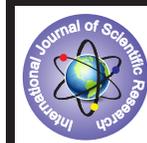


## Identification of EST derived microsatellite markers for heat stress tolerance in *Solanum melongena* by *in silico* methods



### Agriculture

**KEYWORDS :** *Solanum melongena*, EST sequences, heat stress tolerance, microsatellite markers, HSP90

**Reena Rosy Thomas**

Section of Economics & Statistics, Indian Institute of Horticultural Research, Hessarahatta, Bangalore 560089, Karnataka, INDIA

**M. Krishna Reddy**

Head, Division of Plant Pathology, Indian Institute of Horticultural Research, Hessarahatta, Bangalore 560089, Karnataka, INDIA

**Riaz Mahmood**

Professor and Chairman, Dept. of P.G Studies and Research in Biotechnology and Bioinformatics, Kuvempu University, Shivamogga, 577 451, Karnataka, INDIA

**M.K. Chandra Prakash**

Section of Economics & Statistics, Indian Institute of Horticultural Research, Hessarahatta, Bangalore 560089, Karnataka, INDIA

### ABSTRACT

*The extreme environmental conditions disturbs plant growth and development, restricting the genetic potential and yield losses in plants. It is necessary to understand plant responses to stresses that disturbs the homeostatic equilibrium. Newer genes and markers need to be identified to make heat stress tolerant crops and comparative studies in ancestral plant groups can help to understand the complex biological processes. The plants cope with adverse temperature conditions by regulating various mechanism and sustains high temperature stress by producing heat shock proteins (HSPs) and is involved in the innate immunity and abiotic stress tolerance in higher plants. In this context, in silico methods were used to discover new EST derived microsatellite markers associated with heat stress tolerance in Solanum melongena for plant breeding programme as they are directly linked to functional genes. The publically available EST collection of S. melongena were explored to identify EST linked markers associated with heat tolerance. Approximately 529 EST sequences having repeats were analyzed for stress tolerance using insilico methods. These were categorized into 11 different protein super-families implicated in stress tolerance based on conserved domain and protein functions. Among these, five unique EST sequences pertains to HSP90 superfamily associated with heat stress tolerance were found having a microsatellite marker. These are potential markers useful for marker assisted selection in breeding programs and also facilitate the targeted research for development of heat stress tolerant crops.*

### INTRODUCTION

*Solanum melongena* L. is one of the most common vegetable crops grown throughout the year in India which is adapted to a wide range of climatic conditions. However, stresses such as drought, salinity, osmotic and heat affects plant growth and development (Cramer et al., 2010) which has a negative impact on crop productivity (Bray et al., 2000). Plants growing under these stress conditions, responds by modifying several aspects in their metabolic cascade (Dos Reis et al., 2012) for their survival. These response mechanisms help plants to survive during the stress period as well as to recover following cessation of the stress. One major response of plants upon encountering abiotic stresses is the activation of stress responsive genes. Microsatellites are involved in gene regulation, and organism development and evolution (Labbe et al., 2011). Excess numbers of microsatellite repeats play important roles in genomic functions. Since genome sequences of many plant species are still lacking which includes *Solanum melongena*, protein databases available for the model plants are used as reference databases for analyzing data obtained from other plant species. As more sequenced genomes become available, computational methods for predicting protein function from sequence data continue to be of high importance. Bioinformatics is a method of genomic analysis utilizing a database of stored information that is then used for comparison. We can even compare known DNA sequences and coded protein sequences to our unknown gene and protein. Protein domains are often conserved across many species and the characterization of such domains has led to the classification of genes into families, helping to assign function to other genes. HSPs are product of heat shock genes and are named according to their molecular weight. It was reported that they are evolutionarily conserved in Solanaceae crops (Chandra Prakash et al., 2013). Hsp90 (heat shock protein 90) is one of the most conserved and abundant proteins, and is an essential component of the protective stress response. The basic function of HSP 90 is to maintain its targets in an appropriate conformation until they are stabilized by interacting with other components of the pathway (Lewin 2007). As the microsatellite markers are involved in gene regulation, lo-

ating these markers using conserved protein families approach will be useful in identifying genes associated with them and their specific function. The EST sequences of *Solanum melongena* having microsatellites were categorized on the basis of their similar protein superfamily function using *in silico* methods.

### MATERIAL AND METHODS

The EST data available in the public domain for brinjal (*Solanum melongena*) were downloaded in fasta format from <http://www.ncbi.nlm.nih.gov/nucest> for identifying markers implicated in stress tolerance. These EST collections were analysed for repeat sequences comprising of di, tri, tetra repeats using repeat masker, Tandem repeat finder software. The EST sequences containing low complexity repeat sequences with an average length of around 500 nucleotides were processed as batch files against NCBI nucleotide database. From the BLAST results, the alignments of similar sequences were computed on the basis of expectation (E) values and hit value above 500 with gene entry. The protein IDs of the nucleotide sequences from BLAST results were retrieved using the E-utility server program which provides an interface to the Entrez query and database system. A custom URL syntax has been developed embedded with nucleotide IDs to retrieve the corresponding protein Gen bank IDs using the E-utility program. NCBI CD search program searches against the protein database and generates output as a text file having list of conserved domain hits and their superfamilies description against each protein queries.

