FUNCTIONAL GENOMICS, TRANSCRIPTOMICS AND METABOLOMICS OF HEAT STRESS IN BARLEY (*Hordeum vulgare* L.)

Ph.D. THESIS

by

REEKU CHAUDHARY

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247667 (INDIA)
SEPTEMBER, 2019

FUNCTIONAL GENOMICS, TRANSCRIPTOMICS AND METABOLOMICS OF HEAT STRESS IN BARLEY (*Hordeum vulgare* L.)

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

in

BIOTECHNOLOGY

by

REEKU CHAUDHARY

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE–247667 (INDIA)
SEPTEMBER, 2019

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in this thesis entitled, **"FUNCTIONAL GENOMICS, TRANSCRIPTOMICS AND METABOLOMICS OF HEAT STRESS IN BARLEY (*Hordeum vulgare* L.)"** in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from July, 2015 to September, 2019 under the supervision of Dr. Harsh Chauhan, Assistant Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institution.

(REEKU CHAUDHARY)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Harsh Chauhan)
Supervisor

The Ph.D. Viva-Voce Examination of **Ms. Reeku Chaudhary,** Research Scholar, has been held on .....

Chairman, SRC **Signature of External Examiner**

This is to certify that the student has made all the corrections in the thesis.

Signature of Supervisor **Head of the Department**

Dated:.....

ABSTRACT

Adverse impacts of global climate change including extreme high temperature on cereal crop production have been evidenced at the global level. In plants, the elevated temperature above the optimal range dramatically impairs the process of both vegetative and reproductive phases of the development, mainly due to disruption in their vital metabolic processes. However, plants possess an inherent rescue machinery to combat such stressful conditions. Heat shock proteins (HSPs) and heat shock factors (HSFs) are the crucial components of these stress associated rescue mechanisms and also required for normal biological processes.

In the present study, a genome wide identification, evolutionary relationship and comprehensive expression analysis of *Hsp70*, *Hsp90* and *Hsp100* gene families have been performed in barley. Barley genome was revealed to possess 13 members of *Hsp70* gene family, along with 4 members of *Hsp110* subfamily, 6 of *Hsp90*, and 8 members of *Hsp100* gene family, and *Hsp* genes were found to be distributed across all 7 chromosomes of barley. The encoded protein members of these genes were predicted to be localized to cell organelles such as cytosol, mitochondria, chloroplast, and ER. Despite a larger genome size, lesser members of these *Hsp* genes in barley were observed owing to less duplication events. The variable expression patterns obtained for genes encoding the proteins localized in same subcellular compartment suggest their possible diverse roles and involvement in different cellular responses. Expression profiling of these genes was performed by qRT-PCR in an array of 32 tissues, revealed differential and tissue specific expression of various members of *Hsp* gene families. We found the upregulation of *HvHspc70-4, HvHsp70Mt70-2, HvHspc70-*5a, *HvHspc70-5b*, *HvHspc70-N1, HvHspc70-N2, HvHsp110-3*, *HvHsp90-1, HvHsp100-1*, and *HvHsp100-2* upon exposure to heat stress during reproductive development. Furthermore, their higher expression during heat stress, heavy metal stress, drought, and salinity stress has also been observed.

HSFs of class A are the regulators of several heat shock and other abiotic stress responsive genes. In the present investigation, the constitutive overexpression of class A heat shock factor *HvHSFA2c* was achieved in barley. Several overexpressing transgenic lines were identified and analyzed for transgene expression by performing histochemical gus assay, genomic PCR, RT and q-RT PCR. The transcriptomics studies of overexpressing transgenic lines through RNA-Seq revealed the potential putative target genes of *HvHSFA2c*. Here, we inferred that *HvHSFA2c* may be acting by inducing the transcription of chaperone encoding genes belonging to all five classes of Hsps regulating the protein homeostasis of plants. Apart from chaperones; *HvHSFA2c*

may also be regulating the expression of genes associated with calcium-mediated signal transduction pathways, which have been known to be associated with the heat shock response of plants. Also the upregulation of genes involved in biosynthesis of jasmonic acid and its signaling, projected the involvement of JA in heat stress tolerance associated mechanisms. Similarly, the differential expression of early and late auxin responsive genes and gibberellin homeostasis associated genes suggested that *HvHSFA2c* may be regulating the activity of these phytohormone associated growth mechanisms of plants.

Both photosynthetic and respiratory processes are vital for the plants and adversely affected by HS. The overexpression of genes related to light and dark phases of photosynthesis and mitochondrial respiratory electron transport chains suggested that *HvHSFA2c* may be providing thermotolerance by stabilizing the activity of these pathways. The differential expression (up and downregulation) of several other stress associated transcription factors such as WRKY, Myb, AP2/ERF etc. suggested the convergence of stress associated signaling and transcriptional regulation.

We further inferred that *HvHSFA2c* may also be potentially targeting the ROS scavenging pathways by activating flavonoid biosynthesis, and antioxidative enzymes such as superoxide dismutase catalase and ascorbate peroxidase. The higher activity of these enzymes was further validated in our overexpressing transgenic lines. Here, we concluded that, the genes associated with these described pathways may potentially be observed under the direct or indirect regulation of *HvHSFA2c*. Furthermore, the comparative metabolic profiling of wild type plants of barley subjected to HS and plants grown under controlled conditions was performed to analyze the metabolic adaptation of these plants during heat stress. The changes were revealed in metabolomics fluxes of non-reducing sugars, several amino acids, and accumulation of secondary metabolites which may be protecting plants during heat stress conditions.

This study provided the insights of heat stress tolerance mechanisms in barley plants through comparative genomic studies of high molecular weight heat shock proteins, and functional study of a heat shock factor *HvHSFA2c*. The metabolomics study suggested the alterations in key metabolic processes such as citric acid cycle, carbohydrate metabolism and secondary metabolism.

ACKNOWLEDGEMENTS

It is by lavish love and blessing of the supreme God that I have been able to complete my studies successfully hitherto and present of work for which I am eternally indebted.

First of all, I am feeling great pleasure to evince my profound sense of veneration and gratitude to my supervisor, Dr. Harsh Chauhan, Assistant Professor, Department of Biotechnology, I.I.T. Roorkee, for his keen interest, ever enlivening, constructive criticism and encouragement throughout the pursuit of the present research and vital suggestion during the preparation of this manuscripts, this work owes a lot to him. I am also very thankfull to Dr. Debabrata Sircar Associate Professor, Department of Biotechnology, I.I.T. Roorkee for his unconditional help and valuable suggestions during my Ph.D. journey.

I humbly express my profound gratitude and heartfelt thanks to Dr. A.K. Sharma Professor and Head, Department of Biotechnology. I extend my thanks and sense of indebtness to the members of my research committee, Dr. Pravindra Kumar, Professor, Department of Biotechnology, Dr. Partha Roy Professor, Department of Biotechnology, Dr. Debabrata Sircar, Associate Professor, Department of Biotechnology, Dr. Kaushik Ghosh, Professor, Department of Chemistry, for their valuable suggestion and liberal help rendered during the course of study and research work.

I am also very thankful to Prof. Randhawa, and all the teaching and nonteaching staff of Department of Biotechnology, Indian Institute of Technology Roorkee, for their valuable help through my course-cum research programme.

I would also like to thank my lab mates and friends Sumit, Anuj, Parul, Sweta, Sampath, Bidya, Shifa, Amita, and Vandana for their co-operation and constant support. It is indeed a great pleasure to acknowledge my dear friends, Ruchi, Pooja, Sweta and Archana for their affection, indispensable help and providing friendly, homely environment as well as unmatchable support, which made my study time in the institute cheerful. With profound regards in a more personal sense, I owe deepest debts to my parents, who taught me the value of wisdom, based on erudition but without enslaved by it. Their persistent inspiration, selfless sacrifice, continuous encouragement and blessings gave untiring help and have enabled me to be so today. I would also like to give a heartfelt gratitude to my husband Dr. Saurabh Badoni, for his love, support, continual encouragement and affectionate co-operation, without that this work would have been impossible. UGC is duly acknowledged for financial support.

REEKU

List of Publications

- Chaudhary R, Baranwal VK, Kumar R, Sircar D, Chauhan H (2019). Genome-wide identification and expression analysis of Hsp70, Hsp90, and Hsp100 heat shock protein genes in barley under stress conditions and reproductive development. Functional & integrative genomics. 27:1-6.
- Gupta BB, Selter LL, Baranwal VK, Arora D, Mishra SK, Sirohi P, Poonia AK, Chaudhary R, Kumar R, Krattinger SG, Chauhan H (2019). Updated inventory, evolutionary and expression analyses of G (PDR) type ABC transporter genes of rice. Plant Physiology and Biochemistry. 142:429-329.
- Reeku Chaudhary and Harsh Chauhan. Heat shock protein gene families in barley: genome wide identification, organization and expression profiling in response to various abiotic stresses and reproductive development. International Plant Physiology Congress, CSIR- National Botanical Research institute, Lucknow, India, 2-5 December, 2018, Page-213.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abbreviations

%	:	Percentage
˚C	:	Degree celsius
µg	:	microgram
µl	:	microliter
2-4 D	:	2,4-Dichlorophenoxyacetic acid
6-BAP	:	6-Benzylaminopurine,
AHA	:	Aromatic acidic hydrophobic
AT	:	Acquired thermotolerance
ATP	:	Adenosine triphosphate
BLAST	:	Basic local alignment search tool
bp	:	Base pair
BSA	:	Bovine serum albumin
cDNA	:	Complementary DNA
CS	:	Co-supression
CTAB	:	Cetyltrimethyl ammonium bromide
CTD	:	C-terminal domain
DBD	:	DNA binding domain
DEG	:	Differentially expressed genes
DEPC	:	Diethyl pyrocarbonate
DMSO	:	dimethyl sulfoxide
DNA	:	Deoxyribonucleic acid
dNTP	:	Deoxynucleotides
DTT	:	Dithiothreitol
EDTA	:	Ethylenediaminetetraacetic acid
EGTA	:	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
EtBr	:	Ethidium bromide

F	:	Forward
FPKM	:	Fragments per kilobase of transcript per million reads
g	:	Gram
Gb	:	Giga base
GO	:	Gene ontology
HS	:	Heat stress
HSE	:	Heat shock element
HSF	:	Heat shock factors
HSP	:	Heat shock proteins
Hz	:	Hertz
kDa	:	Kilo dalton
KEGG	:	Kyoto encyclopedia of genes and genomes
M	:	Molar
MAPK	:	Mitogen-activated protein kinase
MCT	:	Micro centrifuge tube
MES	:	2-(N-morpholino)ethanesulfonic acid
mg	:	Milligram
$MgCl_2$	:	Magnesium chloride
ml	:	Milliliters
mM	:	Millimolar
MOPS	:	(3-(N-morpholino)propanesulfonic acid)
mQTL	:	Metabolomics quantitative trait loci
N	:	Normal
Na_2EDTA	:	Disodium ethylenediaminetetraacetic acid
NAA	:	1-Naphthaleneacetic acid
NaOH	:	Sodium hydroxide
NBD	:	Nucleotide binding domain

NBT : Nitro blue tetrazolium

NES : Nuclear export signal

no. : Number

OD : Oligomerization domain

OE : Over expression

ORF : Open reading frame

PCR : Polymerase chain reaction

psi : Pounds per square inch

PSII : Photosystem II

PVP : Polyvinylpyrrolidone

q-RT PCR : Quantitative real time polymerase chain reaction

QTL : Quantitative trait loci

R : Reverse

RIN : RNA integrity number

RNA : Ribonucleic acid

RNaseA : Ribonuclease A

RNA-seq : High throughput RNA sequencing

ROS : Reactive oxygen species

rpm : Revolutions per minute

SD : Standard deviation

SDS-PAGE: Sodium dodecyl sulphate-Polyacrylamide gel electrophoresis

SOD : Superoxide dismutase

SSH : Suppression Subtractive Hybridization

TBE : Tris borate EDTA

TE : Tris EDTA

TF : Transcription factor

Tm : Melting temperature

UTR	:	Untranslated region
WT	:	Wild type
X-gluc	:	5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid
β-ME	:	Beta mercaptoethanol

Introduction

Introduction

1.1 Barley a model crop of Triticeae tribe

Barley (*Hordeum vulgare* L. 2n=2x=14) is the earliest domesticated cereal crop and an important member of *Poaceae* family. It comes after maize, rice, and wheat in terms of global cereal production. Barley grains are used for animal feed, human food and have wide applications in malt and brewing industry. Furthermore, they are the excellent source of β-glucan, tocotrienols and tocopherols (tocols). β-glucan is a soluble dietary fiber and required for maintaining gut health and lowering bad cholesterol, whereas tocols have antioxidant properties (Qureshi et al. 1991, Cavallero et al. 2002). Thus, barley is also considered as a functional food with several medicinal benefits (Quershi et al. 1991). Barley is a diploid, self-fertilized crop with a short life cycle and less complex genome as compared to its closet relative wheat, which make it ideal for plant genetic research. Additionally, the whole genome of barley has been sequenced with a high-quality map-based reference genome assembly (Mascher et al. 2017, Wicker et al. 2009, Mayer et al. 2009). These genomic resources are valuable for functional genomic studies and transgenic research for crop improvement (Keller and Krattinger, 2017, Bettgenhaeuser J and Krattinger SG, 2019). Furthermore, the genetic manipulation of barley is comparatively easier than its closest relative wheat (Goedeke et al. 2007, Hensel et al. 2007, Hensel et al. 2008). Hence, availability of genetic resources, less genomic complexity, and well established genetic manipulations protocols make barley a model crop for *Triticeae* family.

1.2 Heat stress and its impact on cereal production

Nowadays, global climate change has become a major area of concern. Due to climate change, rise in temperature is affecting the growth and development of crop plants. With an increase in global warming, El Niño phenomenon (the warm phase of El Niño Southern Oscillations) was recorded highest in the year 2015-2016 for various countries which fall in tropical and subtropical regions (Fig. 1.1) (Hu and Fedorov, 2017). Consequently, most parts of Asia-Pacific region observed hot spring and summer seasons. Therefore, crop failures due to extreme climates cause food insecurity by a decrease in crop production, land degradation which eventually lead to increase in food price (Holman et al. 2018). Hence, the alarming rise in earth temperature is majorly threatening food security worldwide (Bita and Gerats, 2013).

a.)

b.)

Fig. 1.1 Trends of global rise in temperature and agricultural productivity. a.) Rise in global mean surface temperature (Allen et al. 2019) b.) The global agricultural productivity index of the year 2018, revealing slow and non-sustainable agricultural productivity to meet food demands for increasing population till 2050. (https://globalagriculturalproductivity.org/data-resources/).

Almost, sixty percent of the total world population depends on agriculture (FAO, 2018) and extreme temperature has a dramatic effect on agriculture production. The increase in temperature by 2-3°C leads to 15-35% reduction in crop yields (Oritz et al. 2008). At global level, 3-6% reduction in yield was observed for each degree rise in temperature for four major crops comprising wheat, rice, maize and soybean (Zhao et al. 2017). As elevated temperature above the optimal range impaired the process of proper growth and development of plants, thus ultimately have detrimental effects on crop production (Ahuja et al. 2010). High temperature induced increasing pace of phenological development includes the accelerated progress of both wheat and barley plants through early reproductive phases (Rawson and Richards, 1993). When plants encounter high temperature in mid-anthesis or post-fertilization period, produce less number of seeds per spike with lower grain weight (Ferris et al. 1998) which ultimately reduces the crop production. Furthermore, in case of barley, metabolic shifts in the processes associated with the synthesis of storage compounds such as carbohydrate or starch production also has been found responsible for deprived grain filling and grain weight (Mangelsen et al. 2011, and Savin et al. 1996). Wheat and Barley both are winter (Rabi season) crops, sown and harvested between October and April in India. Kalra et al. (2008) reported the severe decline in production of Rabi crops due to high temperature stress (Fig. 1.2)

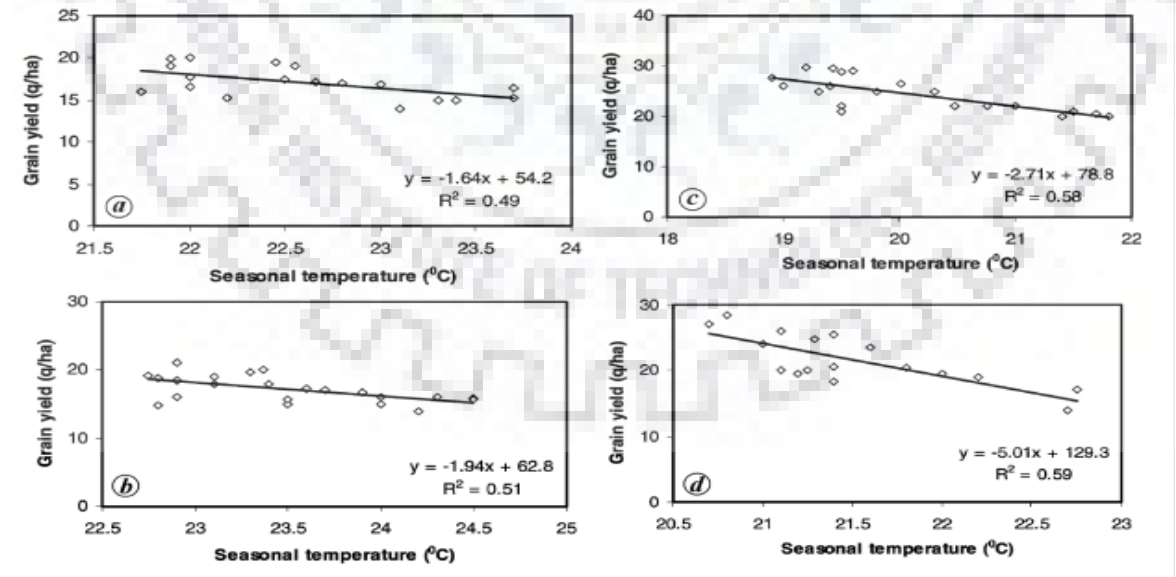

Fig. 1.2 Effect of heat stress on yield of barley crop in four states of India. (a) Uttar Pradesh, (b) Rajasthan, (c) Punjab, and (d) Haryana.

Thus, plants are vulnerable to high temperature stress during both vegetative and reproductive phases of development, due to disruption in the metabolic processes and gametophytic development (Hedhly, 2011, and Barnabas et al. 2008). However, plants possess a remarkable ability to survive in such adverse conditions. Hence, it is necessary to understand the molecular mechanisms evolved by plants to cope up thermal stress and their responses to re-establish cellular homeostasis.

1.3 Heat shock signaling and response in plants

Plants evoke an interconnected signaling cascade in response to heat stress and initiate heat shock response to survive in unfavorable conditions. The sensors of heat stress are, an alteration in the fluidity of the plasma membrane (Saidi et al. 2009), disruption in protein homeostasis at cytosolic and subcellular level, and some nuclear proteins (ARP6) (Kumar and Wigge 2010). These signals are finally perceived by heat shock factors which play as key regulators in response to heat and other abiotic stresses, by triggering the alterations in transcriptional reprogramming of a set of heat shock responsive genes (Scharf et al. 2012). Notably, in response to heat stress, the role of class-A HSFs is widely studied in plants. In Arabidopsis, class-A HSFs are divided into 9 groups (A1-A9) (Nover et al. 2001). The role of A1 HSFs as a regulator of heat shock responsive genes including the genes encoding other transcription factors e.g. HSFs, DREB2A, DREB2B, bZIP28, etc. is well studied in both Arabidopsis and tomato (Liu et al. 2011, Yoshida et al. 2011 and Scharf et al. 2012).

Moreover, HSFA2 has been found as a highly induced transcription factor during heat stress in tomato, Arabidopsis, rice and wheat (Busch et al. 2005, Wunderlich et al. 2007, Chauhan et al. 2011, Xue et al. 2014). Interestingly, HSFA2 become dominant after induction and its transcript, remain accumulated during prolonged high temperature and recovery phases. HSFA2 forms hetero-oligomer complex with HSFA1 in nucleus (Chan-Schaminet et al. 2009), to synergistically regulate the expression of genes involved in establishing protein, metabolic and redox homeostasis (e.g. *small HSPs*, *HSP70* and *HSP101* isoforms *ascorbate peroxidase 2* (*Apx2*), *galactinol synthase* (*GolS*)) (Nishizawa et al. 2008, Schramm et al. 2006, Mishra et al. 2002). Furthermore, the knockout plants of Arabidopsis for *HSFA2* gene showed significantly reduced tolerance towards thermal, high light and oxidative stresses (Schramm et al. 2006, and

Nishizawa-Yokoi et al. 2009). In contrast, enhanced tolerance to thermal and other environmental stresses (salt, osmotic, anoxia, and oxidative stress) has been reported in *Arabidopsis* showing overexpression for *HSFA2* gene (Zhang et al. 2009, Ogawa et al. 2007 and Nishizawa et al. 2006). The elevated transcript levels of *HSFA2* gene in rice and wheat during heat stress, suggested its major regulatory role in combating heat shock response (Chauhan et al. 2011 and Xue et al. 2014).

In plants, the HSF family is highly diversified with multiple and functionally non-redundant members (Miller and Mittler 2006). It has been shown that the ectopic overexpression of *HSFA3* of tomato and wheat in Arabidopsis increased the thermotolerance of plants (Li et al. 2013 and Zhang et al. 2013). Furthermore, Arabidopsis plants showing overexpression of *AtHSFA6a* showed improved tolerance towards salinity stress (Hwang et al. 2014). Similarly, overexpression of *OsHSFA7* has improved salinity and drought tolerance of rice plants (Liu et al. 2013). However, HSFA5 are the repressors of HSFA4 in both Arabidopsis and tomato (Baniwal et al. 2007), and HSFA9 regulate the process of seed development in sunflower and leads to the expression of small heat shock proteins involved in seed maturation (Diaz-Martin et al. 2005, Almoguera et al. 2002). These studies suggest the functional complexity of HSFs among different plant species, and their indispensable role in providing tolerance against a plethora of stress encountered by plants during their life cycle. Hence, HSFs could be the potential targets for genetic manipulation, in order to generate thermotolerant crops.

