science*, 9(5): 244-252.
- Wisén, S., Bertelsen, E. B., Thompson, A. D., Patury, S., Ung, P., Chang, L. and Carlson, H. A. (2010). Binding of a small molecule at a protein–protein interface regulates the chaperone activity of hsp70–hsp40. *ACS chemical biology*, 5(6): 611-622.
- Zou, C., Lehti-Shiu, M.D., Thomashow, M. and Shiu, S. H. (2009) Evolution of Stress-Regulated Gene Expression in Duplicate Genes of *Arabidopsis*



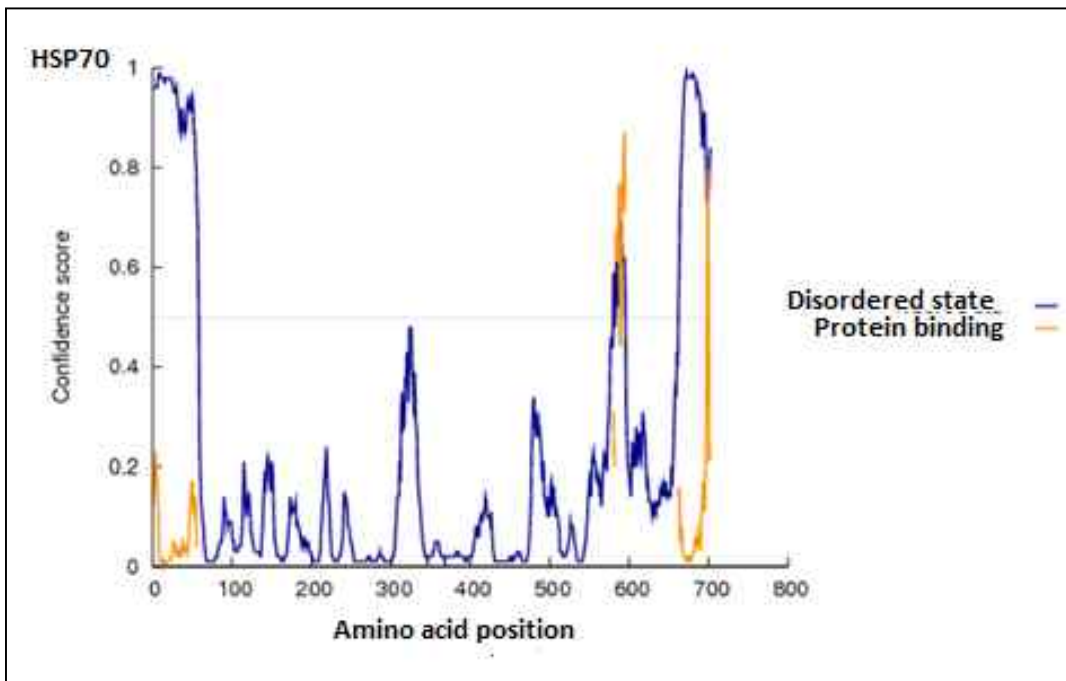
**Fig. 1: Prediction of biological process for HSP70 in *Sorghum bicolor* (Overlapping items can be listed under figure for more clarity)**



**ig. 2: Molecular function predicted of HSP70 in *S. bicolor***



**Fig. 3: Cellular localization prediction of HSP70 in *Sorghum bicolor***



**Fig. 4: Disorder / Order prediction for HSP70**

The image shows a plot of the probability of disorder (on the y-axis) for each numbered amino acid in the sequence (on the x-axis). The disorder/order probability threshold is shown as a dashed line on the plot. Residues above the threshold could be considered as mostly disordered and below as mostly ordered. Amino acids in the input sequence are considered disordered when the blue line is above the grey dashed line, that is the confidence score is higher than 0.5. The orange line shows the confidence of disordered protein binding residue predictions.