These superfamilies along with protein domains are useful in analyzing the conservedness, homology and the respective superfamily it belongs to. The protein Domains associated with abiotic stress involving Heat stress, also Osmotic, Saline and other environmental stress tolerant were extracted to spread sheet file from public domain as source file which is compared with NCBI CD search output results comprising of protein domains. The matching ESTs of *Solanum melongena* with protein domains and their respective superfamilies are extracted and clustered into several subgroups associated with abiotic stress tolerance.

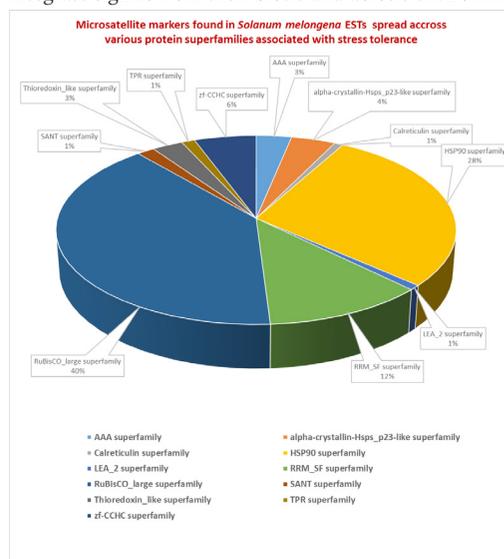
These are potential markers useful for marker assisted selection in breeding programs and would also expedite the targeted research for heat stress tolerant crops development.

**RESULTS AND DISCUSSIONS**

The EST sequences of *Solanum melongena* having microsatellite markers associated with stress tolerance has been identified based on conserved protein domains and their respective superfamilies. Among the 529 EST sequences having repeats were analyzed for stress tolerance using *in silico* methods and it was found that 43 sequences having markers linked to stress tolerance, and belongs to 11 different superfamilies. Further, it was classified based on protein functions implicated in stress tolerance and clustered into major categories like RuBisCO, HSP 90, RRM, alpha crystallin Hsps p23, AAA, TPR superfamily etc. From this classification, five unique EST sequences that belongs to heat shock protein 90 family were found having a microsatellite marker. Distribution of *Solanum melongena* marker sequences associated with several protein superfamilies is given as pie chart (Fig-1). These EST sequences are having more than 90% similarity with Solanaceae crops. It was found that five unique EST sequences be belongs to HSP 90 superfamily. The major proteins present with this superfamily were HSP 90, heat shock cognate protein 80 in various crops including Solanaceae crops.

**Heat Shock Protein 90:** Hsp90 is a chaperone protein that assists other proteins to fold properly (Lee et al., 2009), stabilizes proteins against heat stress, and aids in protein degradation as well as facilitating cell signaling. In this study, alignment of EST sequences of HSP90 revealed that these sequences is highly homologous to published HSP90 sequences of Solanaceae crops. HSP is one of the most common of the heat-related proteins and identifying markers linked to heat tolerance would enhance the efficiency of breeding programs in developing heat tolerant cultivars. Heat stress induces the expression of several heat shock proteins (Kollipara et al., 2002; Renaut et al., 2004). It was re-

ported that the HSP 90 genes were induced by heat, salt and osmotic stresses but the response times and expression abundances were diverse (Xu et al., 2013). HSP90 machinery is important for stomatal closure and serves essential functions in plants to integrate signals from their biotic and abiotic environments.



**Fig-1: Distribution of *Solanum melongena* EST sequences associated with several protein superfamilies**

**EST linked marker sequences:**

The 5 unique EST ids of *S. melongena* having markers that belongs to HSP 90 superfamily associated with heat stress tolerance are listed in table-2 and the corresponding match with several Solanaceae crops along with matching bit score range is given in table-3.

**Table-2: EST Ids and markers present in respective ESTs of *S. melongena***

S.No	EST ID	EST length	Marker sequences
1	FS045050	611	CTTCTTTCTCTTTTTCCTCATCAACCTCTCTACCTTCCCTTCTTCGTCCTTCTTCTCCTCTTCATCC TCATCATC
2	FS039747	581	CCTCCTTACCTTCCCCTCTCATCTTTTCTTCTCCTCTTCTCCTCTC
3	FS043436	453	AGGAAAAGGAGAAGGAAGAGAAGAAAAGAAG
4	FS013618	490	GAGGAGGAAGAAGAGAAGAAAGATGAGGAGGGGAAGGTAGAAGAGTTGATGAGGAAAAGGAGAA GGAAGAGAAGAAAAGAAGAAG
5	FS027845	366	AGAGGAGTTGATGAGGAAAAGAGAAGAAGAGAAGAAAAGAAGAAA

**Table - 3: The *S. melongena* EST IDs that are matching against HSP sequences of several crops along with their bit score range**