Heat shock factors regulate the expression of many heat shock inducible genes including heat shock proteins (HSP) to activate heat shock response. The induction of these HSPs as molecular chaperones takes place early during the onset of heat stress in most of the plants, to proofread and maintain the quality control of distorted proteins in cells (Glover and Lindquist 1998). During heat stress, molecular chaperones bind reversibly to their substrate proteins in an ATP-dependent manner and ensure their correct folding, disaggregation, and solubilization (Hendrick and Hartl 1993, Parsell and Lindquist 1993). For example, high temperature leads to denaturation and misfolding of cytosolic proteins, consequently, free HSP70 molecular chaperone pool decreases in the cytosol, in order to repair distorted polypeptides (Schoffl et al. 1998). These cellular pools are maintained by induced transcription of HSP70 proteins, regulated by HSFs (Czarnecka-Verner et al. 2000). Heat stress induced Ca^{+2} ion signaling and engagement of HSPs

in chaperoning, activate the HSFs which in their trimeric form binds with *heat shock element* (HSE) in the promoters of HSP70, and other heat shock inducible genes (Schoffl et al. 1998). In turn, after stress, the free HSP70 binds to the HSF and attenuate their activity by negative feedback control (Zou et al. 1998). Moreover, the functional and expression analysis of *HSP70* family genes in various crops including Arabidopsis, rice, tobacco, soybean etc. define their role in providing tolerance against heat and other abiotic stresses.

The strongly induced expression of HSP90 genes is observed during various abiotic stresses along with heat stress (Krishna and Gloor 2001 and Banilas et al. 2012). In addition to their molecular chaperone function, HSP90 proteins have a specialized role in signal transduction cascades, by binding with the substrates including transcription factors, protein kinases and steroid receptors (Picard et al. 1990, Shigeta et al. 2015 and Clement et al. 2011). Furthermore, the HSP100/ClpB proteins with HSP70 family proteins form bi-chaperone machinery and lead to the clearance of degraded and harmful polypeptides (Parsell et al. 1994). HSP100/ClpB proteins are also crucial for maintaining acquired thermotolerance and acclimation to harsh temperature (Hong and Vierling 2000). Genome-wide analysis revealed the presence of eight and nine members of HSP100 family genes respectively in Arabidopsis and rice (Agarwal et al. 2001 and Singh et al. 2010). In Arabidopsis, rice and wheat the abundant accumulation of transcripts of *ClpB* genes were observed during heat, cold, drought and ABA treatments (Campbell et al. 2001, Lee et al. 2007 and Singh et al. 2010). Interestingly, their role in grain development of cereals like rice, wheat, and maize under normal and heat stressed conditions has been reported in several studies (Singla et al. 1998).

Heat shock proteins are crucial for the survival of plants in physiological and stressed conditions. Therefore, genome-wide identification and functional analysis studies of these proteins in cereal crops are pivotal, in order to get insights of their functional diversification, and elucidation of mechanisms involved in providing tolerance to plants under heat and other abiotic stresses.

1.4 Alterations in metabolic fluxes due to heat stress

Given that heat stress alters molecular and physiological metabolism of plants including a change in production of secondary metabolites. Hence, these altered metabolic fluxes need to be restored, in response to temperature fluctuations. Synthesis of certain secondary metabolites is

required for osmotic adjustments, stabilization for cellular membranes, chelating toxic ions and alternative energy sources (Thomashow, 1999 and Guy, 1990). These metabolites comprise of carbohydrates, polyamines, amino acids, lipids, and certain organic acids, contributing towards stress tolerance (Diamant et al. 2001). The newly synthesized compounds upon stress have also been found to be involved in signal transduction. These signaling molecules include hormones like salicylic acid, ROS produced due to membrane damage, certain phenolic compounds and antioxidants, which produced early during temperature fluctuations and prevent irreversible damage to cells (Kotak et al. 2007 and Shulaev et al. 2008). Here, the metabolomics offers the comprehensive analysis of metabolites involved in establishing metabolic homeostasis and stabilizing plant growth in adverse environmental conditions (Kaplan et al. 2004). Moreover, the flow of information from genes to metabolites consists of several intermediate pathways. Consequently, the slight alteration in gene expression causes significant alterations in metabolite concentrations (Bino et al. 2004). Hence, the exploitation of metabolic approaches revealed the overall biochemical status of metabolites at the cellular level, required for acclimatization of plants under adverse temperature fluctuations.

1.5 Objectives for the present study

Keeping aforementioned points in mind in the present study we have chosen barley for functional genomics and metabolomics studies with the objectives as follows:

1. Identification, phylogenetic and expression analysis of high molecular weight heat shock protein families viz. *HSP70*, *HSP90* and *HSP100* in vegetative and seed developing stages.

2. Over expression of an A type *HvHSFA2c/A6b* gene in barley through transgenic approach.

3. Analysis of transgenics for functional analysis of *HvHSA2c/A6b* gene and response towards heat stress.

4. Comparative study of changes in metabolome of barley under normal and heat stress conditions.

Review of literature

Review of literature

2.1 Climate change and crop production

Global warming and changing weather patterns are a serious threat to global agriculture productivity and food security. A recent report of IPCC (2018) on climate change showed that earth is going to be 1.5°C warmer by 2030-2050 (Fig. 2.1 a,b). The world population is also most likely to grow around 2.3 billion more till 2050 (Fig. 2.1c) (FAO, 2018).

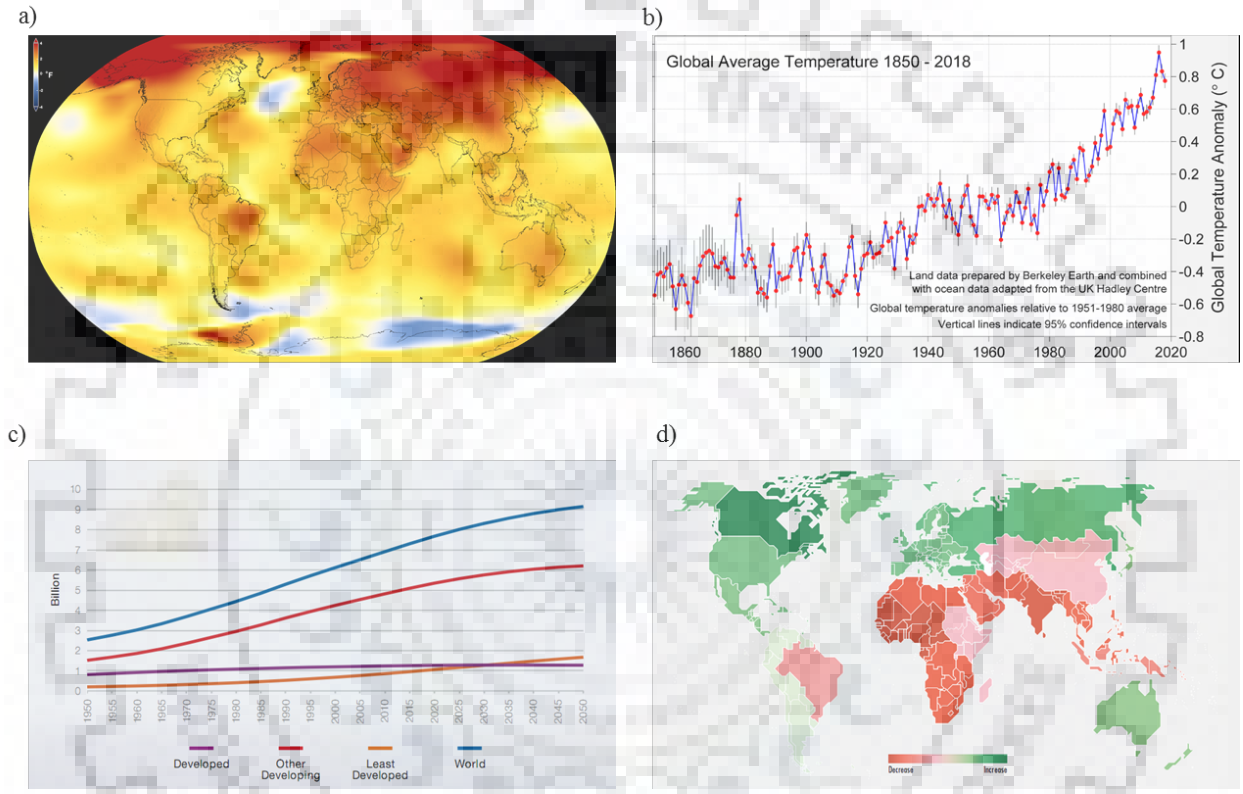

Fig. 2.1 Global temperature, population and agriculture production trends. (a) Comparative rise in average earth's temperature in 2017 as compared to 1980 (NASA's Goddard Institute for Space Studies. (b) Global mean temperature during since 1850 and highest rise in 2018 (http://berkeleyearth.org/2018-temperatures/). (c) Projected increase in global population till 2050 (UN population division, from van der Mensbrugghe et al. 2009). (d) Influence of climate change in GDP relative to baseline.

Cereals are feeding almost two third of the global population, and slowdown in their production has been observed due to adverse environmental conditions (FAO, 2018). South and south east Asia including India and China make most populated area of earth and agriculture production in these areas is highly vulnerable to climate change (Fig. 2.1d) (FAO, 2018). It is estimated that with each degree rise in temperature, the average decrease in yield of three important cereals viz.

rice, wheat and maize are projected to be 3-10% (FAO, 2018, Asseng et al. 2017) and overall decline in production is estimated around 30% in South Asia alone.

Seasonal changes with longer summers and smaller winters limit the crop production as higher temperature generally disturbs the plant phenology and shortened the plant life cycle with reduced yield (Lobell and Gourdji 2012). These climatic fluctuations have more severe impacts on developing countries like India, where 60% of its population is dependent on agriculture. Furthermore, an estimated loss in agriculture productivity in India is around 2.6% due to increasing trends of temperature (FAO, 2018). Therefore, there is an urgent need to intensify the agriculture production, in order to meet the global food demands.

2.2 Effects of high temperature stress on plants

2.2.1 Vegetative development

Heat stress has devastating effects on the growth and development of plants. It has severe impacts on all the aspects of plant life cycle i.e. from seed germination to seed maturation (Porter and Moot 1998). Most importantly vital processes like photosynthesis and respiration are perturbed, which reduces the plant fitness and lifecycle (Stone, 2001). Cellular and thylakoid membranes are highly prone to heat injury. Photosystem II (PSII) is present on thylakoid membranes and alteration in membrane fluidity due to heat stress could disturb the assembly of PSII components (Havaux, 1992). Furthermore, heat stress (temperature ranging between 30°C to 42°C) , disrupt the process of electron flow through Z cycle, denature the enzymes involved in oxygen evolving complex of PSII, and cause reduction in regeneration of RuBP pools which thereby leads to photosynthesis inhibition (Salvucci and Crafts-Brander et al. 2004). Enzymes involved in energy generating and carbon assimilation process are highly sensitive to high temperature, such as ribulose 1,5 bisphosphate carboxylase (Rubisco), rubisco activase and other enzymes of Calvin cycle, reduction in their activity directly disrupt the processes of carbon assimilation (Sharkey, 2005). Moreover, high temperature enhanced the activity of chlorophyllases which destroys the functional activity of chlorophyll (Todorov et al. 2003). On top of this, disruption in the activity of photosystems I and II makes them a factory of production of reactive oxygen species (ROS) of various types, such as superoxides and peroxides, which cause irreparable damage to plant cells (Asada, 2006). The ROS production due to heat stress

cause, membrane peroxidation, which disrupt the permeability and fluidity of plasma membrane. Ougham et al. (2008) suggested that heat stress induce leaf senescence, by disrupting the functions of chlorophyll, cytoskeleton, and cellular and thylakoid membranes. All these processes ultimately decrease the photosynthetic and respiratory ability and consequently limit the production of energy reserves required for growth and reproduction of plants. The premature senescence of leaves disturbs the process of distribution of photosynthetic assimilates form source (leaves) to sink (grains) through phloem. Also, the process of seed filling in plants is highly dependent on the availability of carbohydrate assimilates produced by leaves doing photosynthesis, and the reserves already available in vegetative tissues like stem and leaves, which later on translocate to developing seeds (Plaut et al. 2004).

Heat stress affects all the physiological and metabolic process of plants at all growth stages. Leaf scorching, senescence, paleness due to chlorophyll degradation and retarded growth are some visible symptoms of heat stress on plants (Ismail and Hall, 1999). High temperature lowers the germination percentage, seedling vigor, seedling survival and establishment in several plant species including wheat and rice (Essemine et al. 2010, Piramila et al. 2012). Iloh et al. (2014) reported the stagnant growth of rice, maize and sorghum at higher temperature ranges between 42-50°C. Further, in case of wheat and rice, plants produce less number of tillers with decreased biomass accumulation and stunted plant growth (Hasanuzzaman et al. 2013). Thus, heat stress reduced the reproductive fitness and yield of crop plants.

Heat stress hampers the reproductive development phase of plants including flowering, fertilization, seed filling and maturation of cereal plants resulting into yield deprivation.

2.2.2 Reproductive development

Reproductive phase of plant starts from flower development and ends at developed matured seed. The whole reproductive phase includes several intermediate stages including development of male and female gametophyte, pollination, fertilization, grain filling and maturation. Heat stress adversely affects the plant reproduction process and mainly the development of male gametophyte (Sakata et al. 2000). However, the female gametophyte is comparatively less sensitive to heat stress. Sakata et al. (2000) reported that heat stress in barley during development of male gametophyte cause cytoplasm abortion, formation of short anthers without pollen grains

and less accumulation of starch in pollen grains at different stages such as early stage of panicle differentiation, prior and post meiosis of pollen mother cells respectively. Moreover, Oshino et al. (2007) studied the effect of high temperature during reproductive stages of barley and found the formation of short anthers devoid of pollen grains. They observed the abnormal development of tapetum and pollen mother cell upon heat stress which leads to complete sterility of male gametophyte, although the pistil development was normal. Similarly, Endo et al. (2009) also reported the abnormal tapetum development and pollen sterility due to heat stress in rice.

In angiosperms, the process of fertilization is immediately followed by the seed development, which is divided into two phases i.e. seed enlargement and seed filling. The initial phase involves rapid cell division followed by cell enlargement and in second phase the mobilization of stored photoassimilates to the developing grains takes place (Briarty et al. 1979). Yang and Zhang, (2006) described that assimilates for grain filling is provided by photosynthetically active leaves and reserved in vegetative tissues like stems, stored during pre-anthesis stages, majorly contribute to the final grain size and weight. As described above, heat stress during vegetative phase of plants hampers the process of photosynthesis resulting into plant starvation, leads to the poor grain filling and yield loss of crop plants. In addition to this, sucrose transporters play crucial role for mobilization of sucrose from source to sink (Xu et al. 2018). However, *sucrose transporter* (*SUT*) gene showed plant wise differential expression pattern, for example, in heat stressed Arabidopsis, rice and barley plants *SUT2*, *OsSUT1*, *HvSUT1* and *HvSUT2* respectively were downregulated, whereas, in Poplar *PtaSUT4* was upregulated. Downregulation of this gene further lowers the quality of grains due to disturbed carbohydrate allocation. Heat stress further fastens the process of seed filling (Xu et al. 2018).

2.3 Signaling of heat stress

Plants have evolved an ability to withstand nonlethal stressful conditions. For instance, for heat stress tolerance they have sophisticated mechanisms to transduce a signaling cascade, which initiates at the perception of mild changes in ambient temperature as a signal and ends at expression of genes involved in plant defense. Perception of signal initially takes place at plasma membrane, and transient and mild changes in temperature alter the membrane fluidity (Vigh et al. 2007). These heat induced changes in fluidity of plasma membrane favors the influx of Ca^{+2} ions (Saidi et al. 2009). Ca^{+2} ions are the secondary messengers which convey information along

with calmodulin proteins (CaM) to their downstream target proteins and evoke a variety of cellular responses (Snedden and Fromm, 2001). For confirmation of their role in heat stress signaling, Liu et al. (2005) used transgenic lines of Arabidopsis overexpressing *AtHsp18.2 promoter/GUS* fusion gene. They found the enhanced activity of reporter gene when calcium chloride was provided to the plants in presence or absence of heat stress. The inverse effect was observed in the presence of EGTA a chelating agent of calcium ions. Further the induced expression of *AtCaM* genes during heat stress supported the Ca^{2+}-CaM dependent expression of *Hsp* genes. Calcium ions are also known to govern the activity of heat shock factors (HSFs) which are the terminal and crucial components of heat stress signaling cascade. Li et al. (2004) identified that the Ca^{2+}-CaM enhanced the binding affinity of HSFs to conserved heat shock elements (HSE). At 44°C, addition of EGTA a calcium chelator to the whole cell extracts of maize reduced the binding ability of HSFs to HSE which was later restored by adding $CaCl_2$ and not by $MgCl_2$. Similar observations were obtained by adding CaM antagonist chlorpromazine and restoration in activity by adding CaM to the cell extracts and not by BSA. They further proposed that Ca^{2+}-CaM mediate the expression of *Hsp* genes through HSF regulation. In their later study, Liu et al. (2008) further reported that AtCBK3 a CaM-binding protein kinases (CBK) leads to the phosphorylation of AtHsfA1a and regulates the gene expression of Hsps. They found that knockout lines of *AtCBK3* showed defect in basal thermotolerance whereas, overexpression reversed the phenotype. It has been reported by Sangwan et al. (2002) that influx of Ca^{2+} during heat stress is essential for the activation of mitogen-activated protein kinase (MAPKs). Calcium ion mediated signaling activates several protein kinase enzymes including calcium dependent protein kinase, (CDPKs), MAPKs, and RboH enzymes during heat stress (Suzuki et al. 2011, Sinha et al. 2011, Sinha et al. 2015). Recently, Wang and Huang, (2017) determined the effects of Ca^{+2} ions and other signaling molecules on thermotolerance of tall fescue. Exogenous applications of calcium chloride and phosphatidic acid significantly up-regulated *HSFA2c* gene which was thereby considered to regulate the expression of heat shock protein genes including *Hsp18*, *Hsp70*, and *Hsp90*. These studies suggested that heat stress sensed by plasma membrane is further propagated in cell using calcium ions as secondary messengers.

Furthermore, heat induced accumulation of proteins in cell cytosol and endoplasmic reticulum (ER) evokes an unfolded protein response. The cytosol specific unfolded response is regulated by HSFA2 in Arabidopsis plants (Sugio et al. 2009). Whereas, unfolded protein response in ER

is mediated by brassinosteroid signaling and transcription factors of bZIP class i.e. bZIP17 and bZIP28 which resides in ER during nonstressed conditions. These transcription factors are translocated to Golgi for proteolytic cleavage by site 2 protease and then exported to nucleus where they induce the expression of ER specific chaperones (e.g. *BIP3*) genes which provides stress specific acclimation to plants (Che et al. 2010).

In wheat, through a transcriptomics study Chauhan et al. (2011) found that in vegetative stage, the heat shock induced changes in gene expression returned to normal within two days of recovery, however, this is not the case during reproductive stage, suggesting different type of response in vegetative and reproductive stages. Interestingly, Kumar and Wigge, (2010) on the basis of transcriptome analysis found that nucleosome remodeling gene (*ARP6*) is involved in providing thermosensory responses to Arabidopsis plants. The replacement of classical histone H2A by H2A.Z in nucleosome, promoted by *ARP6* gene was found essential for perceiving the signals from warm environment. Nucleosomes containing modified histone H2A.Z could sense heat stress and regulate the gene expression. It was further proposed that such modifications were primarily found in the promoters of heat shock responsive genes for example Hsp encoding genes (Mittler et al. 2012).

Heat stress leads to the production of reactive oxygen species, and ROS mediated signaling is found conserved during several biotic and abiotic stresses (Mittler et al. 2004). To maintain the normal cellular processes, the scavenging of ROS and their involvement in signaling occurs under tight regulation (Mittler et al. 2004). In plants, key ROS producing enzymes are NADPH oxidase and respiratory burst oxidase homologues (Rboh) family proteins (Suzuki et al. 2011). RbohD mediated ROS signaling cascade involve MAPKs, MBF1c, SnRKs and certain HSFs (Mittler et al. 2004, Jalmi and Sinha, 2015). Moreover, ROS production can also directly lead to the activation of MAP kinases. These MAP kinases further modulate the expression of genes involved in stress associated signaling cascades, as revealed in MAPKK6 overexpressing transgenic lines of rice (Kumar and Sinha, 2014). During HSR, two MAP kinases namely MAPK3 and MAPK6 regulate the expression of heat shock protein genes (Wang and Huang, 2017) and promote the nuclear localization of heat shock factor HsfA2.