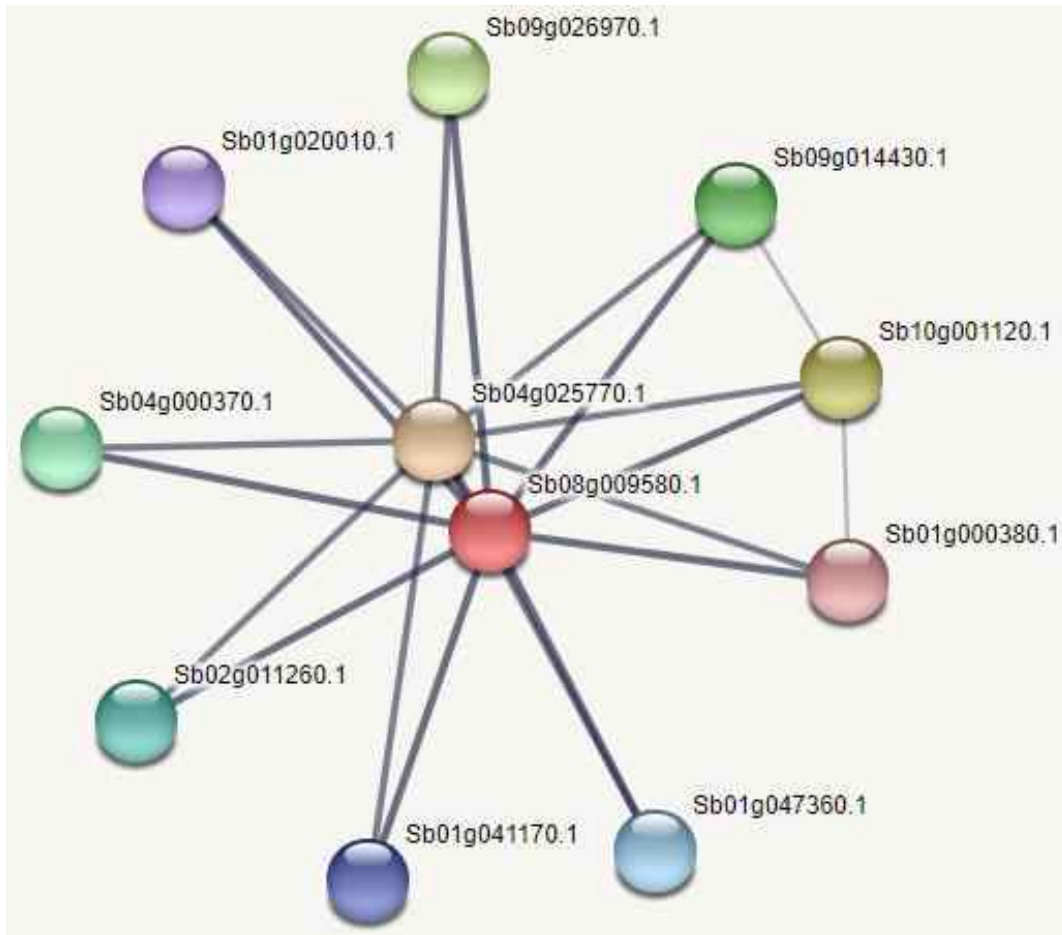


Fig. 5: Analysis of protein-interaction of *Sorghum bicolor* HSP70 with others protein