EST No.	Crop name	Protein name	BIT score range
FS013618	<i>Solanum tuberosum, Solanum lycopersicum, Capsicum chinense, Nicotiana benthamiana, Nicotiana tabacum, Glycine max</i>	Hsp90-2 heat shock cognate protein 80 Unchar protein LOC100819568	503-767
FS027845	<i>Solanum tuberosum</i>	heat shock cognate protein 80	599
FS039747	<i>Solanum lycopersicum, Solanum tuberosum, Capsicum chinense, Nicotiana tabacum, Nicotiana benthamiana, Glycine max, Theobroma cacao, Capsella rubella, Citrus sinensis, Eutrema salsugineum, Citrus clementina</i>	Heat shock cognate protein 80 Hsp90-2 Heat shock protein 81.4 hypothetical protein	503-869
FS043436	<i>Solanum tuberosum, Solanum lycopersicum, Capsicum chinense, Nicotiana tabacum, Nicotiana benthamiana, Populus trichocarpa</i>	Hsp90-2 heat shock cognate protein 80 Hypothetical protein	503-732
FS045050	<i>Nicotiana tabacum, Capsicum chinense, Solanum tuberosum, Solanum lycopersicum, Theobroma cacao, Ricinus communis, Eutrema salsugineum, Arabidopsis thaliana, Arabidopsis lyrata subsp. lyrata, Populus trichocarpa, Citrus sinensis, Medicago truncatula, Glycine max, Citrus clementina, Camellia sinensis, Phaseolus vulgaris</i>	heat shock protein 90 Heat shock protein 81.4 Hypothetical protein Early-responsive to dehydration	507-987

**Conclusion:**

From the above results, it is observed that 43 unique EST sequences having micro satellite markers was identified in *Solanum melongena* which is spanned across 11 protein super-families. These EST sequences are highly conserved in several Solanaceae crops and are proved to be tolerant to abiotic stress in various organisms. Among these, 5 unique EST sequences having conserved domains of HSP90 associated with heat stress tolerance has a marker in respective EST sequences, also these sequences are highly similar to HSP 90 genes. The HSP 90 genes are strongly induced by heat stress and also responsive to salt and osmotic stresses.

These 5 EST sequences are showing more than 95% similarity with 71 published HSP sequences across several crops specifically against Solanaceae crops. These EST linked markers of HSP90 serve as potential genetic markers useful in marker assisted selection for identifying the heat stress tolerant parental lines in plant breeding programs. Information on the relative abundance of these micro satellite markers in the EST sequences of *Solanum melongena* may also provide clues to the functionality of micro satellites in gene regulation and to develop crops that are tolerant to heat stress in future breeding programs.

**REFERENCE**

- Bray, E.A., Bailey-Serres, J., Weretilnyk, E.(2000). Responses to abiotic stresses. In Gruissem, W., Buchanan, B., Jones, R. eds, *Biochemistry and Molecular Biology of Plants*. American Society of Plant Physiologists, Rockville, MD, pp 1158–1249. | • Chandra Prakash, M.K., Thomas, R.R., Reddy, M.K., Mohandas, S. (2013). Molecular evolutionary conservedness of small heat shock protein sequences in Solanaceae crops using in silico methods. *Journal of Horticultural Sciences*, 8(1), 82-87. | • Cramer, G.R. (2010). Abiotic stress & plant responses from the whole vine to the genes. *Aust. J. Grape Wine Res.*, 16, 86 - 93. | • Dos Reis, S.P., Lima, A.M., de Souza, C.R. (2012).Recent molecular advances on downstream plant responses to abiotic stress. *Int. J. Mol. Sci.*, 13(7), 8628-8647. | • Kollipara, K.P., Saab, I.N., Wych, R.D., Lauer, M.J., et al. (2002). Expression profiling of reciprocal maize hybrids divergent for cold germination and desiccation tolerance. *Plant Physiol.*, 129, 974-992. • Labbe, J., Murat, C., Morin, E., Le Tacon, F., Martin, F. (2011). Survey and analysis of simple sequence repeats in the *Laccaria bicolor* genome, with development of microsatellite markers. *Current Genetics*, 57(2), 75 - 88. | • Lee,D.G, Ahsan, N., Lee, S.H., Lee, J.J., Bahk, J.D., Kang, K.Y., et al. (2009).Chilling stress-induced proteomic changes in rice roots. *J. Plant Physiol.*, 166, 1–11. | • Lewin, B. (2007). *Genes IX*, Jones & Bartlett publishers, pp. 226. | • Renaut, J., Lutts, S., Hoffmann, L., Hausman, J.F. (2004).Response of poplar to chilling temperatures: proteomic and physiological aspects. *Plant Biol.*, 6, 81-90. | • Xu, J., Xue, C., Xue, D., Zhao, J., Gai, J., et al. (2013). Overexpression of GmHsp90s, a Heat Shock Protein 90 (Hsp90) Gene Family Cloning from Soybean, Decrease Damage of Abiotic Stresses in *Arabidopsis thaliana*. *PLoS ONE*, 8(7), e69810. |