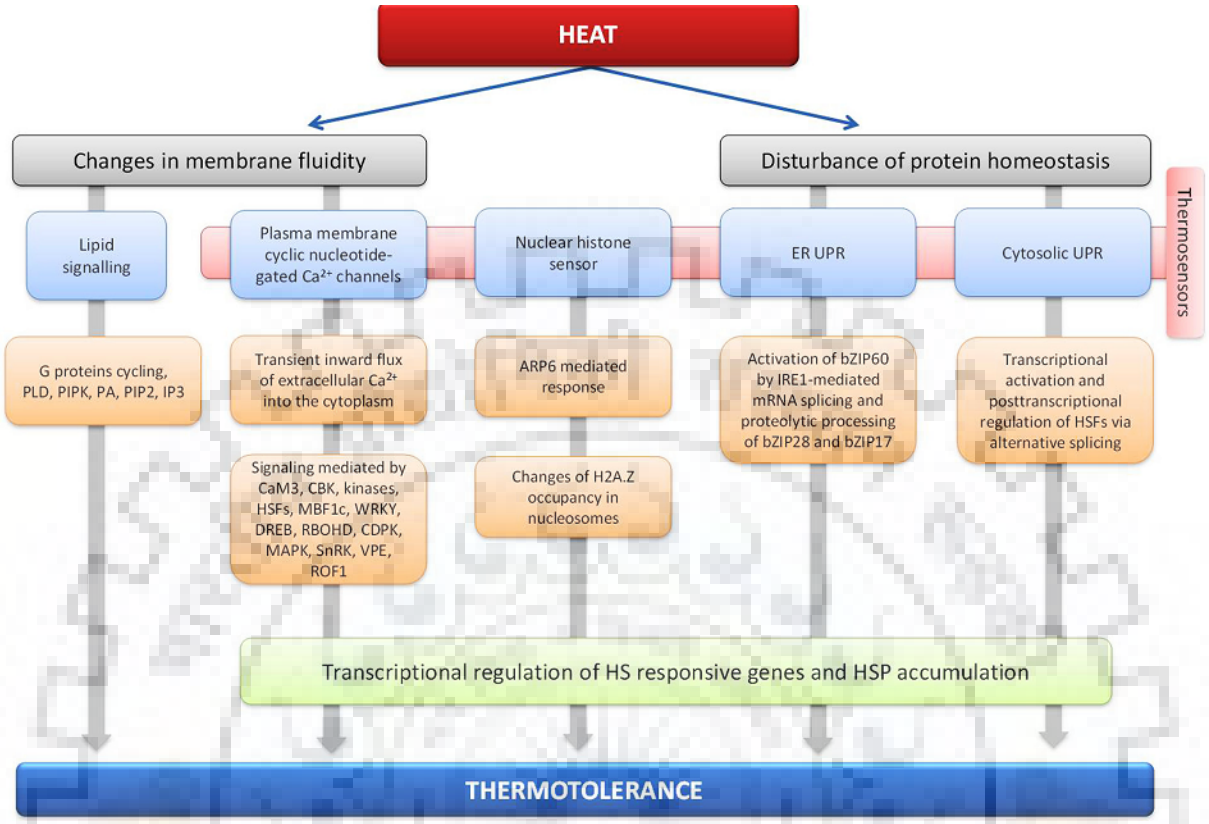

Fig. 2.2. The thermosensors of plants and their downstream signaling cascade evoked to establish heat stress tolerance. The heat stress signals are perceived by different thermosensors including plasma membrane, cytosolic and ER unfolded protein response and nuclear histone sensors which evokes a signaling cascade leading to the expression of heat shock responsive genes for providing tolerance against high temperature. Source (Bokszczanin et al. 2013)

2.4 Heat shock factors identification

Heat shock response in plants is regulated by transcription factors called as heat shock factors (HSFs). These HSFs binds to the heat shock elements (HSE), which are the conserved motifs of defined consensus nAGAAnTTCT and present in the upstream regulatory regions of heat shock responsive genes (Pelham and Hugh, 1982, Amin et al.1988, Nover et al. 2001 and Kroeger and Morimoto, 1994). Earlier, to understand about the regulators of heat shock response, Sin, (1975) injected the extracts of heat shocked cells to the larval salivary glands of Drosophila and observed the puffing in polytene chromosome. Similarly, Compton and McCarthy, (1978) incubated the polytene chromosomes with disrupted heat shock treated Kc cells of Drosophila and also observed the formation of puffs. These results suggested the active transcription at the heat shock loci of polytene chromosome. However, the purification of these promoter binding

transcription factors was become possible by employing the foot-printing analysis of *hsp70* gene of Drosophila (Parker and Topol, 1984) and HSE based affinity chromatography assay (Topol et al. 1985). These purification studies efficiently allowed the cloning of first HSF from yeast (*S. cerevisiae*) (Sorger and Pelham, 1988).

In case of plants, HSFs were first studied in tomato by Scharf et al. (1990). They provided the heat stress to tomato cell cultures which leaded to the activation of a pre synthesized heat shock element binding protein. They isolated three different HSF genes from λgt11 cDNA expression library through South-Western screening. Among these three putative HSFs genes, one (*Lp-HSF8*) was constitutively expressed whereas expression of the other two genes (*Lp-HSP24* and *Lp-HSF30*) was observed as heat induced. The identification of these three functionally independent HSF genes in tomato indicated the possibility for the presence of more HSF genes in plant system. Furthermore, with the advancement of whole genome sequencing technology and availability of myriads of genomic resources, the genome-wide identification studies for HSF family genes have been performed among various plant species. The HSF family in plants has various and functionally diversified members resulting due to gene duplication and whole genome duplication events during evolution (Scharf et al. 1990). The number of HSF family members identified in different plant systems ranges from lowest 21 in Arabidopsis, 22 in barley, 25 in rice, and all the way to highest 56 in wheat (Nover et al. 2001, Chauhan et al. 2011; Reddy et al. 2014 and Xue et al. 2014).

2.5 Structure of heat shock factors

Despite being multiple in numbers and their functional diversification, all HSFs possess a modular structure (Fig. 2.3). These heat shock factors are transcription regulatory proteins and consist of a conserved DNA binding domain (DBD) at their N-terminus (Scharf et al. 1990). The NMR and X-ray crystallography studies of 3-D structure of HSF N-terminal domain revealed the presence of three helical and four beta stranded structures (Damberger et al. 1994, Harrison et al. 1994 and Schulthesis et al. 1996). The presence of helix turn helix motif in the hydrophobic core of this N-terminal domain is required for its precise binding to the heat shock element consensus sequence. DNA binding domain is connected to oligomerization (OD) through a flexible linker. The OD provides a remarkable ability of protein-protein interaction i.e. trimer formation of HSFs by forming a leucine zipper like coiled coil structure (Peteranderl et al. 1999). This domain

consists of two heptad regions of hydrophobic amino acids HR-A and HR-B, and the inserts of variable number of amino acids between both the regions. Peculiarities like length of linker connecting C-terminal domain of DBD to OD and insertion of 21 and 7 amino acids between HR-A/B regions categorized HSFs into class-A and –C, respectively. However, class-B HSFs has compact HR-A/B regions (Nover et al. 1996). Furthermore, heat shock factors are dynamic proteins possess both nuclear localization and nuclear export signals (Nover et al. 1996). The AHA motif is present in almost all class-A HSFs except HSF8 types (Kotak et al. 2004). Doring et al. (2000) studied the C- terminal domain of HSFA1 and A2 of tomato and reported that the presence of AHA motifs is required for their transactivation potential. Interestingly, HSFs categorized in class B consist of a characteristic -LFGV- motif and act as transcriptional repressors (Czarnecka-verner et al. 2004). Hence, trans-activating potential of class A HSFs make them interesting targets for further studies.

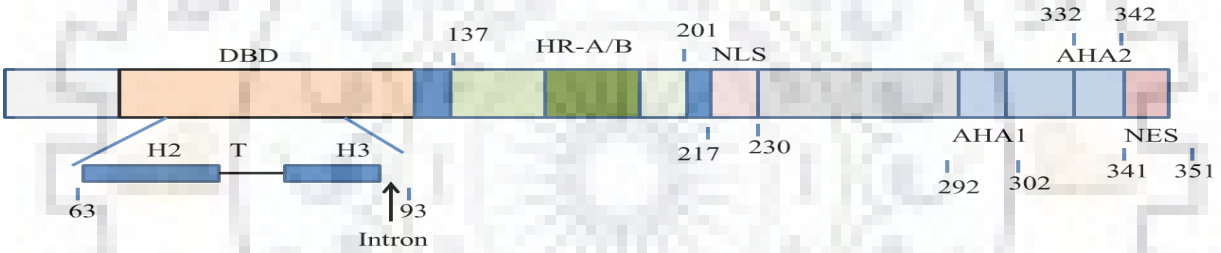

Fig. 2.3. Structure of AtHSFA2. A typical Hsf consist of a DNA binding domain, oligomerization domain (HR-A/B), nuclear localization signal (NLS), a transactivation domain AHA1 and AHA2) and a nuclear export signal (NES). Adapted from Nover et al. (2001)

2.6 Expression and regulation of heat shock factors

Heat shock factors are the terminal components of heat stress signaling cascade, which leads to the transcriptional reprogramming of many heat shock responsive genes. The studies on expression profiling of heat shock factors in various abiotic stress conditions have been performed in many plant species including Arabidopsis, rice, wheat, maize, tomato, and apple etc. Nover et al. (2001) found the strong induction of *HSFA2* during heat stress in *Arabidopsis* cell cultures versus no expression under controlled conditions. Likewise, the expression profiling of wheat HSFs was performed by Xue et al. (2014), where they observed the higher accumulation of transcripts of *HSFA2* and *HSFA6* during heat stress. Similarly, Lin et al. (2011) observed the upregulation of class-A2 HSFs *ZmHSF-01* and *ZmHSF-04* upon providing the heat stress. Moreover, Chauhan et al. (2011) studied the expression patterns of rice HSFs in tissue

specific manner, comprising shoot, root and developing panicle. They found that transcripts of *OsHSFA2a* showed higher accumulation during early (30 minutes) and prolonged heat stress (2 hour). Furthermore, the expression studies in plant species like apple and pepper by Giorno et al. (2012) and Guo et al. (2015) respectively also showed the accumulation of mRNAs of *HSFA2* group in response to heat stress. These studies indicated that the class HSFA2 is of major importance in mediating heat shock response in plants.

2.6.1 Transcriptional and post transcriptional regulation of HSFs

The expression of heat shock factors is regulated at level of transcription, translation, post transcription, and post translation (Fragkostefanakis et al. 2015). The gene regulation at transcriptional level is mediated by transcription factors (TF) which recognize the *cis*-acting upstream regulatory regions lying in their promoter sequences. In case of HSFs, Kotak et al. (2007) found that in *Arabidopsis thaliana* the expression of HSF9, a seed specific transcription factor is controlled by *ABSCISIC ACID–INSENSITIVE3* (*ABI3*) gene. Furthermore, Yoshida et al. (2008) reported the high upregulation of *HSF3* gene in transgenic Arabidopsis plants overexpressing *dehydration response element binding protein 2A* (*DREB2A*) transcription factor during heat stress. Nishizawa-Yokoi et al. (2011) demonstrated that *HSFA2* consist of HSE in the 5' promoter region and showed induced expression during high light and temperature stress. They further performed transient reporter assays and postulated that *HSFA1d* and *HSFA1e* bind to the HSE region of *HSFA2* and regulate its expression in *Arabidopsis*.

Moreover, Liu et al. (2013) reported the post transcriptional regulation of *AtHSFA2*, through alternative splicing induced during heat stress. They found that an intron is removed by utilizing cryptic 5' splice sites from DBD during alternative splicing, resulting into the formation of a splice variant *HsfA2-III*. This is translated into a functional protein S-HsfA2, which recognize the HSE consensus sequence present proximally to TATA box in the promoter region and auto regulates the expression of *HSFA2* gene under heat stress conditions. Similarly, in rice, three alternatively spliced variants of *OsHSFA2d* (*OsHSFA2dI*, *OsHSFA2dII*, *OsHSFA2dIII*) were reported by Cheng et al. (2015), where heat stress was found to induce the expression of *OsHSFA2dI,* which encodes a nuclear localized protein consisting of 357 amino acids involving in transcriptional activation. During heat stress, the same variant also regulates the expression of *OsBip1* gene involved in mediating the unfolded protein response in plants. Whereas, *OsHSFA2dII* lacks DBD, OD and CTD and is a transcriptionally inactive dominant variant

which is reconstructed during splicing of *OsHSFA2d* under normal conditions or induced by photoperiod treatment. Furthermore, He et al. (2008) identified the phenomenon of alternative splicing in alfalfa *MsHSF1* and found the presence of premature stop codon in all the transcripts. These transcripts were degraded by non-sense mediate decay. Sugio et al. (2009) reported that the alternative spliced form *HSFA-II* of *AtHSFA2* gets accumulated and regulate the cytosolic protein response, during heat stress. Such studies suggested that alternative splicing of HSFs is a conserved phenomenon in plant species.

2.6.2 Translational and post translational regulation of HSFs

HSF regulation is also monitored at translational and post translational level. Conserved peptide in upstream ORF regions regulates the expression of several transcription factors (Jorgensen and Dorantes-Acosta, 2012). The involvement of upstream open reading frames (uORF) in HSF regulation is reported only in Arabidopsis so far (Zhu et al. 2012). They reported the presence of uORFs in seven members of HSF family. Among which *HsfA1d*, *HsfA1e*, *HsfA2*, *HsfA4a* consist of one, *HsfB2b*, *HsfB1* two and *HsfC1* possess three uORFs. They further observed that translation of main ORF of *HSFB1* is repressed by its uORF2 and released during heat stressed conditions. Furthermore, after translation, Evrard et al. (2013) found the protein kinase MAPK6 phosphorylates HSFA2 at T239 position and induce its nuclear shuttling during heat stress. Nishizawa-Yokoi et al. (2010) demonstrated that the accumulation of polyubiquitinated proteins and 26S proteasome inhibition induce the HSFA2 expression. These studies suggested the posttranslational regulation of HSFs.

Small RNAs such as mi-RNAs (micro RNAs) was initially found in metazoans and also widely reported in plants (Reinhart et al. 2002). miRNAs are the small stretches of 20-24 non-coding nucleotides which regulates plethora of functions during growth, development and adverse environmental conditions in plants (Baker et al. 2005, Zhou et al. 2007, Mutum et al. 2013, and Balyan et al. 2017). Furthermore, Guan et al. (2013) found the role of miRNA in providing heat stress specific responses in Arabidopsis. They observed the higher induction of miR398 in Arabidopsis plants exposed to high temperature stress. However, plants overexpressing the mutated form of miR398 exhibiting hypersensitive responses towards heat stress. Authors also suggested that HSFA1b and HSFA7b were associated with the heat stress-induced transcriptional regulation of this miRNA.

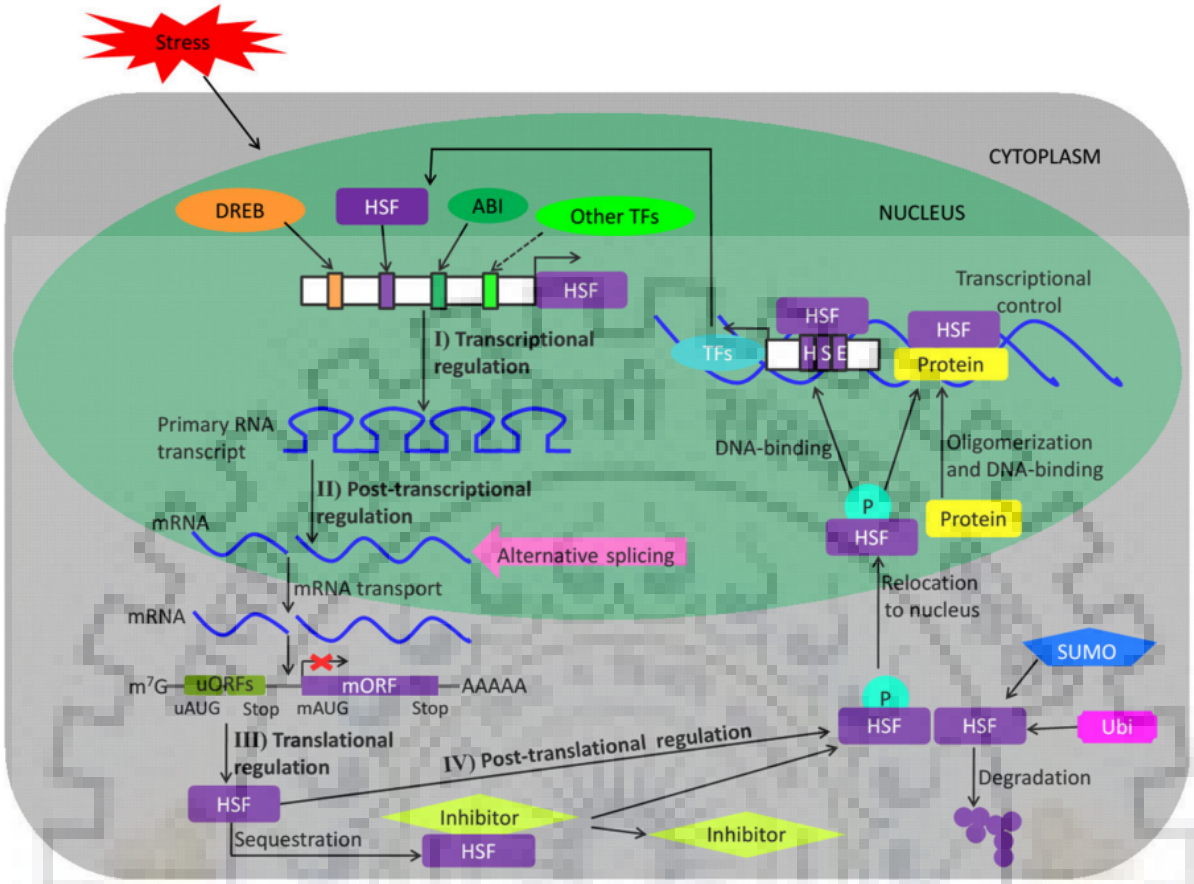

Fig. 2.4. Regulation of heat shock factors at transcription, post transcription, translation and posttranslational level. (I) Binding of DREB, ABI and other TFs to cis-regulatory region of HSF encoding genes regulate their transcription. (II) Alternative splicing mediating regulation of HSFs. (III) uORFs present in UTRs of HSFs governs their translation. (IV) Post translational modifications like phosphorylation, sumoylation, ubiquitination etc. govern the activation and degradation of HSF proteins. (Source: Guo et al. 2016)

2.7 Functions of heat shock factors

2.7.1 *HSFA1a* as a master regulator

As discussed above, the HSF family in plants is highly diversified with multiple members having unique functions. Mishra et al. (2002) performed the functional analysis of *HSFA1* in tomato plants for which they developed the transgenic plants showing overexpression (OE), and co-suppression (CS) through RNA silencing of *HSFA1a* gene. Under normal conditions, they did not observe any change in growth parameters of wild type (WT), OE and CS lines of tomato. However, the plants and fruits of CS lines were highly susceptible to heat stress due to deprived synthesis of HSFs and HSPs. They further suggested the role of *HSFA1a* as a master regulator of

other HSFs during heat shock. Interestingly, Liu et al. (2011) performed the knockout studies in *Arabidopsis* to dissect the function of class-A1 HSFs (*HSFA1a*, *A1b*, *A1d* and *A1e*) as master regulator of thermotolerance. They observed HSFs viz. *HSFA1a* and *A1d* as strong inducers of heat shock response as compared to that of *A1b* and *A1e*. Moreover, upon transcriptome analysis upregulation of nearly 65% of HSR responsive genes was observed as *HSFA1* dependent. Zhu et al. (2006) identified a heat shock factor *GmHSFA1* in *Glycine max* through homology searches with tomato *LpHSFA1* from EST database. Moreover, its constitutive expression pattern was identified under different conditions. Simultaneously, they examined the overexpression of *GmHSFA1* by developing transgenics in soybean. These plants showed enhanced tolerance towards heat stress as compared to non-transgenic wild type plants. The upregulation of *GmHsp70* was also observed in transgenic plants during normal conditions.

2.7.2 *HSFA2* and its role in thermotolerance

Nishizawa et al. (2006) isolated *HSFA2* gene by suppression-subtractive hybridization along with other 76 genes induced during high light and temperature stress. The transcripts of this HSF were most abundant during stressed conditions. They subsequently identified the role of *HSFA2* in providing tolerance against several abiotic stresses including heat, high light, and hydrogen peroxide stress. The overexpression of *HSFA2* under the *CaMv35s* promoter in *Arabidopsis* induced the expression of approximately a set of 46 genes including HSPs and *Apx2* (*ascorbate peroxidase 2*). Comparatively, in knockout plants of *Arabidopsis* for *HSFA2* gene, 26 genes were found to be strongly down regulated, during thermal (2 hours) and high light stress. Moreover, the overexpressing transgenic plants were found to be tolerant towards adverse environmental challenges. As a result, they inferred *HSFA2* as a potential target for genetic modifications to generate thermotolerant plants. Charng et al. (2007) analyzed the T-DNA insertion mutant lines of *Arabidopsis* under severe heat stress provided for two days after recovery from sub lethal temperature. They found that knockout mutants for *HSFA2* was highly sensitive towards heat shock among all other 48 heat shock genes used for mutant analysis. Moreover, thermotolerance of these plants was restored by supplementing the copy of *HSFA2* gene. The downregulation of several heat shock induced genes in these mutant lines was also observed, including the *HSP* genes. These findings suggested the importance of *HSFA2* gene in providing acquired thermotolerance.

Expression analysis studies showed the dominant higher expression of *HSFA2* during heat stress. *HSFA2* works in cooperation with *HSFA1* by forming hetero oligomer complexes. Through developing transient expression system in tobacco protoplast cell cultures Scharf et al. (1998) revealed that *HSFA2* resides in cytosol and requires the co-expression of *HSFA1* for its nuclear translocation, and the interaction between two HSFs was confirmed by two hybrid tests in yeast. In addition, it was revealed that HSFA2 remain in cytosol in form of soluble protein during ambient conditions, nonetheless in response to HS it mainly remains nuclear localized and also stored in cytoplasmic granules. Eventually, they inferred HSFA2 as a stable protein, which remain accumulated during longer HS durations and recovery phases. *HSFA2* was determined to behave as a dominant HSF in providing thermotolerance and also to function independently of *HSFA1* which is considered as master regulator of other heat shock factors (Liu and Charng 2013). Authors found that ectopic constitutive expression of *HSFA2* can complement the function of *HSFA1* in the quadruple knockout mutants of *HSFA1* (*A1a, A1b, A1d, A1e*). Subsequently, transcriptome analysis revealed that both of these transcription factors have different target genes under their regulation. *HSFA2* mainly regulates the expression of genes mediating redox homeostasis and metabolism, whereas *HSFA1* mostly regulates the genes responsible for transcriptional functions and regulation (Liu and Charng 2013). Interestingly, in plants a significant overlap in tolerance mechanisms towards different stresses has been observed in several cases, for example, Banti et al. (2010) gave individual and combined heat and anoxia stresses to *Arabidopsis* seedlings. They analyzed their responses employing microarray analysis at genomic level and found overlap between genes involved in both the responses. Evidently, *HSFA2* involved in providing tolerance towards various form of abiotic stresses, also showed maximum induction towards anoxia stress in *Arabidopsis*. Furthermore, authors observed that heat stress treated plants showed acclimation towards anoxia stress, whereas, mutant plants lacking *HSFA2* and double mutants lacking both *HSFA1a*/*A1b* and *HSFA2* become sensitive towards anoxia stress (Banti et al. 2010). These findings implicate the role of all three HSFs in conferring tolerance towards anoxia stress. Moreover, the overexpression of *HSFA2* in Arabidopsis induces the expression of genes involved in providing tolerance towards anoxia stress. Ogawa et al. (2007) performed overexpression studies of *HSFA2* in Arabidopsis. For overexpression *HSFA2* gene was cloned under *CaMV35S* promoter having enhancer like sequences and omega sequences of tobacco mosaic virus (*El2Omega*) for increasing translational

efficiency. Simultaneously, authors also generated the mutated plants carrying truncated *HSFA2* without AHA and NES at its C-terminal. The lines overexpressing *HSFA2* gene (around 430 fold) showed dwarf phenotype and enhanced basal and acquired thermotolerance (Ogawa et al. 2007). However, plants with dominant negative mutation showed compromised thermotolerance. Therefore, these findings confirmed the role of *HSFA2* in providing tolerance against heat stress. Moreover, genes imparting tolerance towards heat and other abiotic stresses were identified by microarray analysis of overexpressing transgenic lines. These plants also showed elevating tolerance in response to salinity and osmotic stress concluded that *HSFA2* gene play wide role against different stresses.

