The Arabidopsis thaliana Heat Shock Transcription Factor A1b Transcriptional Regulatory Network

Waleed S. Albihlal

A thesis submitted for the degree of Doctor of Philosophy

School of Biological Sciences

University of Essex

Date of submission 12/01/2015

Table of contents

CHAPTER 6 The Arabidopsis thaliana HSFA1b gene in yeast

CHAPTER 7 Final discussion and future direction

List of tables:

List of figures:

Acknowledgments:

I would like to start by thanking my supervisor Prof. Philip M. Mullineaux for the superb guidance and support from the early days until the last second of this project. Without his continuous help and amazing advice I would not have reached this level of knowledge and this work would not have seen the light of day. This, without a doubt, will be one of the most memorable stages in my life and I will always be proud that I worked with such an excellent scientist.

I would also like to express my appreciation to all the lab members, especially to Dr. Rubén Álvarez-Fernández whose excellent technical advice has been of great importance for the success of this project and Dr. Igor Chernukhin, the director of the bioinformatics unit, for the extraordinary computing advice he provided throughout the research which undeniably aided solving some of the trickiest puzzles in this research.

I would also like to thank Prof. Dennis J. Thiele and his research technician Ms. Carianne Jones from Duke University School of Medicine, NC, USA, for providing the yeast strain PS145 and all the supporting information which have been crucial for the success of a major part of this work.

I would also like to very much thank the ministry of higher education in the Kingdom of Saudi Arabia for funding and sponsoring my PhD program. This work would not have been possible without the extremely generous financial support provided by the government of the Kingdom of Saudi Arabia.

Finally, a very special thank you goes to my mother and brothers who kept motivating me, pushing me, and providing all the support I needed to carry on despite all the hard times and shocking losses we all have been through in the last four years. This meant so much to me and this work is the least I can provide as a token of gratitude.

Author's declaration:

This thesis has been written by myself and has not been submitted to any previous application for any degree. Unless otherwise stated, the work in this thesis has been carried out by myself. This thesis is written in accordance with the regulations for the degree of Doctor of Philosophy at the University of Essex.

Abstract:

Plants as sessile organisms have adapted highly sophisticated cellular processes to cope with environmental stress conditions, which include the initiation of complex transcriptional regulatory circuits. The heat shock transcription factors (HSFs) have been shown to be central regulators of plant responses to abiotic and biotic stress conditions. However, the extremely high multiplicity in plant HSF families compared to those of other kingdoms and their unique expression patterns and structures suggest that some of them might have evolved to become major regulators of other non-stress related processes. *Arabidopsis thaliana* HSFA1b (AtHSFA1b) has been shown to be a major regulator of various forms of plant responses to abiotic and biotic stresses. However, it has also been suggested that overexpression of *AtHSFA1b* results in a subtle developmental effect in *Arabidopsis thaliana* and *Brassica napus* in the form of increased seed yield and harvest index. Through genome-wide mapping of the AtHSFA1b binding profile in the *Arabidopsis thaliana* genome, monitoring changes in the AtHSFA1b-regulated-transcriptome, and functional analysis of AtHSFA1b in *Saccharomyces cerevisiae* under non-stress and heat stress conditions, this study provides evidence of the association of AtHSFA1b with plant general developmental processes. Furthermore, the outcome of this research shows that AtHSFA1b controls a transcriptional regulatory network operating in a hierarchical manner. However, in an agreement with a previously suggested model, the results from this study demonstrate that the involvement of AtHSFA1b in the regulation of heat stress response in *Arabidopsis thaliana* is possibly limited to the immediate and very early phases of heat stress response which also results in a collapse in its transcriptional network which seems to be accompanied by a general shutdown in plant growth and development.

Abbreviations:

35S: Cauliflower mosaic virus promoter.

A. brassicicola: Alternaria brassicicola.

A. thaliana: Arabidopsis thaliana.

ABA: Abscisic acid.

AtHSF: A. thaliana Heat Shock Transcription Factor.

AtHSFA1b: *A.thaliana* Heat Shock Transcription Factor A1b.

AtHSFA1b::mRFP: AtHSFA1b tagged with mRFP.

AtHSP: A. thaliana Heat Shock Protein

B. napus: *Brassica napus*

bp: Base pair

BiFC: Biomolecular Fluorescence Complementation.

cDNA: complementary DNA.

ChIP: Chromatin Immunoprecipitation.

ChIP-CHIP: Chromatin Immunoprecipitation followed by tiling array.

ChIP-PCR: Chromatin Immunoprecipitation followed by PCR.

ChIP-SEQ: Chromatin Immunoprecipitation followed by Sequencing.

Col-0: *A. thaliana* Colombia-0 ecotype.

CTF: Constitutively Expressed Transcription Factor.

DBD: DNA Binding Domain.

DEGs: Differentially Expressed Genes.