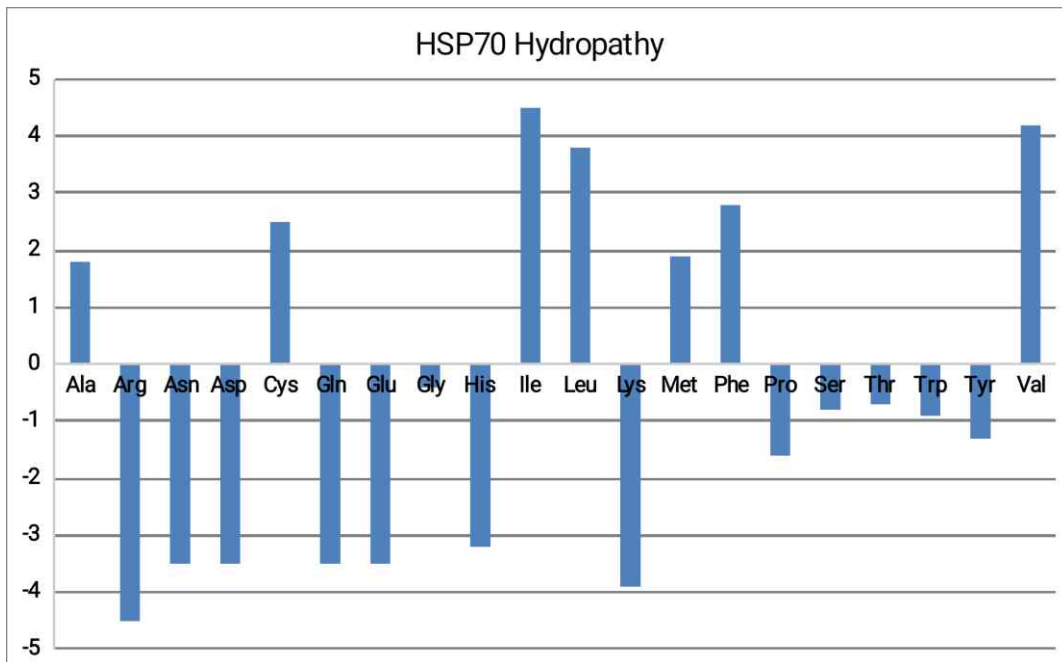
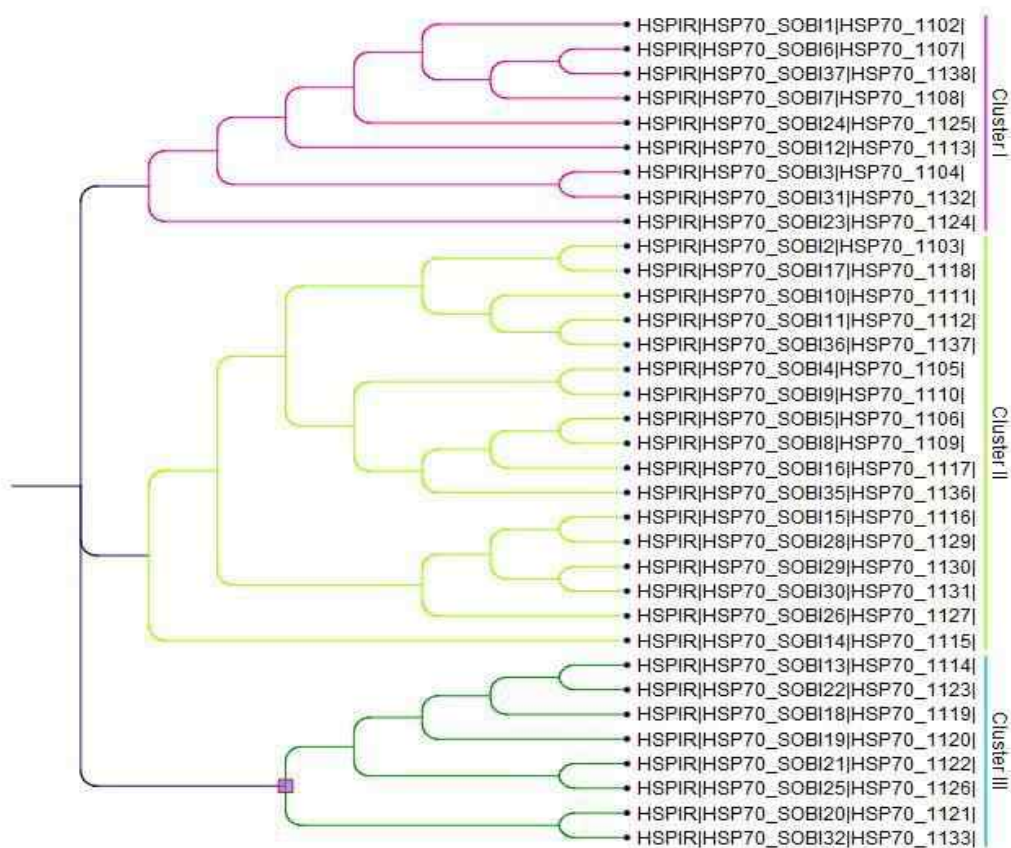


Fig. 6: Hydropathy analysis of amino acid sequences in *Sorghum bicolor* HSP70 explaining the up and down regulation of amino acid sequences





**Figure 7: Phylogenetic tree of *Sorghum bicolor* HSP70 proteins.**

The sequences were aligned by CLUSTALW at CLC and the unrooted phylogenetic tree was deduced by neighbor-joining method. The proteins were classified into three distinct clusters. Each family was assigned a different colour according to well-known members in other species.

**Table 1: Homology study of HSP70 genes in *Sorghum bicolor***

HSP70	Identity	Similarity	E-value	Query cover
<i>Geobacillus kaustophilus</i> HTA426	62%	100%	0.0	83%
<i>Mycobacterium tuberculosis</i> H37Rv	58%	100%	0.0	83%

**Table 2: *Sorghum bicolor* top 10 conserved motifs of HSP70 sequences found by the MEME system**

HSP70	Width	E-value	Best possible match
1	41	1.6e-713	AYFNBSQRQATKDAGTIAGLNVMRIINEPTAAAJAYGLDKK
2	50	3.3e-588	TACERAKRTLSSSTAQTTEIDSLHDGIDFSETITRARFEELNKDLFRKCM
3	41	7.9e-549	FTDTERLVGEAAKNQAAMNPTNTIFDVKRLIGRRFSDSESVQ
4	38	1.5e-426	PVEKCLRDAMDKSSVHDVVLVGGSTRIPKVQSLLODF
5	50	4.3e-546	MTVLIPRNTTIPTKKEQVFSTYSDNQITGVLIQVYEGERARTKDNNLLGKF
6	35	7.8e-444	GVFEVKATAGDTHLGGEDFDNRLVDHFVREFKRKH
7	41	1.8e-475	YKGEKQFAPEEISSMVLAKMKETAAYLGTTVKNAVITVP
8	39	1.5e-393	LSGIPPAPRGVVPQIEVTFDIDANGILNVSAEDKTTGRKN
9	29	8.4e-381	TTYSCVAVWRHDRVVIANDQGNRTTPSY
10	28	3.1e-340	GKELCKSINPDEAVAYGAAIQAAILSGE