2.7.3 *HSFA2* regulates the memory of heat stress

Besides conferring thermotolerance to plants, HSFA2 also has been known for providing a memory of heat stress in plants by regulating the expression of genes involved in HSR and extending the duration of acquired thermotolerance (Charng et al. 2007). At molecular level, the accumulation of H3K4 (histone-3-lysine-4) hyper-methylation at memory loci, marks the transcriptional activation patterns of recently active genes during heat stress (Lamke et al. 2016). Sharma et al. (2019) found that glucose mediated signaling induced the accumulation of hypermethylation (H3K4me3) on genes associated with heat stress memory. DNA hyper-methylation patterns at different genic or regulatory regions have important roles in conferring abiotic stress tolerance. Rajkumar et al. (2019) studied the dynamic methylation patterns of sensitive (Nagina 22) and tolerant (Pokkali) cultivars of rice towards drought and salinity stress. They correlated the hypermethylation and elevated expression of abiotic stress responsive genes in tolerant cultivar Pokkali during both the stresses. Similarly, Yang et al. (2018) found the elevated H3K4me2 methylation in soybean during heat stress. Recently, Liu et al. (2018) showed that the sustained H3K4 methylation during heat stress has been associated with HSFA2 for establishment of stress associated memory in Arabidopsis.

2.7.4 Studies of class A2 HSFs in monocots

HSFs are much explored in model crops tomato and Arabidopsis, both consisting of single *HSFA2* member as compared to cereal crops like rice, wheat, maize and barley. All the monocot cereal crops consist of at least five members in HSFA2 class (A2a to A2e), each having unique

and distinct function. Yokotani et al. (2008) studied the function of rice *OsHSFA2e* by performing its ectopic overexpression in Arabidopsis. They observed that *Arabidopsis* plants were showing thermotolerant phenotype when exposed to heat stress. In addition, on performing microarray analysis authors identified that certain stress responsive genes were regulated by *OsHSFA2e* including heat shock protein genes, given their expression was also found during ambient temperature conditions in transgenic plants. Furthermore, authors observed these transgenic plants tolerant to salinity stress, indicating the importance of *OsHSFA2e* in alleviating different form of environmental stresses (Yokotani et al. 2008). Chauhan et al. (2013) isolated a seed preferential heat shock factor belonging to subclass-A2 (*TaHSFA2d*) from developing seeds of heat stressed wheat plants which was found homologous to rice HSF *OsHSFA2d*. Given the presence of transactivation domain as a characteristic of Class-A HSFs, they identified the transactivation potential of the newly identified HSF by performing protein-protein interaction assay in yeast. Additionally, the overexpression of *TaHSFA2d* in *Arabidopsis* plants was identified to confer tolerance against heat and other abiotic stresses. Moreover, in mutant lines *TaHSFA2d* was found to partially complement the role of *AtHSFA2* during exposure to stress conditions. On performing the qRT-PCR based analysis authors revealed the upregulation of abiotic stress responsive genes in constitutively expressing *TaHSFA2d* Arabidopsis plants under normal conditions.

It is clear from above that most of the overexpression studies have been reported mainly in model dicot plants *Arabidopsis* and tomato and limited for monocots plants. Wang et al. (2017) identified and characterized a heat shock factor *FaHSFA2c* in tall fescue plant which is a temperate grass. They developed transgenics of tall fescue and Arabidopsis plants simultaneously overexpressing *FaHSFA2c*. To examine the subcellular localization of *FaHSFA2c*, a *GFP*::*FaHSFA2c* fusion protein was created and Arabidopsis protoplast cells were transformed. The overlapping GFP fluorescence and DAPI (a nucleus staining dye) designated its localization to nucleus. In Arabidopsis three transgenic lines were selected for physiological analysis on the basis of their better survival rate at 45°C. Wherein, a severe exposure of heat stress for 3 days at 45°C resulted in non-recoverable wilting of wild type *Arabidopsis* plants, whereas transgenic lines showed higher recovery with fully expanded leaves, though, few leaves remained yellowish. Furthermore, overexpressing transgenic lines of both tall fescue and *Arabidopsis* revealed better membrane stability and photosynthetic rates than their respective

wild type plants. Moreover, production of more number of tillers was also observed in transgenic lines during the exposure of severe heat stress in case of tall fescue. The expression pattern of downstream target genes in plants constitutively expressing *FaHSFA2c* was examined using q-RT PCR, where the higher expression of several small and higher molecular weight HSP genes and genes required for redox homeostasis like ascorbate peroxidase was evident (Wang et al. 2017). Interestingly, *AtApx2* gene showed upregulation under the regulation of *FaHSFA2c* in *Arabidopsis*, although *FaApx2* did not show any induced expression during early and prolonged heat stress in tall fescue plants. Therefore, authors inferred that plants expressing *FaHSFA2c* exhibited enhanced basal and acquired thermotolerance and hence it might be suggested as a potential candidate gene to generate thermotolerant crop, especially belonging to *Poaceae* family through genetic engineering or molecular breeding. Furthermore, biotechnological and next generation sequencing-based approaches allowed the comprehensive functional analysis of target genes predicted to be involved in alleviating the impacts of abiotic stresses. For example, Jain et al. (2014) performed the NGS-based whole genome re-sequencing of three contrasting cultivars of rice possessing variable tolerance towards abiotic stress. Authors provided the valuable information about the putative target genes and molecular markers which could be utilized for abiotic stress specific molecular breeding-based crop improvement programs.

2.8 Heat shock proteins (HSPs)

Cellular protein homeostasis is very prone to heat injury. High temperature can distort the original folding and conformations of proteins, causing their aggregation and consequently render them nonfunctional and possibly toxic to cells. Heat shock proteins or molecular chaperones are present ubiquitously from prokaryotes to eukaryotes and help maintaining the protein homeostasis within the cells in normal and stressed conditions. Initially, Tissieres et al. (1974) gave the first clue for the existence of heat shock proteins and their induced expression by giving heat shock treatment in *Drosophila melanogaster*. The role of heat shock protein in providing tolerance against heat stress is widely known for plants as well. Burke et al. (1985) found the accumulation of HSPs in leaves of cotton plants incubated at 40°C for three hours and growing in field conditions experiencing the same degree of temperature for several days. In another study, Kimpel and Key, (1985), observed that during heat stress response the synthesis of heat shock proteins takes place. In etiolated seedlings of soybean, they found the strong

induction of HSPs upon gradual increase (2-3°C/hour) in the temperature of the growth chamber. Analogously, in field conditions also they obtained similar results regarding induced synthesis of HSPs and HS mRNA. The one-dimensional SDS-PAGE profile of in-vitro translated proteins of these mRNAs provided evidence of high and lower weight HSPs synthesis. On the basis of their findings, they proposed that the synthesis of HSPs is a general part of heat shock response in plants. Plants exhibit both basal and acquired thermo-tolerance in response to heat stress. Krishnan et al. (1989) used two contrasting varieties of wheat Mustang (tolerant) and Sturdy (sensitive), showing a difference in providing acquired thermo-tolerance, to analyze the effect of the difference in genotype on HSP production. A significant difference in the production of small HSPs was observed among the two cultivars including sHSP16, sHSP17 and sHSP26. They suggested that the production of some low molecular weight proteins was directly correlated with the ability of plants to tolerate heat stress. These studies evidenced that the production of HSPs takes place in temperate crops including pea, wheat, and soybean, etc. as they encounter a temperature above the optimum level. The involvement of HSP in heat shock response for providing thermo-tolerance is completely well known and also evolutionarily conserved (Vierling, 1991).

Multigene families of HSPs are evolved due to the duplication events during evolution and their members exhibit complex functions. The expression of HSPs in plants is both constitutive and abiotic and biotic stress dependent. Constitutively expressed HSPs assist in proper folding and translocation of newly translated proteins and hence considered as essential component of plant growth and development (Lindquist and Craig, 1988). During stress conditions, these HSPs promote the refolding of protein structures to their native form and aids in their stabilization (Vierling, 1991). According to their molecular weight, HSPs are categorized into the classes viz. small HSPs (sHSPs), HSP60, HSP70, HSP90 and HSP100. All heat shock proteins are comprised of a conserved heat shock domain at their C-terminal, whereas the N-terminal domain is variable for each class (Vierling, 1991 and Helm et al. 1993). These high and low molecular weight HSPs are localized in the cytosol and other subcellular compartments including mitochondria, endoplasmic reticulum, nucleus, chloroplast etc. (Boston et al. 1996). The low molecular weight proteins or sHSPs have molecular weight ranging from ~ 15kDa to 40kDa and possess a characteristic C-terminal α-crystalline domain. The expression of these proteins during different developmental stages and various stress conditions (heat, cold, salinity etc.) has been

reported by Sarkar et al. (2009) in rice. Additionally, the functional characterization of wheat sHSP26 was performed by Chauhan et al. (2012) by applying transgenic approaches. They found that transgenic Arabidopsis plants overexpressing sHSP26 of wheat showed improved thermotolerance by protecting PSII from thermal injury. The transgenic lines showed higher biomass accumulation and seed yields. However, the reverse was observed in case of antisense plants. Through promoter analysis they revealed the involvement the sHSP26 in stress tolerance and developmental processes like seed germination and maturation. Similarly, Reddy et al. (2014) studied the microarray based differential expression pattern of all 20 sHsps of barley. The seed specific expression of most of the Hsps suggested their role in seed development of barley. The involvement of *HvHSFB2c* in regulating the seed specific expression of several sHSPs (HvHsp16.88-CI, HvHsp17.5-CI and HvHsp17.7-CI) was also deduced by authors' by exploiting gene network analysis. They also suggested the role of these sHSPs in ameliorating the effects of adverse environmental conditions.

Therefore, heat shock proteins are the major class of heat and other abiotic stress responsive proteins. They maintain the cellular activity of proteins and prevent their functional damage through cross talk with other cellular components involving in stress response and provide tolerance to plants.

2.9 Hsp70: Structure and functional relevance

During heat and other stress conditions, proteins of HSP70 superfamily were found to be accumulated abundantly in the cell cytosol. Their upregulation was also observed in normal conditions, indicating their role in different developmental processes of plants (Nover and Scharf, 1997). HSP70 proteins work in an ATP-dependent manner; they bind to non-native hydrophobic residues of target proteins and prevent their aggregation.

2.9.1 Structure of HSP70 proteins

Zhu et al. (1996) studied the crystal structure of the Hsp70 protein of *E.coli* known as DnaK. Proteins of this class possess adenosine triphosphate binding domain required for substrate exchange; a substrate binding domain (SBD) which binds to the hydrophobic amino acid residues of proteins exposed in response to stress (Zhu et al. 1996). Authors proposed that the C-terminal SBD consist of both β-sandwich and alpha-helical structures, where alpha-helical

structure help to stabilize the DnaK-substrate complex without forming any direct contact to the polypeptide. Furthermore, their model suggested that substrate binding to the HSPs is dependent on their confirmations. Moreover, Erbse et al. (2004) reviewed that Hsp70 chaperones associate with their target proteins, either a nascent or stressed protein, by forming non-covalent transient associations in the presence of ATPs. This chaperone mechanism is further mediated by the co-chaperone J-domain (~70 amino acid residues) containing proteins. J-domains are exclusively conserved and responsible for the functional specificity of their client proteins (Sahi and Craig, 2007). Moreover, the expression of these co-chaperones could be orchestrated according to cellular need and physiological status of plants (Verma et al. 2017)

2.9.2 Functions of Hsp70 family proteins

HSP superfamily consists of many members and found across the plant kingdom. Duplication and divergence lead to the formation of multiple members of a gene family. For instance, Arabidopsis genome has duplicated three times with extensive rearrangements which led to the formation of different sets of gene families (Proost et al. 2011). Lin et al. (2001) studied the *Hsp70* gene family in Arabidopsis and identified 18 members in this class at genome-wide level by utilizing already known gene sequences of yeast. Out of 18 genes, they found 14 genes belong to DnaK while four genes were form SSE/110 subfamily. Authors further performed the expression analysis of these genes under various stresses, such as thermal, drought and chilling stress, and observed their induced upregulation across entire suite of stresses (Lin et al. 2001). In a similar study, Sung et al. (2001) performed the expression analysis of *Hsp70* genes of *Arabidopsis* under the temperature and chilling stress. They found the induction of all genes in response to the heat stress, whereas only chloroplastic and mitochondrial members showed induced expression under the chilling stress. Authors also suggested that *Hsp70* genes exhibit development-associated roles as they showed induced expression during seed maturation stage (Sung et al. 2001). Hsp70 proteins localized in ER, known as Bip proteins, are essentially required for nascent protein targeting to ER lumen, and are also a part of unfolded protein response in plant cells. To furthermore analyze the role of ER lumen specific Bip proteins during stress, transgenic tobacco plants overexpressing *Bip* gene were developed, where tobacco overexpression lines exhibited enhanced tolerance towards drought stress by maintaining the protein structure and activity of secretory proteins required for osmotic adjustments during

stressed conditions (Alvim et al. 2001). Also, enhanced stomatal conductance and rate of transpiration was also reflected in tolerant lines. Conversely, antisense lines showed deprived photosynthetic activity, lower transpiration rates and enhanced anti-oxidative enzyme synthesis. Authors further inferred that constitutively expressed Bip proteins enhance the tolerance towards oxidative stress, aside from performing the key regulatory roles (Alvim et al. 2001). Likewise, in genome-wide perspective, 25 Hsp70 family genes in rice, a few of which were segmentally duplicated, were observed for expression pattern during the heat stress and further compared with the reported data for drought, cold and salinity stress (Hu et al. 2009). As a result, the overlapping expression of most of the genes were observed across most of the stressed conditions, whereas, some genes showed expression specific for particular stress including heat stress.

Upon identifying the interaction between HSF1 and Hsp70 proteins of Arabidopsis it was suggested that during the later stages of heat shock response, such interactions attenuate the transcriptional activity and act as negative regulators of HSFs (Kim and Schoffl, 2002). Furthermore, these proteins are observed essential for maintaining plant growth and development under several adverse conditions. On the other hand, in transgenic tobacco plants overexpressing cytosolic *NtHsp70-1* were observed more tolerant towards water deficit conditions (Cho and Hong, 2006). The chloroplast localized Hsp70 proteins in model crop *Arabidopsis thaliana* were investigated by constructing the knockout mutants of *ΔcpHsc70-1* and *Δ70-2,* which have showed impaired shoot and root development (Su and Li, 2008). Furthermore, authors observed that seed germination of *ΔcpHsc70-1* mutant lines, in particular was adversely affected by high temperature stress. Interestingly, the double knockout mutants for both of the genes studied were lethal which suggested that these genes are crucial for plant development (Su and Li, 2008). Similarly, Duan et al. (2011) isolated a *TaHsc70* gene from wheat leaves infected by a rust pathogen *Puccinia striiformis* utilizing SSH cDNA library, which was phylogenetically a homolog of chloroplastic Hsp70 protein. *TaHsc70* gene showed enhanced expression when plants were challenged with fungal pathogen and heat stress as well. Furthermore, exogenous application of only methyl jasmonate among enhanced the *TaHsp70* mRNA production, nonetheless, its expression remains uninfluenced when plants were treated with other phytohormones including other jasmonates and ethylene etc. Their findings suggest that

TaHsp70 is involved in both abiotic and biotic defense related responses encompassing JA-mediated signaling.

Similarly, the functional characterization of 32 *Hsp70* genes in rice, mined and identified through rice database was described as vital for plants during elevated temperatures (Sarkar et al. 2013). They further found that yeast strains mutated for Ssc1 protein shows compromised growth at 37°C, which loss in function mutation was successfully complemented by rice mitochondria specific Hsp70 proteins mt70-1, 70-2 and 70-3 at different extents. Similar to other studies, authors found that protein isoforms belonging to same subcellular compartments are not functionally redundant, and instead exhibit the differential expression (Sarkar et al. 2013). Likewise, 61 Hsp70 genes were identified throughout the soybean genome, highest across plant species hitherto, wherein 53 of them interestingly showed differential expression among different tissues particularly during heat and drought stresses (Zhang et al. 2015). Out of them, the promoters of 45 harbor HSE, presence of which mark the involvement of genes in heat stress specific response, while one gene consist of dehydration response element. These results suggest that most of the *GmHsp70* genes are required for providing heat shock response. Liu et al. (2018) found that *Chenopodium quinoa* comprise of 16 *Hsp70* genes. They performed a more comprehensive evolutionary study of Hsp70 gene family among 16 different plant species by constructing a phylogenetic tree utilizing Hsp70 protein sequences. They found that in case of *Chenopodium quinoa* allopolyploidy is responsible for expansion of *Cqhsp70s* gene family. Higher and differential expression patterns of these genes during drought stress revealed that they are potentially involved in tolerance mechanisms and possess functional diversity.

2.10 Hsp90: Structure and functional relevance

2.10.1 Structure of Hsp90 proteins

Hsp90 proteins consist of an N-terminal domain of 25-kDa, which exhibits the ability of nucleotide binding and hydrolysis with a linker connecting to C-terminal domain (55-kDa). Linker is of variable sizes and not necessarily found to present in all Hsp90 proteins, such as mammalian Hsps, whereas, both N and C-terminals are conserved across all species (Pearl and Prodromou, 2000).

2.10.2 Functions of Hsp90 family proteins

Molecular chaperones of Hsp90 class are not the prerequisite for folding of newly formed proteins, although they assist the proteins to maintain their native conformations (Nathan et al. 1997). The most interesting substrates of these proteins are the molecules which are involved in mediating signaling cascade of cell involving as receptors of steroids (e.g. glucocorticoid), some kinases (e.g. tyrosine, serine or threonine kinases) and NO synthase (Young et al. 2001). Further loss of function mutations in Hsp90 impaired the mitogen associated signaling cascade which further halts the process of cell division (Young et al. 2001). Similar to *Hsp70*, *Hsp90* gene family also contains multiple members. *Arabidopsis thaliana* genome possesses 7 HSp90 family genes which exhibit plastid, mitochondria, cytosol, and ER specific localization on the basis of the deduced structure of all these proteins (Krishna and Gloor, 2001). Authors further revealed the presence of both of the conserved N and C-terminal domains, whereas linker domain was found to be present especially among cytosol and ER-localized proteins, similar to animals and yeast Hsp90 proteins. They further described that the expression of Hsp90 genes increased in presence of heat, cold, heavy metal stress and also during the exogenous applications of the brassinosteroids. These proteins localized within different subcellular compartments exhibit unique functions and developmental roles (Krishna and Gloor, 2001). For example, in *Arabidopsis*, Ishiguro et al. (2002) deciphered the function of *SHEPHERD* (AtHsp90-7), an ER residing protein in maintaining the correct folding and functional conformation of CLAVATA proteins which regulates the development of shoot and flower meristem. Similarly, Cao et al. (2003) cloned a CR88 gene in *Arabidopsis* encoding chloroplast specific Hsp90 protein, which remains constitutively expressed under normal conditions, as well as with induced expression during high light and temperature conditions. Moreover, the higher expression of this gene was also observed at different plant developing stages i.e. immediately after germination, in vegetative tissues including leaves, and in reproductive stages (Cao et al. 2003). Intriguingly, given the mutants of this gene showed defective plastid development, chloroplast CR88 chaperone was proposed to exhibit role in proper maturation and translocation of newly synthesized and imported proteins, and vital for chloroplast biogenesis (Cao et al. 2003). Moreover, Sangster et al. (2007) suggested that the inhibition of Hsp90 adversely affects the quantitative traits, governing time of flowering and seed setting. In a separate study, Sangster et al. (2008) further described that altered activity of Hsp90 proteins could modulate the phenotype

of organisms including plants. This may be attributed to the modulated activity of Hsp90 leading to the exposure of cryptic genetic polymorphisms and influencing the canalization of genetic traits. Hence, these proteins are vital for maintaining phenotypic variability and stability (Sangster et al. 2008).

Apart from maintaining the conformation, folding and sorting of their target proteins to the organelles, Hsps also lead to their proteasomal degradation, whose structural damages are beyond the limits of repair. Pratt et al. (2010) described that Hsp90/Hsp70 co-chaperone machinery leads to the proteasomal degradation of proteins damaged due to oxidative stress. Conversely, under non-stress conditions Hsp90 and Hsp70 regulate the activity of HSFs (HSFA1, HSFA2 and HSFB1) in tomato (Hahn et al. 2011). Hsp70 act by inhibiting the DNA-binding activity of HSFA1, however, Hsp90 leads to the degradation of HSFA2 transcripts and regulates it at cellular level. Moreover, HSFB1 act as a client protein of Hsp90 and undergo proteasomal degradation (Hahn et al. 2011).

Furthermore, several reports suggest that Hsp90 molecular chaperones are also involved in cellular signal transduction pathways. Hsp90 and Hsp70 proteins revealed influencing the expression of ABA-induced genes and these chaperones might be the part of signaling cascade initiated by this phytohormone (Clement et al. 2011). Moreover, Shigeta et al. (2014) presented the evidence of AtHsp90-3 interaction with a transcription factor BRI1-EMS-SUPPRESSOR 1 (BES1) which is involved in brassinosteroid (BR) signaling. Upon providing geldanamycin a universal inhibitor of Hsp90 in T87 cultured cells they observed the inhibition of two genes i.e. *constitutive photomorphogenesis* and *dwarfism,* and *dwarf4*, which further suggested that genes involved in biosynthesis of BR are controlled by HSP90/BES1 interactions. Gupta et al. (2015) through whole-genome transcriptional profiling of Arabidopsis suggested that brassinosteroids and glucose co-regulate the expression of several genes conferring tolerance towards various abiotic and biotic stresses.

The role of Hsp90 in imparting tolerance to abiotic stress conditions and regulating plant development is also well documented. The genetic and functional analysis of cytosol-localized three Hsp90 genes (*TaHsp90-1*, *90-2* and *90-3*) in wheat revealed that *TaHsp90-1*, *TaHsp90-2* and *TaHsp90-3* were located on short arms of chromosome 2, 5 and long arm of chromosome 7 respectively (Wang et al. 2011). They further observed the differential expression of these genes

wherein higher number of transcripts of *TaHsp90-1* gene was found in reproductive tissues such as pollen grains while its marginal expression was observed in vegetative tissues especially in roots. Conversely, the higher expression of *TaHsp90-2* and *TaHsp90-3* genes was found in both of the vegetative and reproductive tissues (Wang et al. 2011). Moreover, authors determined the role of one *TaHsp90-1* gene during seedling growth and the other two genes *TaHsp90-2* and *TaHsp90-3* in providing tolerance against rust pathogens by applying virus induced silencing approach (Wang et al. 2011). Keeping in view the importance of Hsp90 family genes, Zhang et al. (2013) performed the analysis of this gene family at genome-wide scale and identified 10 members in *Populus* tree. Authors further revealed their responses towards various abiotic stresses such as heat, drought, nitrogen deficiency and methyl jasmonate treatment on the basis of their expression profiles. With the temperature increase above of the optimum levels, the upregulation of all *Hsp90* genes of *Populus* was noticed (Zhang et al. 2013), while during nitrogen deficiency the *Hsp90* genes depicted differential expression patterns. For example, *PtHsp90-1a* and, *90-1b*, *90-5a*, *90-5b* showed upregulation in 4 weeks and 8 weeks old leaves. However, most of the genes remain downregulated in case of mechanical wounding and methyl jasmonate treatment. Their findings unraveled that *Hsp90* genes are required for abiotic stress tolerance. Furthermore, Agarwal et al. (2016) performed a comprehensive survey of Hsp90 chaperones in four different legume crops namely, chickpea (5 members), pigeon pea (7 members), medicago (5 members), common bean (6 members) and lotus (5 members). In order to observe the expression patterns of these genes during heat stress they utilized the RNA-seq data of different growth stages of three tolerant and susceptible varieties of chickpea and validated it with qRT-PCR. They found that among all six chickpea varieties, *Ca_09578* and *Ca_14133* genes showed significant upregulation. However, *Ca_23799* was only candidate upregulated in all of the three tolerant varieties. At reproductive phase in leaf tissues, *Ca_02170* was upregulated in all six genotypes, whereas, *Ca_08436* and *Ca_23799* genes were highly expressed in only tolerant genotypes. They suggested that results of expression profiles of these genes in tissue and specific manner could provide candidate genes for further studies.

2.11 Hsp100: Structure and functional relevance

2.11.1 Structure of Hsp100 family proteins

Gottesman et al. (1990) initially identified and purified an ATP-dependent caseinolytic protease in *E.coli* and later on these proteins were identified in almost all organisms including plants, except animals (Wawrzynow et al. 1996). These Clp proteins are classified on the basis of presence of nucleotide binding domains (NBDs). The NBD has highly conserved structure and consist of walker A (for nucleotide binding) and walker B (for Mg^{+2} binding) domains (Schirmer et al. 1996). Class I Clp protease possesses two NBDs and distributed in subfamilies from ClpA to ClpD, whereas class II possesses one NBD in subfamilies ClpM, N, X and Y (Schirmer et al. 1996).

2.11.2 Functions of Hsp100 family proteins

Clp ATPase or Hsp100 proteins are a specialized class of molecular chaperones having ability of protein disaggregation along with performing conventional chaperone activity. The proteins belonging to this family are also widely studied in plants including Arabidopsis, rice, maize, wheat, soybean etc. (Agarwal et al. 2001, Singh et al. 2010). Like Hsp70 and Hsp90 chaperones, Hsp100 proteins also aid plants to survive in stressful conditions. Singla and Grover, (1993) recognized a protein of 110-kDa in rice seedling and specifically in shoot tissues utilizing antibodies raised for yeast HSP104 protein. They proposed it as a heat and ABA-inducible protein. Schirmer et al. (1994) isolated an *AtHsp101* gene form Arabidopsis and found it orthologous to yeast *HSP104* gene. The accumulation of AtHsp101 protein was heat inducible and not observed in ambient conditions. Heterologous expression of this gene complemented the deficiency of endogenous *HSP104* in mutated yeast strains. Similarly, Lee et al. (1994) isolated a cDNA clone of *Hsp101* from heat induced soybean seedlings, which further restored the compromised acquired thermo-tolerance of mutated yeast deficient for *Hsp104*. Nieto-Sotelo et al. (1999) characterized a newly isolated gene of Hsp100/ClpB family in maize (Hsp101), which was present as a single copy throughout its genome. *cis*-regulatory region of this gene was enriched with five copies of heat shock elements which are the characteristics of heat shock response genes. Although authors found heat induced higher expression of maize *Hsp101*, however, did not observe any transcripts of this gene during cold and osmotic stress and upon exogenous treatment with ABA. Queitsch et al. (2000) did the functional analysis of *Hsp101* (ClpB1) by its overexpression and suppression in transgenic Arabidopsis plants. The lines showing inhibited expression of *HSP101* were severely affected by heat stress. The expression of

Hsp101 was required during seed germination as a developmental regulator, hence silencing of this gene decreased their tolerance and affected the process of germination. Although, the overexpressing lines were performing better while exposed to elevated temperature. These changes in *Hsp101* expression further changed the basal and acquired thermo-tolerance (AT) status of plants.

Keeler et al. (2000) cloned a chloroplast targeted *HSP100/ClpB* gene of lima bean. The previous reports suggested that these proteins were essential for providing acquired thermo-tolerance to yeast and cyanobacteria. Nevertheless, authors were not able to conclude its role in imparting AT, but due to constitutive expression of this gene in field grown plants and overexpression in laboratory conditions they suggested its heat stress specific function. In another study, Hong and Vierling, (2000) screened genetic loci *hot1-4* in Arabidopsis and found that mutation at *hot1* locus cause complete loss in function of *Hsp101* gene. Plants with *hot1* mutation were not able to tolerate high temperature stress i.e. having compromised basal and acquired thermo-tolerance, which suggested that *Hsp101* was essential for plant survival during heat shock. Aside from providing thermo-tolerance, Hsp100 proteins aids in alleviating other abiotic stresses. Campbell et al. (2001) isolated two *Hsp100* genes in bread wheat namely, *TaHsp101B* and *TaHsp101C*. Both of the genes showed heat stress dependent differential expression. Interestingly, they found the higher transcript levels of these genes in response to drought stress and ABA treatment. It was the first study which reported the expression of Hsp100 genes in presence of both heat and dehydration stress. Katiyar-Agarwal et al. (2003) developed the stable transgenic plants of rice cultivar Pusa Basmati1 (*Oryza sativa L.*) overexpressing *AtHsp101* gene. They performed the analysis of heat shock on T_2 generation plants. Under normal conditions, both control and transformed plants were in healthy condition, however, transgenic plants showed remarkable tolerance and recovery during and after heat shock than that of wild type plants. They further observed that the phenotype of improved thermo-tolerance was solely imparted by overexpression of *AtHsp101*. Moreover, Agarwal et al. (2003) also reported that rice *OsHsp101* encoding protein promotes the disaggregation of misfolded protein granules and required for improved protein homeostasis.

Hsp100 molecular chaperons regulate the developmental process of plants along with maintaining the cellular protein homeostasis. The Arabidopsis ClpB subfamily consists of four

genes i.e. *ClpB1*, *B2*, *B3* and *B4*. Lee et al. (2007) studied the role of three *ClpB* genes viz. *ClpB2*, *B3*, and *B4*. They used GFP proteins fused with the transit peptides of the ClpB3 and ClpB4 proteins and found their respective localization in chloroplast and mitochondria; annotated them as ClpB-p and ClpB-m, respectively. Heat stress specific accumulation of the transcripts in all of the three genes namely *AtHsp101*, *ClpB-p* and *ClpB-c* were observed, nonetheless, mutants for *ClpB-p* and *ClpB-m* did not show any seed specific phenotypes. Intriguingly, mutation of *ClpB-p* was lethal due to failing in chloroplast biogenesis in mutant plants. Therefore, these proteins are also essential for maintaining the housekeeping functions of cells and not only limited to stress tolerance. Yang et al. (2006) identified and studied a chloroplast localized *Lehsp100/ClpB* gene in tomato. Unlike lima bean they found higher and induced expression of this gene only when tomato plants were treated with heat and not in optimal conditions. To further validate its function, they generated the antisense lines for this gene by *Agrobacterium* mediated transformation. When control wild type plants along with antisense lines were initially exposed to mild heat stress (38°C for 2 hours) and then to severe heat shock at 42°C for similar duration, the later did not survived and died after few days. This may be due to nonreversible damage of PSII and halt of photosynthesis process. Consequently, they concluded the involvement of *Lehsp100/ClpB* in imparting AT plants.

A comprehensive study of Hsp100 family at genomic level in rice was conducted by Singh et al. (2010). They identified nine ORFs in Hsp100 family of rice, localized to different cellular organelles. The phylogenetic analysis of rice Hsp100 proteins with other dicot plants i.e. *Arabidopsis* and *Populus* showed the conservation of ClpB family proteins among monocots and dicots. Microarray analysis of these genes revealed their expression specifically during heat stress followed by cold, desiccation and ABA treatment. Hsp100 genes are majorly heat inducible and consist of heat shock elements in their promoter regions. Singh et al. (2012) developed the transgenics of rice with two different gene constructs having 2kb upstream region of ClpB-cyt/Hsp100 and promoter with deletion at HSE region (-273 to -280) with Gus as a reporter gene. The tissue specific induced expression of reporter gene was observed in vegetative and reproductive tissues in response to the high temperature and metal stress. Surprisingly, plants with mutated copy of HSE showed temperature induced as well as constitutive expression of reporter gene. This suggests that the presence of classical HSE is required for the attenuation of ClpB-cyt/Hsp100 gene activity in absence of heat shock stress. Authors further performed the

DNA-protein interaction assay and found the interaction of *OsHsfA2c* with promoter elements of *OsClpB-cyt* and *OsHsfB4b* with all the three isoforms (Singh et al. 2012).

2.12 Metabolomics of temperature stress and other abiotic stresses in plants

The different phenotypes of plants are the result of interactions between gene expression, protein synthesis and metabolite synthesis at a particular time and conditions. Plants respond to heat stressed conditions by altering their transcriptome, proteome and metabolome. Numerous alterations in transcriptome and metabolome of plants take place in response to acquire tolerance against several abiotic and biotic stresses. For example, in an interesting study Panikulangara et al. (2004) identified that HSF *AtHsfA1b* overexpressing Arabidopsis plants showed constitutive expression of *galactinol synthase 1* (*GolS1*) gene, which regulated the synthesis of oligosaccharides of raffinose family. Expression analysis of wild type plants revealed the heat induced expression of *GolS1*. Additionally, the functional studies of this gene unveiled its role in heat tolerance by regulating the sugar metabolism in plants. Furthermore, to identify the novel compounds and examining differential changes in metabolome during acquired heat and cold stress tolerance, Kaplan et al. (2004) performed the metabolite profiling of Arabidopsis using GC-MS. They found a very rapid increase in salicylic acid levels during both heat and cold stress, and implicated its role in heat stress signaling and acquired thermo and chilling tolerance. The other metabolites which were increased during heat and cold stress were some carbohydrates (e.g. sucrose and raffinose) and amino acid molecules such as tysrosine, valine, phenylalanine and tryptophan etc. Their findings suggested that most of the signaling molecules and metabolites accumulated during both stressed conditions were common and collectively mediating the stress tolerance mechanism. Several compatible solutes like glycine betaine also help in providing osmoprotection during abiotic stress (Giri, 2011). Interestingly, different plant species respond specifically to same environmental challenges. For example, the ectopic expression of *PpENA1* gene of *Physcomitrella patens* encoding a Na^+ transporter, in rice and barley plants provides salinity tolerance. The metabolite analysis of leaf tissue of both the rice and barley transgenic plants showed differential metabolite profile when exposed to salinity stress (Jacobs et al. 2007). Moreover, proper growth and productivity of plants depend on the availability of essential macro and micronutrient. The metabolite analysis revealed the modification in carbohydrate and nitrogen metabolism during Pi deficiency in barley plants

(Huang et al. 2008). In addition, Avila-Ospina et al. (2017) studied change in metabolic profiles of leaves during senescence in barley under low and high nitrate conditions, and found decrease in intermediates of glycolysis pathway, TCA cycle, and other carbohydrate molecules such as hexose and sucrose. They suggested that sugars play key roles during leaf senescence.

Usually plants experience both heat and drought stress simultaneously in field conditions. Tolerance to heat and drought is a multigenic complex trait which governs the physiological, biochemical, molecular and developmental process of plants. Several quantitative trait loci (QTLs) have been identified at genomic level which regulates the developmental process of plants during stress conditions. Similarly, Templer et al. (2017) identified several metabolic QTLs in barley utilizing landraces of Mediterranean region and elite breeding cultivars of Central Europe. A total of 81 accessions were used for the identification of mQTLs, and metabolic traits providing adaptations to tolerant cultivars. Fifty-seven metabolites were identified which were involved in antioxidant, carbon and nitrogen metabolism in flag leaves under stressed and non- stressed conditions. The mQTL traits identified were the compounds involved in ROS detoxification viz. tocopherol and glutathione. The metabolic status of plants during stress conditions provide the information about pathways and primary and secondary molecules involved in providing tolerance to plants. Therefore, metabolomics of plants generates a numerous data and provide novel insights to comprehend the effect of complex metabolic changes on providing tolerance to plants against various heat, salinity, drought and other abiotic stresses.

Materials and Methods

Materials and methods

3.1 Materials

The experiments were conducted for comprehensive study of *Hsp70*, *Hsp90* and *Hsp100* heat shock proteins families at genome wide level by utilizing comparative genomic and expression analysis approaches. Furthermore, the functional characterization of *HvHSFA2c* gene was performed to understand the heat stress tolerance mechanism in barley through its overexpression, followed by transcriptomic, physiological and biochemical analysis of transgenic plants. Additionally, the analysis of change in metabolome of barley plants during normal and heat stressed conditions was performed.

a.) Plant material

Seeds of barley genotype Golden Promise and RD2786 were procured from University of Zurich, Switzerland and Indian Institute of Wheat and Barley Research, Karnal India respectively.

b.) Glasswares, Plasticwares and Equipments

Glass wares used in the present investigation were purchased from Borosil (India), and Schott Duran, (Germany). Plastic wares comprising of 1.5 mL and 2 mL microcentrifuge tubes, 0.2 mL PCR tubes with domed cap, 90 mm sterile petri-plates, micro tips and tip boxes were procured from Tarsons, India. Micropipettes were obtained from Eppendorf (Germany). The instruments used in the present study are listed in Table 3.1.

Table 3.1 List of instruments

S. No.	Name of the Instrument	Company/ Model
1	Laminar Air Flow	Diamond clean air devices, India
2	PCR machine	Bio Rad, USA
3	Refrigerated centrifuge	Eppendorf, Germany
4	Microcentifuge Wise Spin	Witeg, Germany
5	Water bath	Rivotek, India
6	UV-Visible Spectrophotometer	Agilent Technologies, USA

7	Microwave oven	Samsung, India
8	Magnetic Stirrer	Tarsons, India
9	Refrigerator	Whirlpool, India
	Deep Freezer (-80°C)	New Brunswick, USA
10	Deep Freezer (-20°C)	Blue Star, India
11	pH meter	Labman, India
12	Digital balance	G%G, USA
13	Gel Documentation System	Bio Rad, USA
15	Agarose Gel Electrophoresis system	Tarsons, India
16	Power pack of Agarose Gel Electrophoresis system	Tarsons, India
17	Nanodrop	Thermo Scientific
18	Real Time PCR	Quantstudio 3, Applied Biosystems, USA
19	Sonicator	Scientz, China
20	Biolistic Particle Delivery System	Bio Rad, USA
21	Tissue LyzerII	Qiagen, Germany

Table 3.2 List of chemicals

Chemicals/Fine Chemicals/ Enzymes	Company
EDTA, SDS, Sodium acetate, Glycerol, DTT, Tris base, NaCl, Potassium dihydrogen orthophosphate, Potassium acetate, Guanidine hydrochloride, Calcium chloride, LB-Agar, LB-Broth, Spectinomycin, Ampicillin, Agarose, Bromophenol blue, Ethidium bromide, Boric acid	HiMedia Laboratories Pvt. Ltd.
Isoamyl alcohol, Chloroform, Bromophenol blue, Isopropanol, Absolute Ethanol, Potassium phosphate dibasic, EDTA disodium	SRL, Pvt. India Ltd.

salt, Triton X-100, X-Gluc, Nitroblue tetrazolium chloride (NBT), Methionine, L-Proline, L-Glutamine, Thiamine HCl, Casein hydrolysate, Gellan Gum, Maltose, Glucose, L-Asparagine, Myo-inositol, Potassium ferrocyanide, Potassium ferricyanide, DMSO, Sarcosyl, MES	
MS minerals with vitamins, 6-BAP, Dicamba, 2-4-D	Duchefa
Hygromycin-B	Thermo Scientific
Spermidine	Sigma
Gold particles	Bio Rad
Oligonucleotide Primers	Eurofins and Sigma
Phusion, Trizol, One Step SuperscriptII RT-PCR, Clone Jet Cloning Kit pJET	Thermofisher, Scientific
Gel elution and PCR Product Purification Kit, Plasmid DNA isolation kit	Favorgen Biotech Corps
Restriction enzymes	New England Biolabs

3.2 Experimental Plan Outline

a.) *In silico* and expression analysis of all three gene families:

i) Identification, phylogenetic analysis, analysis of gene structure, protein domains, and regulatory domains in promoter regions were performed.

ii) Seed germination and seedling growth of barley under controlled conditions in green house.

iii) Various abiotic stresses and chemical treatments given at different growth stages.

iv) RNA isolation from different tissues and cDNA synthesis.

v) Primer designing and quantitative RT-PCR analysis for all the three family genes was done.

b.) Cloning of *HvHSFA2c* gene and transformation for barley plants for its overexpression:

i.) *HvHSFA2c* gene was cloned in PCR vector and then in binary vector.

ii.) Transformation of immature embryos of cultivar Golden Promise through particle bombardment.

c.) Analysis of putative transgenic plants.

d.) Transcriptome analysis of overexpressing lines:

i.) RNA isolation of overexpressing lines and wild type plants. Outsourcing of RNA for RNA-seq experiment.

ii.) Data analysis

iii.) Biochemical analysis

e.) Comparative metabolomic analysis of plants:

i.) Isolation of metabolites for GC-MS analysis

ii.) Data analysis

3.3 Methodology

3.3.1 *In silico* analysis of *Hsp70, Hsp90* and *Hsp100* family genes at genome wide level

a.) Identification and phylogenetic analysis of *Hsp70, Hsp90* and *Hsp100* family genes in barley

Homology search with reported *Hsp70, Hsp90* and *Hsp100* family genes of rice was done in order to identify members of these families in barley. Total 32 genes in *Hsp70* family and 9 genes in each *Hsp90* and *Hsp100* family were reported in rice (Singh et al. 2010; Hu et al. 2009 and Sarkar et al. 2013), and sequences of all these genes were retrieved from TIGR database (http://rice.plantbiology.msu.edu/). Nucleotide sequences of each gene were BLAST searched against *H. vulgare* using BLASTX in 'NCBI' and then confirmed by TBLASTN. The identified Hsp family gene sequences were again deep searched in 'IPK database' (http://webblast.ipk-gatersleben.de/barley_ibsc/) (using BLASTN with "Barley all transcripts HC" and "Barley pseudomolecule contigs masked Apr2016") and 'gramene database' (http://ensembl.gramene.org/Hordeum_vulgare/Tools/Blast?db=core).

After removal of redundant sequences, domains in these gene family proteins were searched using SMART search tool (http://smart.embl-heidelberg.de/smart) for confirming the presence of specific Hsp70, Hsp90 and Hsp100 domains. These identified high confidence protein sequences were subjected to ExPASy (http://web.expasy.org/protparam/) for the calculation of number of amino acid present, analysis of molecular weight and determination of theoretical isoelectric point. Phylogenetic analysis of these newly identified genes of *Hsp100*, *Hsp90* and *Hsp70* families along with their rice homologs was done in Mega 7.0 software (Kumar et al. 2016), ClustalW was selected for multiple sequence alignment and a bootstrap N-J tree was constructed using pairwise gap deletion mode. WoLFPSORT and TargetP 1.1 (http://www.cbs.dtu.dk/services/TargetP/), ChloroP 1.1 databases were accessed for in silico analysis of their subcellular localization (Emanuelsson et al. 1999). The nomenclature for all the identified members of Hsp70, Hsp90 and Hsp100 in barley, were followed according to their rice orthologs, as proposed by Sarkar et al. 2013, Hu et al. 2009, and Singh et al. 2010 respectively.

b.) Gene structures of *Hsp70*, *Hsp90* and *Hsp100* gene families

To study the intron, exons and UTR distribution, GFF files were downloaded from Ensemble ftp server (ftp://ftp.ensemblgenomes.org/pub/plants/current/fasta/hordeum_vulgare/). Predicted representative gene structure were mined from this file and converted into bed files for respective genes. The bed file was used on the GSDS version 2.0 (http://gsds.cbi.pku.edu.cn/) server to draw the gene diagrams as well as their arrangements.

c.) Promoter *cis* element distribution analysis of *Hsp70*, *Hsp90* and *Hsp100* gene families

Genome of *Hordeum vulgare* is downloaded from the Ensemble ftp (https://www.ensembl.org/info/data/ftp/index). The coordinates of the genes were used to get the 1Kb upstream sequences of these genes. To fetch the sequences, a samtools wrapper (Li et al. 2009) was written in PERL. Those 1K upstream sequences which fall in gap regions were discarded from analysis. Further, those with smaller gaps, ambiguous nucleotides were removed from them. A *cis*-element database file was downloaded from PlaceDB and was used to query the upstream regions of various genes. Over-represented motifs and functionally related *cis* elements were used to create a GFF file and same was used to generate the one Kb upstream diagrams.

d.) Protein domains analysis of *Hsp70, Hsp90* and *Hsp100* gene families

Pfam database from the ftp server in form of PfamA.hmm (https://pfam.xfam.org/) was downloaded. This database was used to query against the identified members of representative protein sequences. To query the protein sequences of HSPs representative proteins sequences, hmm scan from HMMER 3.1b2 version was used with default cut off.

3.3.2 Expression analysis of genes belonging to all the three families under study

a.) Plant material and treatments

Barley variety RD2786 was grown in controlled conditions with 20°C day/18°C night temperature and 16 hours light. Fourteen days old seedlings were used for expression analysis. Heat stress treatments for 30 minutes and 2 hours were given to seedlings in growth chamber at 42°C temperature, followed by recovery from heat stress at control conditions and sampling was done after two days of recovery. For drought stress seedlings were left for drying for 2 hours in light. Other stresses were given in aqueous solutions for four hours including sodium chloride (200 mM), cadmium chloride (100 µM), calcium chloride (100 µM), salicylic acid (100 µM) and brassinosteroid (1 µM). Unfertilized flower and developing seeds at different developmental stages i.e. 3 DAP (days after pollination), 7 DAP, 14 DAP, 21 DAP and 28 DAP were subjected to heat stress for four hours. Sampling of all control tissues including shoot, root, flower and at developing stages along with tissues in stressed conditions was done. All the samples were snap-freezed in liquid nitrogen until RNA isolation.

b.) RNA isolation and cDNA synthesis

RNA isolation of two-week-old seedling was done by using SV total RNA isolation system (Promega, USA) according to manufacturer's instruction, including an on-column DNase treatment. RNA from developing seeds was isolated by using methods described by Singh et al. (2003) for carbohydrate rich seeds followed by RNA purification with SV total RNA isolation system (Promega, USA). The detailed procedures for RNA isolation and purification from seeds are as follows.

Table 3.3 Composition of RNA extraction buffer for seeds (Singh et al. 2003)

S. No.	Components	Final Concentration	Amount for 200 mL
1.	Tris-HCl	50 mM	10 mL
2.	Sodium chloride	150 mM	15 mL
3.	Sarcosyl	1%	20 mL
4.	EDTA	20 mM	4 mL
5.	DTT	5 mM	10 µL

Table 3.4 Composition of Guanidine HCl buffer (Singh et al. 2003)

S. No.	Chemical	Final Concentration	Amount for 50 mL
1.	Guanidine hydrochloride	8 M	32.8 g
2.	EDTA	20 mM	1 mL
3.	MES	20 mM	4 mL
4.	β-mercaptoethanol (β-ME)	200 mM	10 µL

Other components required for RNA isolation are:

1. Phenyl, chloroform and isoamyl alcohol in 25:24:1 ratio.
2. 3M Sodium acetate pH 5.2.
3. 2M Sodium chloride
4. Isopropanol
5. 70% ethanol

c.) Protocol for RNA isolation from barley seeds

1. Seed samples were finely ground in liquid nitrogen using mortar-pestle.
2. Approximately, 100 mg of finely ground sample was transferred to 2 mL microcentrifuge tubes (MCTs) containing 500 µL of seed extraction buffer, 5 µL of DTT and 500 µl of Phenol: Chloroform: Isoamyl alcohol was added in a ratio 25:24:1 and vortexed vigorously.
3. The samples were centrifuged at 21910 x g for 5 minutes at 4°C.

4. Supernatant was removed carefully in another 2 mL MCT. 650 µL of guanidine buffer and 350 µL of P:C:I were added to the sample and vortexed for mixing. After vortexing 10 µL of β-mercaptoethanol was added. Then the samples were again centrifuged at 21910 x g for 5 minutes at 4°C.
5. Approximately, 450 µL of upper aqueous phase was transferred to two new1.5 mL MCTs. 45 µL of 3 M sodium acetate of pH 5.2 and 900 µL of prechilled absolute ethanol was added and mixed by inverting 4-5 times. Samples were then incubated at -80°C for three to four hours.
6. After incubation, samples were taken out and centrifuged at 21,910 x g at 4°C for 20 minutes.
7. Supernatant was discarded and resulting pellet was washed by adding 70% chilled ethanol by spinning at 10,130 x g at 4°C for 3 minutes.
8. Again, supernatant was discarded, and pellet was put for air drying at room temperature.
9. Pellet was dissolved in 30 µL autoclaved DEPC treated water and stored at -80°C until further use.

d.) Purification of RNA

RNA was purified using SV total RNA isolation system Promega, USA.

1. 200 µL of RNA was taken in a fresh micro centrifuge tube and added with 200 µL of RNA lysis buffer containing β-ME.
2. 350 µL of RNA dilution buffer was added and mixed by inverting 3-4 times.
3. 300 µL of 95% ethanol was added and mixed well by pipetting.
4. The mixture was transferred to column supplied with the kit and centrifuged for 1 minute at 12,000 x g.
5. 600 µL of RNA wash solution was added and centrifuged for 1 minute at 12,000 x. Elute was discarded from the collection tube.
6. DNase incubation mix was prepared as given in Table 3.5. 50 µL of DNase mix was added to column membrane and left for incubation at room temperature for 15 minutes.
7. To stop the reaction, 200 µL of DNase stop solution was added immediately and centrifuged for 1 minute at 12, 000 x g.

8. 600 µL RNA wash solution was added and centrifuged for 1 minute at 12,000 x g. The flow through was discarded.
9. 250 µL of RNA wash solution was added and centrifuged for 2 minutes at 18,000 x g.
10. Column was transferred to new 1.5 mL MCT. 50 µL of nuclease free water was added to the column and centrifuged for 1 minute at 12,000 x g to elute the RNA.
11. The eluted RNA was stored at -80°C until further use.

Table 3.5 Composition of DNase incubation mix*

S. No.	Components	Amount for 50 µL
1.	Yellow core buffer	40 µL
2.	$MnCl_2$	5 µL
3.	DNase I	5µL

*Note all the components are supplied with the SV total RNA isolation system kit (Promega, USA).

e.) Agarose Gel Electrophoresis

Agarose gel electrophoresis method was used to check the quality of RNA fragments. For resolving the RNA fragments on the gel, the RNA samples were loaded with RNA loading dye in 4:2 ratios on 1% agarose-formaldehyde gel. The RNA was subjected to electrophoresis in 1x MOPS (3-(N-morpholino)propanesulfonic acid) buffer. RNA concentrations were measured using Nanodrop (Thermo scientific, USA).

Table 3.6 Composition of MOPS buffer

S. No.	Components	Concentration	Amount in 1000 mL
1.	MOPS free acid	0.2 M	41.8 g
2.	Sodium acetate anhydrous	0.05 M	4.1 g
3.	Na_2EDTA	0.01 M	10 mL
			pH 7.0

Table 3.7 Composition of RNA loading dye

S. No.	Components	Amount in 1 mL
1.	Formamide	600 µL
2.	Formaldehyde	100 µL
3.	MOPS	125 µL
4.	Ethidium Bromide	50 µL
5.	Sodium acetate	125 µL
6.	Glycerol	50 µL

f.) Components of cDNA synthesis

First strand cDNA was synthesized with 1 µg of total RNA using Go-Script reverse transcription system (Promega, USA) as per manufacturer's instructions (Table 3.8 and Fig. 3.1).

Table 3.8 Components for cDNA synthesis

S. No.	Components	Volume for 20 µL
1.	RNA (1 µg)	10 µL
2.	5x Buffer	4.0 µL
3.	$MgCl_2$	2.0 µL
4.	dNTPs	1.0 µL
5.	Random Hexamers	1.0 µL
7.	MMLV-RT	0.5 µL
8.	Milli Q	1.5 µL

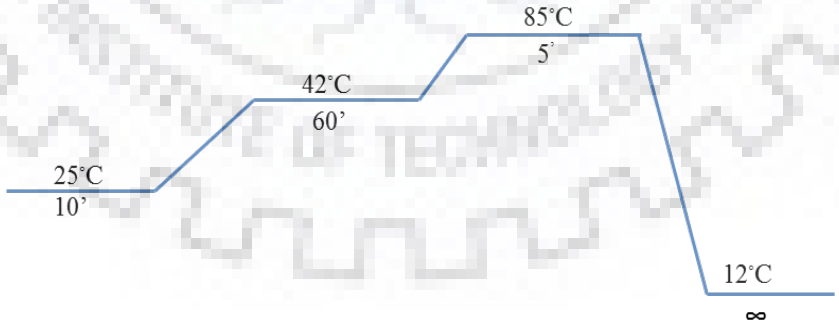

Fig. 3.1 First strand cDNA amplification cycle.

g.) Primer designing and expression analysis of *HSP* genes of barley

Primer express 3.0 software (Thermo scientific, USA) was employed to design primers for qPCR by utilizing coding DNA sequences (CDS) of all identified *HSP* family genes under study. (Table 3.9).

Table 3.9 List of primers of all *Hsp70*, *Hsp90* and *Hsp100* genes

S. No.	Name of Hsps	Primer- Forward	Primer- Reverse
1.	*HvHsp100-1*	*TCGACGACCTGAGGGACAA*	*TGCTTCAGCTTGCGGATCT*
2.	*HvHsp100-2*	*CCGACCGGTGTGGGTAAA*	*GATGCGGACGAGCAGGTT*
3.	*HvHsp100-3*	*CAAGACCGAGCTGGGAAAGA*	*CGTATCAAGGCGTTCTCAGTGT*
4.	*HvHsp100-4*	*GCAGGCCGAGCGTGAGTAT*	*CCGCTGCAATGCATTCAG*
5.	*HvHsp100-5*	*TCCGGGAGCGGTACGA*	*CGGCAATCAGCGCTTCA*
6.	*HvHsp100-6*	*GGTTCGACGGCGACCATA*	*TCCGGATGCGCTTTTTCA*
7.	*HvHsp100-7*	*AAGGGTTGGCTCTTCGGATT*	*CGCTTTGCCACGAGGAA*
8.	*HvHsp100-8*	*TCCCTGACAGGCATCTTCCT*	*GCTCTGCTACCGGCCTCAT*
9.	*HvHsp90-1*	*ACAAGAACGACAAGTCCGTCAA*	*GAGCATGCGGTGGATCCT*
10.	*HvHsp90-2*	*CCTTGCCGTCAAGCACTTC*	*AAGAGGACGGCCTTGAACTCA*
11.	*HvHsp90-4*	*CCTTCGACCTCTTCGACAACA*	*AAGACACGGCGCACATACAG*
12.	*HvHsp90-5*	*TCAATCCCGAACACCCTATTG*	*GGTGCTCTCAGGCTCGTTCT*
13.	*HvHsp90-6*	*AGGTTAATGAGGGCGCAGTCT*	*TCCTGGACCGCATGAAATC*
14.	*HvHsp90-7*	*CCAACATGGAGAAGATCATGCA*	*GCATGTACGCCTGCTTGCT*
15.	*HvHsp70-2*	*GCATTTGCAACCCCATCATC*	*GCCACCGGGCATTTCA*
16.	*HvHsp70-4*	*GCAACGCGCTGGAGAACTA*	*GGAGCTTGGACGCGATCTT*
17.	*HvHsp70-5*	*TCCTGTGAGCGGGCAAAG*	*CAAACAGCGAGTCAATCTCGAT*

18.	*HvHsp70-5a*	*TGGACCTCTTCCGCAAGTG*	*GCACGGTGCTCTTGTCCAT*
19.	*HvHsp70-5b*	*GACCTGTTCCGCAAGTGCAT*	*GCACGGTGCTCTTGTCCAT*
20.	*HvHsp70-N1*	*GCTTGTGCCCAAGATAAATGTG*	*CGCCGACACCTTCAGGAT*
21.	*HvHsp70-N2*	*ATCCCCACAACGAGGAAGAA*	*CCAAGTTCATGGAGGTCTGGTT*
22.	*HvHspmt70-2*	*GCAAGGCGAGCGTGAGAT*	*TGCCGACCAGGTCAAACTC*
23.	*HvHspmt70-3*	*CCACCATCTACAGCATCGAGAA*	*CGACTTCAGCAGGAACCTTGTC*
24.	*HvHspcp70-2*	*CAGTGCAAGAAACCGTTAGGAA*	*TCATCAGGGTTGACGGTAACATT*
25.	*Hv-Bip1*	*CGACGCCCGGAACCA*	*GTCGCCCACCGTGTTCTT*
26.	*Hv-Bip3*	*CGAAGCTGATCCCGAGGAA*	*GGTGGTCTGCTTGTCCTGGTA*
27.	*Hv-Bip6*	*AACGGCGACACCCATCTC*	*CGGATGAAGTGGTCCATTACG*
28.	*HvHsp110-2*	*AACCTCCAAAGGAAGACGCTAA*	*GGCTCGCTGGATGCAGAT*
29.	*HvHsp110-3*	*GCCATCGCTTCAGGAGTCA*	*CCTGGATCGTCAGCAGATCA*
30.	*HvHsp110-4*	*CCCGAGGATGCGCTCTT*	*GGAACTCCCGCACTTTGAAC*
31.	*HvHsp110-7*	*TGCTGGAAGCTCAAGAGCAA*	*TTTCGTTCGCTCGACAAGTTT*

Quantitative real time PCR analysis was performed in Quant studio 3 system and by using PowerUp™ SYBR™ Green Master Mix in 96 well plates (Applied Biosystems, Thermofisher scientific, USA). Barley *Actin* gene (Gene bank accession no. AK356840.1) was used as internal standard for normalization and fold change was calculated using $2^{\wedge-\Delta\Delta Ct}$. Heat-maps were used to depict the qPCR results obtained after the experiments. To plot the heatmaps, fold change values obtained following ddCt method were used. The fold change values less than 2 were not taken into consideration and are shown using 'azure' color blocks. Data transformation was done by taking the log values of fold changes. These log transformed values were used to plot heat-maps. Heat-maps were generated in 'R' version 3.2.3 and its packages 'gplots' and 'RColorbrewer'.

3.3.3 Cloning of *HvHSFA2c* gene

a.) Seed germination and heat stress treatment

Seeds of cultivar Golden Promise were sown in germination trays containing mixture of soil and soilrite (1:1) under controlled conditions in greenhouse 18/16°C (day/night) temperature and 16-hour photoperiod. Two-weeks-old plantlets were subjected to heat stress at 42°C for different durations i.e. for 15, 30, 45, 60 and 120 minutes in a growth chamber. The leaf samples were harvested at every time point and snap freezed in liquid nitrogen.

b.) Cloning of *HvHSFA2c* gene

Leaf samples collected at different time points were pooled and RNA isolation was done by using Trizol reagent (Invitrogen, Thermofisher USA).

c.) RNA isolation from leaf tissues by Trizol method

1. Leaf samples were homogenized using Tissue lyser II (Qiagen, Germany) at frequency of 30 Hz for 30 seconds.
2. One ml Trizol reagent was added to 100 mg of homogenized samples and incubated for 5 minutes at room temperature for proper dissociation of nucleoprotein complexes.
3. 200 µL of chloroform was added to each sample and shaken vigorously for 2 minutes followed by incubation at RT for 5 minutes. Followed by centrifugation at 12,000 x g for 15 minutes at 4°C.
4. Upper aqueous phase was transferred to fresh 1.5 mL MCT and 500 µL chilled isopropanol was added for precipitation of RNA and incubated at -80°C for 1 hour.
5. To pellet down RNA samples were centrifuged at 12000 g for 10 minutes at 4°C.
6. After careful removal of supernatant, pellet was washed with 75% 1 mL ethanol, and centrifuged at 7500 g for 5 minutes at 4°C.
7. Ethanol was removed and pellet was left for air drying.
8. Dried pellet was re-suspended in 30 µL RNase free DEPC treated water and stored at -80°C till further use.
9. Quality and integrity of RNA was checked on 1.2% agarose-formaldehyde gel in MOPS buffer (Table 3.6).

10. To estimate the concentration of RNA, quantification was done by measuring absorbance ratio A260/A280 using Nanodrop (Thermo Scientific, USA).

d.) cDNA synthesis and amplification of *HvHSFA2c* gene

2µg RNA and *HvHSFA2c* forward and reverse primers were taken for cDNA synthesis and PCR amplification of 1119 bp ORF of *HvHSFA2c* gene was performed by using one step SuperscriptII RT-PCR system (Invitrogen, Life Technologies, Rockville, MD, USA) according to manufacturer's instructions. The resulting PCR product was then cloned in pGEMT-easy cloning vector (Promega, USA).

e.) Sub-cloning of *HvHSFA2c* in pJET1.2 PCR cloning vector

PCR was performed using *pGEM-T:HvHSFA2c* plasmid DNA as template, *XmaIHvHSFA2c-F* and *SalIHvHSFA2c-R* primers and Phusion™ DNA polymerase enzyme (Thermofisher Scientific, USA) to amplify *HvHSFA2c* gene of 1119 bp. The composition of PCR mixture and PCR cycle is given in Table 3.10 and Fig. 3.2 respectively. The amplified PCR product was run on 1% agarose gel and TBE buffer (Table 3.11, Fig. 3.3a).

Table 3.10 Composition of PCR reaction mixture for amplification of *HvHSFA2c* gene

S. No.	Components	Volume for 20 µL
1.	Phusion™ enzyme mix	10.0 µL
2.	*XmaIHvHSFA2c-F*	1.0 µL
3.	*SalIHvHSFA2c-R*	1.0 µL
4.	Plasmid DNA (*pGEM-T-A2c*)	1.0 µL (25 ng/µL)
5.	DMSO	1.0 µL
6.	Milli Q	6.0 µL

Table 3.11 Composition of Tris borate EDTA (TBE) buffer

S. No.	Components	Amount for 1000 mL
1.	Tris base	108 g
2.	Boric acid	55 g
3.	EDTA	40 ml (0.5 M) pH=8.0

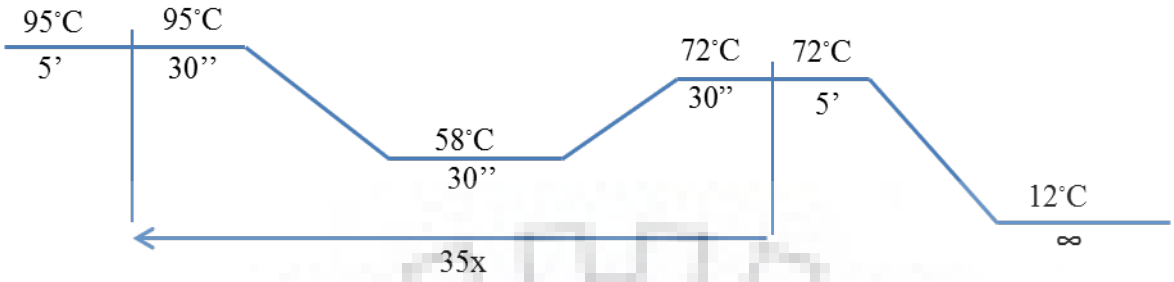

Fig. 3.2 PCR amplification cycle of barley *HvHSFA2c* gene.

a) b)

Fig. 3.3 PCR for *HvHSFA2c* amplification and vector map of PCR cloning vector. (a) Agarose gel (1%) showing the amplified product of *HvHSFA2c* gene with band size 1119 bp. (b) Vector map of PCR cloning vector pJET 2.1. The blunt end PCR product (*XmaI-SalIHvHSFA2c*) was ligated between *XhoI* and *XbaI* blunt sites present in MCS region of vector. (Thermofisher Scientific).

Then the desired band was excised and eluted using Favorgen Gel/PCR purification kit (Favorgen Biotech. Corp.) as per manufacturer's instructions. Ligation reaction of eluted amplicon *XmaI-SalIHvHSFA2c* as insert with pJET1.2 vector (Fig. 3.3b) (Invitrogen, Thermofisher Scientific USA) was put according to the manufacturer's instructions at 25°C for 1 hour and 16°C overnight. The composition of ligation mixture is given in Table 3.12. The transformation of DH5α cells of *E.coli* was done with the ligated product.

Table 3.12 Ligation reaction of *XmaISalIHvHSFA2c* in pJET1.2 cloning vector

S. No.	Components	Volume (10 µL)
1.	Reaction buffer	5.0 µL
2.	pJET1.2 vector	1.0 µL
3.	Insert DNA	3.0 µL
4.	Ligase	1.0 µL

vi) Transformation of DH5α cells

1. *E.coli* DH5α competent cells were taken out from -80°C in ice and 10 µL of ligated product was added to the cells.
2. Cells with ligated product were kept on ice for 20-30 minutes.
3. For transformation, three freeze thaw cycles of 45 seconds at ice and 42°C (in water bath) was repeated with final incubation for 5 minutes at ice.
4. 800 µL of LB broth was added to these transformed cells which were further put in incubator shaker at 37°C at 200 rpm for 1 hour.
5. After one-hour cells were pelleted down at 5000 rpm for 5 minutes, approximately 600 µl of supernatant was discarded. The cells were resuspended in remaining 200 µL of LB by gentle pipetting and spreaded on LB agar media plates containing ampicillin as selection pressure.
6. Plates were put in an incubator at 37°C for overnight.

For plasmid isolation, growth of five transformed colonies was put on LB-broth containing ampicillin antibiotic at 37°C for overnight.

vii) Plasmid isolation from *E. coli* DH5α cells using alkaline lysis method

1. One to 2 mL of full-strength bacterial culture grown overnight was transferred to 2 mL micro centrifuge tubes and pelleted down at 13,500 rpm for 1 minute.
2. Supernatant was discarded and each pellet was resuspended in 200 µL of Lysis I buffer (Table 3.13).
3. 250 µL of freshly prepared Lysis II buffer (Table 3.14) was added to the resuspended pellet, mixed gently by inversions and incubated at room temperature for 5 minutes.

4. After completion of incubation 350 µL of Lysis III (Table 3.15) buffer was added to neutralize the solution and mixed by gentle inversions. Followed by centrifugation at 13,000 rpm for 10 minutes at room temperature.
5. Supernatant was taken to new 1.5 mL MCT and 0-7 x volume of isopropanol was added for precipitation of plasmid DNA. Incubation was done at -20°C for 10-30 minutes.
6. To pellet the precipitated DNA, centrifugation was done at 13,000 rpm for 15 minutes.
7. Pelleted plasmid DNA was washed with 500 µL of 70% ethanol by centrifugation at 13,000 rpm for 10 minutes.
8. Pellet was allowed to air dry and dissolved in 25-50 µL of TE.

Table 3.13 Composition of Lysis buffer I

S. No.	Components	Final Concentration	Amount for 100 mL
1.	Tris- base	25 mM	2.5 ml
2.	EDTA	10 mM	1 ml
3.	Glucose		0.9 g

Table 3.14 Composition of Lysis buffer II

S. No.	Components	Final Concentration	Amount for 5 mL
1.	NaOH	0.2 N	500 µl
2.	Sodium dodecyl sulfate	1 %	250 µl

Table 3.15 Composition of Lysis buffer III

S. No.	Components	Final Concentration	Amount for 100 mL
1.	Potassium acetate	3 M	29.4 g

Cloning was confirmed by restriction digestion using *XmaI* and *SalI* (New England Biolabs, USA) enzymes, followed by sequencing. The reaction mixture for restriction digestion is described in Table 3.16.

Table 3.16 Composition of restriction digestion put to check for insert ligation in pJET1.2 vector

S. No.	Components	Volume for 10 µL
1.	Plasmid DNA	2.0 µL
2.	*SalI* readymix	2.0 µL
3.	*XmaI* HF	0.5 µL
4.	Cut smart buffer	2.0 µL
5.	Milli Q	13.5 µL

viii) Subcloning of *XmaI-SalIHvHSFA2c* in binary vector *p6oAct-UbiZm-LH*

In binary vector *p6oAct-UbiZm-LH* (*p6OAct* hereafter) (Fig. 3.4), barley HSF *XmaI-SalIHvHSFA2c* was subcloned under ubiquitin promoter (*ZmUbi*) and *nos* terminator (*NosT*). Double digestion of empty binary vector and PCR vector pJET1.2 harboring gene of interest (*XmaI-SalIHvHSFA2c*) was performed with both the cloning enzymes *XmaI* and *SalI*. Digested products were run on 1% agarose gel. The bands of insert and vector were excised and extracted using Favorgen Gel/PCR purification kit (Favorgen Biotech. Corp.).

Ligation of vector (*XmaI-SalI* digested *p6OAct*) and insert (*XmaI-SalIHvHSFA2c*) was performed in 3:1 ratio as described previously. Further steps from transformation to plasmid isolation were similar as described in above section only bacterial selection was different i.e. spectinomycin. The cloning was confirmed by restriction digestion with *EcoRV* enzyme. The growth was again put for positive clones. Plasmid isolation was done by using kit (Favorgen Biotech. Corp.), and samples were sent for sequencing. The complete construct formed in *p6OAct* vector is shown in Fig. 3.5a.

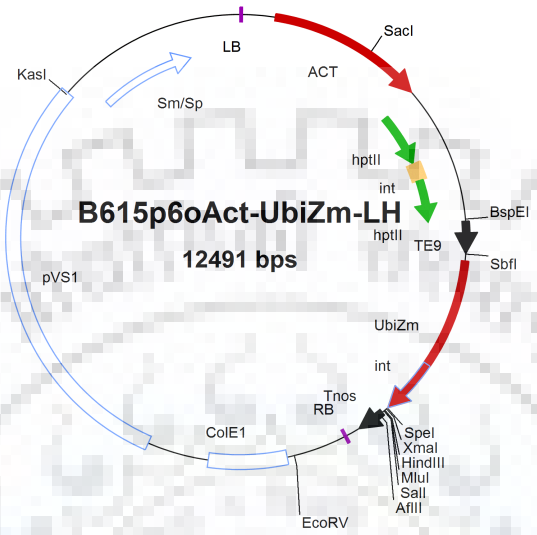

Fig. 3.4 Map of binary vector *B615p6oAct-UbiZm-LH* (Source: DNA cloning services, Hamburg Germany)

i) Sub-cloning of *XmaI-SalIHvHSFA2c* in modified binary vector *pANIC6B*

For sub-cloning of *XmaI-SalIHvHSFA2c* into *pANIC6B*, the complete cassette containing promoter:geneofinterest:terminator (*ZmUbi:HvHSFA2c:NosT*) was isolated from *p6OAct* vector through restriction digestion with *SwaI* and *PmeI* enzymes. *SwaI* and *EcoRV* sites are present in upstream of *ZmUbi* promoter and downstream of *NosT* terminator respectively. Therefore, the restriction double digestion with both the enzymes provides a complete cassette required for sub-cloning (Fig. 3.6a). Simultaneously the restriction digestion of modified *pANIC6B* vector was done with *SwaI* enzyme which linearized the vector downstream to *NosT* of Gus (Fig. 3.5b). Digested product of both the vectors was run on 1% agarose gel. The desired bands of insert and vector backbone was cut, and eluted and ligation reaction was put. The complete cassette cloned in *pANIC6B* is represented in Fig. 3.5c. Cloning was confirmed by restriction digestion with *AflII* enzyme followed by sequencing.

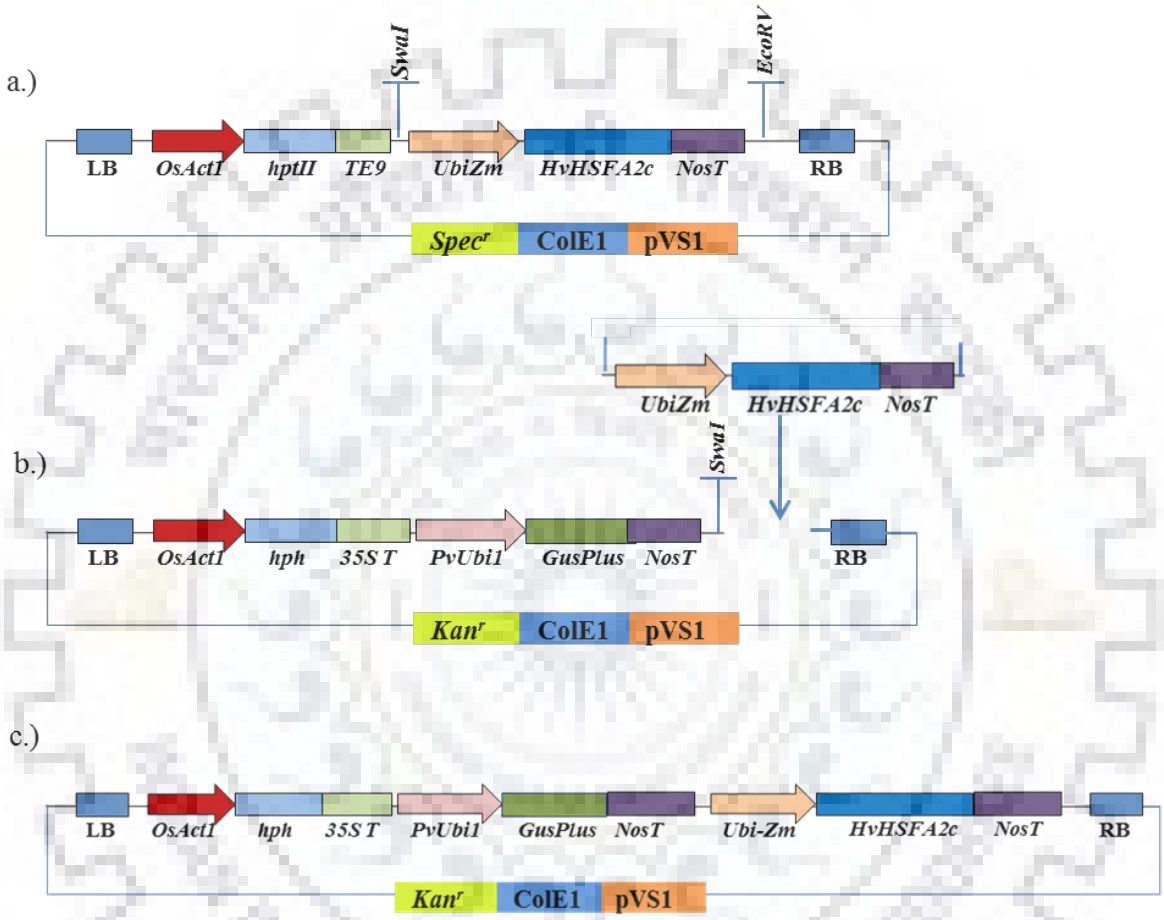

Fig. 3.5 Schematic diagrams of vectors depicting the subcloning strategy of *HvHSFA2c* gene in modified *pANIC6B* vector. a) *p6OAct* vector containing *HvHSFA2c* gene under *ZmUbi* promoter and *NosT* terminator. The restriction sites for blunt end producing enzymes *SwaI* and *EcoRV* enzymes are present in upstream of *ZmUbi* and downstream of *NosT*. b) Modified *pANIC6B* vector digested with *SwaI* enzyme in downstream of *Gus-NosT* which created a blunt end and allowed the ligation of *ZmUbi:HvHSFA2c:NosT* cassette in upstream of right boarder sequence. c) *ZmUbi:HvHSFA2c:NosT* cassette ligated with modified *pANIC6B* vector.

j) Transformation of *Agrobacterium* strain *AGL1* with binary vector *pANIC6B:HvHSFA2c*

1. An aliquot of competent cells of *AGL1* was thawed on ice. 10 µL of plasmid DNA (*pANIC6B:HvHSFA2c*) was added and kept on ice for 15 minutes.

2. Then a single freeze thaw cycle of 5 minutes at liquid nitrogen and at 37°C in water bath was performed.
3. 800 µL of YEM broth (Table 3.17) was added to the cells in laminar air flow and put on incubator shaker for overnight at 28°C.
4. Then the cells were pellet down at 5000 rpm for 10 minutes.
5. 600 µL of YEM was discarded and cells were resuspended in 200 µL of YEM.
6. Spreading was done on YEM-agar plates containing carbenicillin, rifampicin and kanamycin.
7. Plates were incubated for two days at 28°C for growth of transformed colonies.

Afterwards, growth of four transformed colonies was put in 10 mL YEM containing rifampicin, carbenicillin and kanamycin at 28°C for plasmid isolation.

Table 3.17 Composition of YEM (Yeast extract mannitol) medium

S. No.	Components	Amount for 1000 mL
1.	Yeast Extract	1 g
2.	Mannitol	10 g
3.	$MgSO_4$	200 mg
4.	NaCl	100 mg
5.	K_2HPO_4	500 mg

*For YEM agar 15 g/L agar was added to the liquid YEM media.

k.) Isolation of plasmid from *Agrobacterium AGL1*

1. 4 mL culture of *Agrobacterium* was pelleted down the in 2 mL MCT at 13,000 rpm for 1 minute by following the step twice.
2. 250 µL of Lysis I buffer (Table 3.13) was added and cells were resuspended followed by the addition of 20 µL of lysozyme. The samples were incubated for 20 minutes at room temperature.
3. 250 µL of Lysis II buffer (Table 3.14) was added and further incubated for 5 minutes at RT.
4. 350 µL of Lysis III buffer (Table 3.15) was added for neutralization and tubes were inverted for 10-12 times.

5. Samples were centrifuged at 13,500 rpm for 15 minutes at 4°C.
6. Supernatant was taken in fresh 2 mL MCT and equal amount of P:C:I was added and mixed properly.
7. The samples were centrifuged at 12,000 rpm for 10 minutes at 4°C.
8. Again the supernatant was taken in fresh MCT and added with equal amount of isopropanol followed by incubation at -20°C for 1 hour.
9. Samples were centrifuged at 12,000 rpm for 10 minutes at 4°C.
10. Supernatant was discarded and pellet was washed with 70% ethanol and air dried.
11. The pellet was dissolved in 30 µL of TE.

The isolated plasmid was back transformed in DH5α cells. Again, the plasmid was isolated and restriction digestion was put with *AflII* restriction enzyme which confirmed that *AGL1* consisted of the desired plasmid.

3.3.4 Genetic transformation of barley cultivar Golden Promise using particle bombardment and *Agrobacterium* mediated transformation method (Hensel et al. 2009)

a.) Compositions of different media required for barley tissue culture

For barley tissue culture different media were used for osmotic treatment, callus induction, regeneration and plantlet formation. The compositions of all the media used are given in the following tables:

Table 3.18 Composition of osmoticum media

S. No.	Components	Stock concentrations	Amount for 1000 mL
1.	MS minerals		4.3 g
2.	Thiamine Hydrochloride	40 mg/mL	0.4 mL
3.	L-Asparagine		150 mg
4.	Maltose monohydrate		15.0 g
5.	2-4 D	2.5 mg/mL	1 mL
6.	Phytagel		3.5 g
			pH 5.8

Table 3.19 Composition of barley co-cultivation media (BCCM) (Hensel et al. 2009)

S. No.	Composition	Stock concentration	Amount for 1000 mL
1.	MS minerals		4.3 g
2.	Thiamine HCl	1 mg/mL	1 mL
3.	L-Proline		0.69 g
4.	Dicamba	2.5 mg/mL	1 mL
5.	Casein hydrolysate		1 g
6.	Acetosyringone	1M	500 µL
7.	Myo-inositol		0.25 g
8.	Maltose		30 g
9.	Phytagel		3 g`
			pH 5.8

Table 3.20 Composition of barley callus induction media (BCIM) (Hensel et al. 2009)

S. No.	Components	Stock concentration	Amount for 1000 mL
1.	MS minerals		4.3 g
2.	Thiamine HCl	1 mg/mL	1 mL
3.	L-Proline		0.69 g
4.	Dicamba	2.5 mg/mL	1 mL
5.	Casein hydrolysate		1 g
6.	Copper sulphate	2.5 mg/mL	624 µL
7.	Myo-inositol		0.25 g
8.	Maltose		30 g
9.	Hygromycin	50 mg/mL	1 mL
10.	Phytagel		3 g
			pH 5.8

Table 3.21 Composition of barley regeneration media (BRM) (Hensel et al. 2009)

S. No.	Components	Stock Concentration	Amount for 1000 mL
1.	MS minerals		2.15 g
2.	Glutamine		146 mg/L
3.	6-BAP	2.5 mg/mL	90 µL
4.	Maltose		36 g
5.	Hygromycin	50 mg/L	500 µL
6.	Phytagel		3 g
			pH 5.8

Table 3.22 Composition of barley plantlet formation media

S. No.	Components	Stock concentration	Amount for 1000 mL
1.	MS minerals		2.15
2.	Glutamine		146 mg/L
3.	NAA	1 mg/mL	100 µL
4.	Maltose		20 g
5.	Hygromycin	50 mg/mL	500 µL
6.	Phytagel		3 g
			pH 5.8

b.) Growth of donor plants of barley cultivar Golden Promise for tissue culture

Sowing of barley cultivar Golden Promise was done in field conditions and pots during the month of October and November 2018 respectively, at Indian Institute of Technology Roorkee. To synchronize the availability of explants sowing was done at the interval of two weeks. In field conditions, it was difficult to get immature embryos after mid of March. Therefore, to get the continuous supply of explants for longer durations, the plants grown in pots were shifted to greenhouse under controlled conditions at 18/16°C temperature and photoperiod of 16 hours with light intensity 136 µmol/m^2/s photon flux density. Plants were regularly watered according to requirement and fertilizer was given every fortnightly.

c.) Selection, sterilization, and excision of explants

Immature caryopses from donor plants of cultivar Golden Promise were harvested, after 14 days of pollination, when immature embryos started turning translucent to opaque. Immature seeds were put in an Erlenmeyer flask after removal of awns and sterilized with 70% ethanol and washed thrice with autoclaved water. These seeds were further placed in 4% sodium hypochlorite for 15 minutes with continuous shaking at 100 rpm. Washing with sterilized distilled water was done for complete removal of sodium hypochlorite in a laminar airflow.

For excision of immature embryo, seeds were held firmly with forceps. Seed coat was slightly removed from the opposite side of awn using scalpel to expose the immature embryo. The embryo axes were excised using a fine blade followed by removal of immature embryo. These isolated immature embryos were cultivated on barley callus induction media for one day scutellum side up and then transferred to small petriplates (size) containing osmoticum at least five to six hours prior to bombardment (Cardinal et al. 2016). Whereas for *Agrobacterium* mediated transformation immature embryos were directly isolated in six well plates containing barley co-culture medium (Hensel et al. 2009).

d.) Preparation of gene cassettes for particle bombardment

For genetic transformation of barley, cassettes containing *HvHSFA2c* gene, selectable marker and reporter genes were used for co-bombardment of immature embryos of cultivar Golden Promise. Cassette containing *ZmUbi:HvHSFA2c*:*NosT* was isolated through restriction digestion of vector *p6OAct:HvHSFA2c* with *SwaI* and *EcoRV* enzymes (Fig. 3.6a). Additionally, selection and reporter cassette *OsAct1:hph:35T:PvUbi1:GusPlus:NosT* was isolated from modified *pANIC6B* by performing restriction digestion with *EcoNI* and *XbaI* enzymes (Fig. 3.6b)

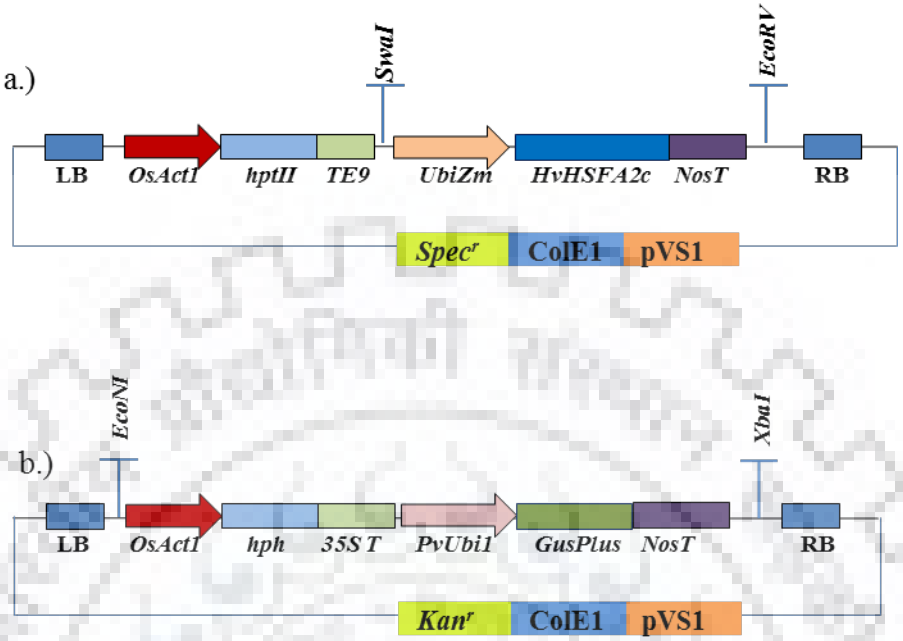

Fig. 3.6 Schematic representation of *p6OAct:HvHSFA2c* and modified empty *pANIC6B* vector. a.) Restriction enzymes *SwaI* and *EcoRV* were used for isolation of *ZmUbi:HvHSFA2c:NosT* cassette from *p6OAct:HvHSFA2c* b.) Restriction enzymes *EcoNI* and *XbaI* were used for isolation of *OsAct1:hph:35T:PvUbi1:GusPlus:NosT* form *pANIC6B* vector.

e.) Preparation of gold particles

Gold particles supplied by Bio-Rad were activated prior to use by using following method

1. 50 mg of 0.6 µm gold particles were resuspended in 1 mL of 70% ethanol by pipetting 10-12 times. These particles were vortexed for 5 minutes, followed by centrifugation at 13,000 rpm for 1 minute.
2. The supernatant was removed carefully. The gold particles were completely resuspended in 1 mL of H_2O.
3. These resuspended gold particles were then vortexed for 1 minute and sonicated for 5 minutes.
4. Centrifugation was done at 13,000 rpm for 1 minute and supernatant was carefully thrown away. The pelleted gold particles were resuspended in 1 mL of sterile H_2O. This step was repeated twice.

5. After completion of step 4, the supernatant was discarded. The remaining gold particle pellets were dissolved in 50% glycerol and stored at -20°C until use.

f.) DNA precipitation on activated gold particles

1. For each bombardment, gold particles were taken out from freezer and resuspended through sonication.
2. According to number of bombardments 5 µL gold particles per shot was taken out from the stock. The dephosphorylated gene cassettes consisting *ZmUbi:HvHsfA2c:NosT* and *OsAct1:hph:35T:PvUbi1:GusPlus:NosT* were added in equal amounts and mixed by pipetting several times.
3. 20 µL spermidine of 0.1 M concentration was added the 50 µL 2.5 M $CaCl_2$. This 70 µL $CaCl_2$ and spermidine mix was added to the gold particles prepared in step 2.
4. The samples were incubated for 1 minute at ice, then vortexed for 15 minutes and centrifuged at 6000 rpm for 1 minute at RT.
5. After removal of supernatant the pellet of gold particles precipitated with DNA was resuspended using 200 µL of 99.8% ethanol, by pipetting up and down several times.
6. Samples from step 5 were vortexed and then centrifuged for 30 seconds at 6000 rpm.
7. Again, supernatant was discarded. The pellet was dissolved in 99.8% chilled ethanol and kept at ice till the bombardment.

g.) Conditions for bombardment

1. 50 immature embryos were put on each osmoticum (Table 3.18) plate approximately 5-6 hours prior to bombardment.
2. Stopping screen and macrocarrier holders were sterilized by autoclaving prior to bombardment. Gene gun, rupture discs and macrocarriers were sterilized with 70% ethanol and isopropanol respectively.
3. Two blank shootings were done before bombardment of explants.
4. Depending on the number of shootings, 10 µL gold particles precipitated with DNA were spreaded on each macrocarrier and left for air drying.
5. After air drying macrocarriers were loaded on macro carrier holders.
6. Rupture disks of 1100 psi were used. A vacuum of 27 Hg was applied for bombardment.

7. The distance between rupture disc and macrocarrier support was 1.3 mm.
8. Petriplates containing immature embryos were put at distance of 9 cm.
9. After bombardment each plate was sealed with parafilm and stored at 4°C for overnight.
10. Next day bombarded embryos were transferred to petridishes containing barley callus induction media (BCIM) without any selection for three days and thereafter on BCIM with 50 mg/L hygromycin as selection pressure for four weeks.

h.) *Agrobacterium* mediated genetic transformation of barley cultivar Golden Promise

1. 50-60 embryos were isolated per well in a six well plate containing 2.5 mL of barley co-cultivation media (BCCM) (Table 3.19).
2. For co-cultivation BCCM was removed and 600 µL of full-strength secondary growth of *Agrobacterium* strain *AGL1* harboring binary vector *pANIC6B:HvHSFA2c* was added to each well.
3. Vacuum infiltration was done for 1 minute at 500 mbar and kept for 10 minutes without agitation in a desiccator.
4. Washing was done with 2.5 mL of BCCM for 10 minutes. For co-cultivation 2.5 mL of fresh BCCM was added to each well.
5. Six well plate was then sealed with parafilm and kept in dark for 72 hours at 21°C, without any agitation.
6. After three days of co-cultivation embryos were cultivated in petridishes containing barley callus induction media (Table 3.20).

i) Callus induction of transformed immature embryos, their regeneration and plantlet formation

1. Bombarded and agro-transformed immature embryos were cultured in dark at 20°C on BCIM (Table 3.20) containing 50 mg/L hygromycin for four weeks, and subcultured at the interval of two weeks on fresh plates containing same media.
2. After four weeks of callus induction, only healthy calli appearing yellowish and compact were considered as transformed and transferred on regeneration media containing 25 mg/L hygromycin (Table 3.21). Whereas, brownish and poorly developed calli were considered as non-transformed and discarded.

3. Transformed calli plated on regeneration media were initially put in indirect light for two days and then transferred under optimum light conditions (136 µmol/s/m^2 photon flux density) at 20°C. Greening of somatic embryos started appearing after two days of transfer in direct light.
4. Subculture on BRM was done every fortnightly till the emergence of regenerants.
5. When plantlets reached length of 2 to 3 cm, they were transferred in individual tubes (Borosil) containing 5 mL of plantlet formation media (Table 3.22) with hygromycin.
6. Plants with well-developed root systems were transferred to pots containing mixture of soil and soilrite (1:1). These plants were put in green house under controlled conditions as described above for the growth of donor plants.

3.3.5 Analysis of transformants and transgenic lines

a.) Histochemical Gus assay

To confirm the stable transformation event, the intrinsic activity of β-glucuronidase (Gus) was monitored by performing histochemical localization. Tissue samples were taken at different stages i.e. immature embryos two days after bombardment, two-week-old calli, six-week-old regenerating calli and plantlet emerging from eight-week-old regenerating calli. For Gus assay, the following procedure was followed:

1. Immature embryos and two-week-old calli (3-4 from each BCIM plate) were put in a 2 mL MCTs containing 1 mL of Gus buffer (Table 3.23), whereas, regenerating plantlet were put in 5 mL MCTs containing 2.5 mL Gus buffer.
2. Vacuum infiltration was done for 15 minutes.
3. Samples were stored in dark at 37°C for overnight.
4. After staining, the samples were rinsed in water. 70% ethanol was added to the regenerating plantlets for removal of chlorophyll.
5. Photographs were taken by using stereo zoom microscope and Sonyα-DSLR camera.

Table 3.23 Composition of Gus buffer

S. No.	Components	Stock concentration	Amount for 50 mL
1.	Sodium phosphate buffer		25 mL
2.	Potassium ferrocyanide	0.1 M	500 µL
3.	Potassium ferricyanide	0.1 M	500 µL
4.	Triton-X-100	10%	250 µL
5.	X-gluc		2.5 mL
6.	Milli Q		21.25 mL

b.) Molecular analysis of transgenic plants

The putative transgenic plants raised through particle bombardment and tissue culture were examined for integrity and overexpression of the gene of interest (*HvHSFA2c*). To examine the integrity of *HvHSFA2c* gene in barely genome, whole genomic DNA from leaf tissue of transgenic plants was isolated according to the protocol described below.

c.) DNA isolation of barley leaf tissue using CTAB method

1. Leaf tissue was taken into 2 mL MCT with stainless steel beads and pre chilled in liquid nitrogen. The homogenization of these leaf tissues was done by using Tissue lyser II at frequency of 28 Hz for 30 seconds.
2. After complete homogenization 1 mL CTAB extraction buffer was added to each tube and samples were put on water bath at 65°C for 45 minutes.
3. When the samples were cooled down to room temperature 500 µL of chloroform: isoamyl alcohol (24:1) was added to the samples and mixed for 5 minutes by continuous shaking and centrifuged at 12000 x g for 15 minutes.
4. A clear layer of supernatant was pipetted out into fresh 1.5 mL MCTs. 2 µL of RNaseA was added to each sample and kept at 25°C for 15 minutes.
5. To precipitate DNA 700 µL of chilled isopropanol was added to each sample, thereafter, stored at -20°C for 2 hours.
6. After completion of incubation, these samples were then again centrifuged at 12000 x g for 10 minutes to pellet down the DNA.

7. Supernatant was discarded, for washing of pellet 500 µL of 70% ethanol was added to each MCT and centrifuged at 7500 x g for 5 minutes.
8. Again, the supernatant was discarded, and DNA pellet was allowed to air dry till the complete removal of ethanol. 70 to 100 µL TE of pH8.0 was added to each sample to dissolve the DNA pellet.
9. Quality of isolated DNA was checked on 0.8% agarose gel prepared in TBE buffer (Table 3.11). Quantification of DNA concentration was done using Nanodrop (Thermo Scientific, USA) by measuring the absorbance ratio of A260/A280.

The PCR amplification reaction was performed for each candidate plant using gene specific primers of *HvHSFA2c* and selectable marker gene encoding *HptII*, enlisted in Table 3.24. PCR reactions of 20 µL were put using GoTaq master mix (Promega, USA) and 50 ng/µL of DNA for each reaction (Table 3.25). The annealing temperature and extension time were set according to T_m (melting temperature) and length of amplification product (Fig. 3.7). The amplified PCR products were run on 1% agarose gel and visualized on gel documentation system.

Table 3.24 List of primers used for the analysis of transgenic plants

S. No.	Primer	Sequence 5'-3'	T_m ˚C	Amplicon length (bp)
1.	*XmaIHvA2c-F*	*CCCGGGATGGACGCGGCGGTG*	58	1119
2.	*SalIHvA2c-R*	*GTCGACCTAGAGCGGGCTAGTAGAACC*	58	1119
3.	*HvA2c500-F*	*CATCTTGCCTGGAGGTTTGGTGAG*	58	619
7.	*6B-ACTHPT-F*	*AGGGGAAAAGGGCACTATGG*	60	450
8.	*6B-ACTHPT-R*	*TTCGGTTTCAGGCAGGTCTT*	60	450
9.	*pANICHpt-F*	*ATGAAAAAGCCTGAACTCACCG*	60	1000
10.	*pANICHpt-R*	*CTATTTCTTTTGCCCTCGGACG*	60	1000
11.	*NosT-F*	*AATAGGATCCGATCGTTCAAACATTTGGC*	60	253
12.	*NosT-R*	*GCGGCCGCGATCTAGTAACATAGATG*	60	253
13.	*qHvA2c-F*	*GATTCGGGTGTCCTCAGTGATC*	*60*	
14.	*qHvA2c-R*	*CTGAATGTTCTGCGCCAAATT*	*60*	

15.	*qHvActin-F*	*CCAAGAACAGCTCCTCAGTTGA*	*60*	
16.	*qHvActin-R*	*AATCGTGATCACCTGACCATCA*	*60*	

Table 3.25 Components of PCR reaction

S. No.	Components	Volume for 20 µL
1.	Milli Q	6.0 µL
2.	DMSO	1.0 µL
3.	Primer-F	0.5 µL
4.	Primer-R	0.5 µL
5.	GoTaq	10 µL
6.	DNA	2 µL

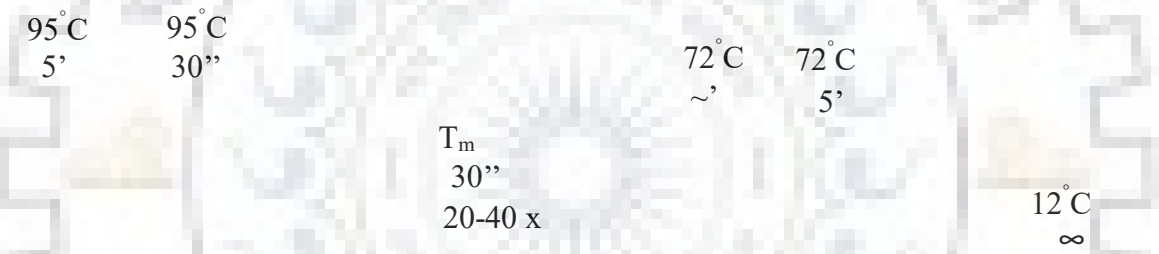

Fig. 3.7 PCR amplification cycle for analysis of putative transgenic plants. Annealing is done at Tm for each primer pair used for analysis and accordingly on the basis of length of amplicon the duration of extension (~) was set.

To identify the overexpression lines, total RNA isolation from leaf tissues of all the T0 plants was done. The quality and quantity of RNA was checked as described previously. cDNA was synthesized with 1 µg/µL of RNA concentration using MMLV-RT enzyme (Promega, USA) Table 3.26.

The plants confirmed positive for transgene through PCR were further analyzed for overexpression. Semi-quantitative PCRs were performed using full length cloning primers (*XmaIHvHSFA2c-F* and *SalIHvHSFA2c-R*) and internal primer pairs (*HvA2c500-F* and *SalIHvHSFA2c-R*) listed in Table 3.24.

Table 3.26 Components for cDNA synthesis

S. No.	Components	Volume for 20 µL
1.	RNA (1 µg)	10 µL
2.	5x Buffer	4.0 µL
3.	$MgCl_2$	2.0 µL
4.	dNTPs	1.0 µL
5.	Oligo Dt	1.0 µL
6.	*SalIHvHSFA2c-R*	0.5 µL
7.	MMLV-RT	0.5 µL
8.	Milli Q	1.0 µL

Standard PCR reactions were put (Table 3.25 and Fig. 3.7) followed by agarose gel electrophoresis and gel documentation. The overexpressing lines showed higher band intensity as compared to other lines. For further confirmation quantitative PCRs analysis were performed using *Actin* gene of barley as internal standard and relative fold change in expression was calculated according to $2^{-\Delta\Delta Ct}$ method.

In similar way, the analysis of T1 plants generated from confirmed T0 transgenic lines was performed. Two overexpressing T1 transgenic lines selected on the basis of qRT-PCR analysis along with two wild type plants were further analyzed at transcriptome level through RNA-seq.

3.3.6 RNA-seq and data analysis of overexpressing transgenic lines

a.) RNA isolation of overexpressing transgenic lines

RNA isolation from leaf tissues of five-week-old non-transformed wild type and overexpressing transgenic lines (two replicates of wild type and two transgenic lines) were isolated as per manufacturer's instruction (Qiagen, Germany). The quality and quantity of RNA was checked as described in previous sections. High quality RNA samples with 200 ng/µL concentration were outsourced for RNA-seq.

b.) Sample preparation and cDNA synthesis

1 µg of total RNA was taken for extraction and purification of mRNA, using oligo-dT beads (TruSeq RNA Sample Preparation Kit, Illumina). The integrity and purity of extracted mRNA was checked on Bioanalyzer 2100 system (Agilent Technologies, CA, USA). RNA with RNA integrity number (RIN) value above 7.0 was used for construction of cDNA library. Fragmentation of purified mRNA was performed at 90°C in the presence of divalent cations. First strand cDNA synthesis was done using random hexamers and Superscript II Reverse Transcriptase system (Life Technologies, USA). Second strand cDNA synthesis was performed by using first strand as template, DNA polymerase I and RNaseH enzymes. Moreover, the cleaning of cDNAs was done using Beckman Coulter Agencourt Ampure XP SPRI beads.

c.) Library preparation for transcriptome sequencing

cDNA libraries for sequencing were constructed using TruSeq RNA Library Prep Kits for Illumina® (NEB, USA) according to the manufacturer's instructions, followed by indexing of each sample. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using the TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. Finally, 100 bp paired end sequencing was carried out at Illumina Hi-Seq platform with a read depth of ≥25 p million reads per library.

d.) Quality check and data analysis

Quality control of all the reads obtained for each sample was done using NGSQC Tool kit (Patel et al. 2012) and reads with Phred score >Q20 were selected for further analysis. Reference genome of barley from ENSEMBLE (plants/release40/gff3/hordeum_vulgare) was used for alignment of reads and transcript identification. TopHat pipeline (Trapnell et al. 2009) was used for the alignment of reads on reference genome. Cufflink and Cuffdiff pipelines (Trapnell et al. 2009) were used for identification of differentially expressed transcripts in wild type and transgenic replicates using default parameters. Transcripts with log2 ratio ≥2 were considered as differentially expressed. Unsupervised hierarchical clustering of differentially expressed genes was done using Cluster 3.0 and visualized using Java Tree View. Gene ontologies and pathways

that harbor expressed transcripts were identified using DAVID Functional Annotation Tool (DAVID Bioinformatics Resources 6.8, NIAID/NIH) (Huang et al. 2009).

e.) Statistical analysis of differentially expressed transcripts

CuffDiff pipeline was used for analysis of differentially expressed transcripts among samples of wild type and transgenic replicates. The threshold of absolute fold change was set ≥2 and statistically significant P value threshold adjusted for false discovery rate of less than 0.001 by applying Student's t-test. Statistically significantly enriched functional classes with a P value adjusted for false discovery rate of less than 0.05 derived using the hypergeometric distribution test corresponding to differentially expressed genes were determined using Student's t-test with Benjamini Hocheberg FDR test.

f.) Biological Pathways and Gene Ontology Enrichment Analysis

GO ontologies and KEGG pathways analysis was done using DAVID Functional Annotation Tool (DAVID Bioinformatics Resources 6.8, NIAID/NIH) (Huang et al. 2009).

3.3.7 Biochemical analysis of overexpressing transgenic lines

a.) Preparation of plant extract

For the estimation of anti-oxidative enzyme activities, 500 mg of leaves of transgenic and wild type plants were homogenized in 5 mL of extraction buffer (Table 3.27) in a pre-chilled mortar and pestle using liquid nitrogen. The homogenate was centrifuged at 22,000 × *g* for 20 minutes at 4˚C. The supernatant was re-centrifuged at 4˚C. The supernatant (enzyme extract) was collected in fresh 5 mL MCT and stored at ice till the completion of experiment.

Table 3.27 Composition of one step assay buffer

S. No.	Components	Concentration
1.	Phosphate buffer	100 mM
2.	Na_2EDTA	2 mM
3.	PVP	1%

b.) Estimation of superoxide dismutase activity

The activity of superoxide dismutase was determined by measuring its ability to inhibit the photochemical reduction of nitroblue tetrazolium chloride, as described by Giannopolitis and Ries (1977). The assay mixture consisted of 300 µL of the enzyme extract, 100 mM phosphate buffer (pH 7.8), 3 mM EDTA, 200 µM methionine, 2.25 µM nitroblue tetrazolium and 60 µM riboflavin (added at last) in a total volume of 3.2 ml. The tubes were inverted for 2-3 times for proper mixing and placed under fluorescent light. After 15 minutes lights were switched off to stop the reaction. Then the tubes were immediately put in dark. Absorbance was measured at 560 nm, and one unit of SOD activity (U) was defined as the amount of enzyme required to cause 50% inhibition of the nitroblue tetrazolium photoreduction rate. The results were expressed as Umg^{-1} protein.

c.) Estimation of ascorbate peroxidase activity

The activity of ascorbate peroxidase was assayed according to Nakano and Asada (1981). The assay mixture consisted of 150 µL of the enzyme extract, 50 mM phosphate buffer (pH 7.0), 0.2 µM EDTA, 0.2 mM ascorbate, and 2.0 mM H_2O_2 in a total volume of 3 mL. Ascorbate oxidation was monitored by reading the absorbance at 290 nm at the moment of H_2O_2 addition and 1 min later. The difference in absorbance (ΔA_{290}) was divided by the ascorbate molar extinction coefficient (2.8 $mM^{-1}.cm^{-1}$) and the enzyme activity expressed as µmol of H_2O_2 min^{-1} mg^{-1} protein, taking into consideration that 1.0 mol of ascorbate is required for the reduction of 1.0 mol of H_2O_2 (McKersie and Leshem, 1994).

d.) Estimation of catalase activity

Catalase specific activity was measured according to Aebi et al. (1974). The assay mixture consisted of 100 µL enzyme extract, 100 mM phosphate buffer (pH 7.0) and 200 mM H_2O_2 in a total volume of 1.5 mL. The H_2O_2 was added in the last and absorbance was measured at 240 nm. The extinction of substrate (H_2O_2) by enzymatic decomposition was measured by decrease in absorbance till 3 minutes.

3.3.8 Comparative metabolomics of wild type plants of cultivar Golden Promise under controlled and heat stressed conditions

Comparative metabolomics analysis of barley was performed under normal and heat stressed conditions using leaf tissue.

a.) Sample collection and heat treatment

Seeds of barley cultivar Golden Promise were grown in green house at 20°C day/18°C night temperature and 16 hours light. For metabolomics analysis heat stress treatment at 42°C was given to five-weeks-old seedlings in triplicate for two hours and leaf samples were harvested and snap freezed in liquid nitrogen. Similarly, leaf tissues from plants under controlled conditions were harvested in triplicate.

b.) Sample preparation for GC-MS analysis

1. After plugging leaves was immediately rinse in distilled water for 5 minutes with vigorous shaking followed by air drying (hot air, preferably using hair dryer for 30 minutes).
2. Dried leaf samples were crushed in liquid nitrogen. Extraction mixture was prepared by adding methanol/water/chloroform in 2.5:1:1 (v/v/v) ratio.
3. One milliliter of pre-cooled (at -20°C) extraction mixture was then added to 200 mg powdered leaf samples in a 1.5 mL micro centrifuge tube and vortexed vigorously at room temperature for 5 minutes.
4. In order to identify extraction efficiency, 50 µL of 2-phenylphenol (2 mg/mL methanol stock) was spiked in the extraction mixture as the internal standard (IS) and re-vortexed for 1 minute. The homogenized extracts were then centrifuged at 14000 g for 5 minutes.
5. The resulting supernatant (0.8 mL) was transferred into a new 1.5 mL tube. Water (0.4 mL) was added to the supernatant, vortexed for 10 s and then centrifuged at 14000 g for 5 minutes.
6. The polar upper phase (methanol/water) was transferred to a new tube and used for metabolite measurements by GC-MS.
7. Prior to GC-MS analyses, first extract was dried out in a vacuum concentrator (Labconco, Centrivap; USA) without heating for 2 hours followed by freeze drying for 12 hours in a lyophilizer.

8. Finally, dried material was subjected to double derivatization for GC-MS analyses.
9. First derivatization was performed by dissolving the dried samples in 40 µL of methoxyamine hydrochloride (stock solution: 20 mg/mL in pyridine) and incubating the solutions at 37°C for 2 hours.
10. Then second derivatization was performed by adding 80 µL of N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) at 37°C for 30 minutes, empty reaction tube was taken as control and is also subjected to derivatization process.

c.) GC-MS analysis using Agilent GC-MS system

GC- MS analysis was performed on Agilent GC-MS system comprising of Agilent 7890A gas chromatograph (Agilent technologies, CA, USA) coupled with an Agilent 5975C mass detector (Agilent technologies, CA, USA). Derivatized sample (1 µL) was injected in GC-MS by auto sampler (7683 B series, Agilent Technologies) with a split ratio of 1:5. Samples were separated on fused silica capillary column DB-5 MS (5% phenyl methyl polysiloxane: 30m x 0.25mm i.d. x 0.25 µm, Agilent technologies). The temperature program was as follows: Initial temperature of 70°C for 5 minutes, followed by increase in temperature to 270°C at the ramp rate of 5°C/min and was hold for 2 minutes. Further the temperature was increase to 300°C at the ramp rate of 10°C/min and was kept on final hold temperature of 300°C for 15 minutes. Total run time calculated was 65 minutes. Helium gas of ultra-high purity was used as carrier gas at a flow rate of 1mL/min. The inlet temperature and interface temperature were set 280°C. The MS unit was tuned to its maximum sensitivity and the total ion current was recorded for mass range was m/z 80-700, and the detector voltage was set at 1700 V. Each sample was replicated three times. Scan was started after solvent delay of 5 minutes with scan frequency 4 S^{-1} (2.0 HZ).

d.) Metabolomics data analysis

Metabolites present in the barley leaf samples were identified by library matching of mass spectra of each compound using the NIST-17 mass spectral library (National Institute of Standards and Technology), and our in-house library that include several secondary metabolites, amino acids, organic acids, and sugar standards. Metabolite identity was obtained and reported only when the matching value of the mass spectra comparison was more than 75, and identity of compounds was further confirmed by spiking the sample with the available standards (when available). No co-elution detected in any of the identified peaks.

Raw GC-MS data files obtained after analysis with Agilent ChemStation^{TM}software were deconvoluted by Automated Mass Spectral Deconvolution and Identification System (AMDIS) linked with WsearchPro (www.wsearch.com.au) and by utilizing tools available with WsearchPro software. Metabolite data obtained from different samples (control vs heat treated) was further combined in a single file and is further converted into .csv (comma separated values) format before uploading to Metaboanalyst 3.0. Finally, the data obtained was normalized using internal standard. After that, the data were, log transformed with Pareto scaling (mean-centered and divided by the square root of standard deviation of each variable) followed by normalization before statistical analyses. Multivariate statistical analyses like ANOVA (using Fisher's LSD method; p value < 0.05), principal component analyses (PCA) were performed by using interactive online tool Metaboanalyst 3.0 (http://www.metaboanalyst.ca). The output for PCA data consisted of score plots for visualizing the contrast between different leaf samples and loading plots to explain the cluster separation. A heat map was created using interactive heat map tool of Metaboanalyst 3.0.

Results