D. melanogaster: Drosophila melanogaster.

dmHSF: *D. melanogaster* HSF.

EDTA: Ethylenediaminetetraacetic acid.

HSE: Heat Shock Element.

HSF: Heat Shock Transcription Factor.

hHSF: Human Heat Shock Transcription Factor.

HSP: Heat Shock Protein.

HSR: Heat Shock Response.

ITF: Inducible Transcription Factor.

MEME: Multiple Expectation-maximisation Motif Elucidation.

mHSF: Mouse Heat Shock Transcription Factor.

mRFP: Modified Red Fluorescent Protein.

mRNA: Messenger RNA.

O. sativa: Oryza sativa.

OsHSF: *O. sativa* Heat Shock Transcription Factor.

PCR: Polymerase Chain Reaction.

PMSF: Phenylmethylsulfonyl fluride.

qRT-PCR: Quantitative Real Time PCR.

redox: Reduced/Oxidised

rGADEM: R Genetic Algorithm guided formation of spaced Dyads coupled with Expectation maximisation

RNA-SEQ: RNA sequencing.

ROS: Reactive Oxygen Speices.

S. cerevisiae: Saccharomyces cerevisiae.

SlHSF: *S. lycopersicum* Heat Shock Transcription Factor.

sHSP: Small Heat Shock Protein.

S. lycopersicum: Solanum lycopersicum.

STEME: Suffix Tree Expectation-maximisation Motif Elucidation.

TAD: Trans-activational Domain.

TAE: Tris Acetic acid EDTA.

TBP: TATA-binding Protein.

TF: Transcription factor.

TRD: Trans-repressional Domain.

Y1H: Yeast one-hybrid.

Y2H: Yeast two-hybrid.

yHSF: Yeast HSF.

VPD: Vapour Pressure Deficit.

YPDA: Yeast Extract Peptone Dextrose Adenine.

YPGA: Yeast Extract Peptone Galactose Adenine.

WHTH: Winged Helix-Turn-Helix.

DE: Differential Expression.

CHAPTER 1

Background

1.1 The heat shock response:

The Heat shock response (HSR) was first discovered in 1962 by Ferruccio Ritossa when he showed that heat induces puffs in the chromosomes of salivary glands in *Drosophila melangoaster* (*D. melanogaster*) larvae. Later, it was shown that those puffs are caused by the activation of genes coding for heat shock protein (HSP) chaperones (Lis et al., 1981). HSR is a highly conserved process among all eukaryotes (Wang et al., 2004; Schlesinger, 1990; Lindquist 1986). Despite being called HSR, it has become widely accepted as a general stress response mechanism where it is highly induced by various forms of stress stimuli including heavy metals, oxidative stress, and pathogens as well as heat (Bechtold et al., 2013; Wang et al., 2004; Carper et al., 1987). The high conservation of HSR among all eukaryotes suggests that it is a crucial regulatory mechanism for survival under stress conditions (Åkerfelt et al., 2010).

The HSR is typically characterised by the strong and fast induction of genes coding for HSP chaperones (Wang et al., 2004; Linquist, 1987). During stress, HSPs accumulate in cells to aid refolding of denatured proteins, prevent the aggregation of damaged proteins and maintain protein homeostasis under stress (Åkerfelt et al., 2010; Miller and Mittler, 2006; Kregel, 2002; Vierling, 1991). This process helps cells to cope with the deleterious states caused by protein damage in stressful conditions (Åkerfelt et al., 2010). In addition, one of the main characteristics of HSR is the downregulation of some non-HSP genes in favour of expression and synthesis of HSPs (Åkerfelt et al., 2010). The HSR and the expression of HSPs are regulated at a transcriptional level by a family of transcription factors (TFs) known as heat shock transcription factors (HSFs) (Scharf et al., 2012; Åkerfelt et al., 2010; Nover et al., 2001; Pirkkala et al., 2001).

1.1.1 The heat shock transcription factors:

Similar to HSPs, the heat shock transcription factor (HSF) family is one of the most conserved TF families across all species (Scharf et al., 2012; Åkerfelt et al., 2010; Nover et al., 2001). This group of TFs belongs to a family of proteins known as winged helix-turn-helix (WHTH) DNA-binding proteins (Sakurai and Enoki, 2010; Liu and Thiele, 1999; Littlefield and Nelson, 1999). All HSFs in all species share very similar structures. The basic structure of an HSF monomer consists of, a highly conserved DNA binding domain (DBD), which recognises and binds the DNA heat shock *cis*-element (HSE), a proximal hydrophobic oligomerisation domain (HR-A/B) which is the region where HSF monomers interact with each other to form functional trimers, nuclear localisation signal (NLS) element that allows HSFs to enter the nucleus, nuclear export signal (NES) which allows HSFs to exit the nucleus, and trans-activation domain (TAD) (Scharf et al., 2012; Åkerfelt et al., 2010; Nover et al., 2001) (figure 1.1)

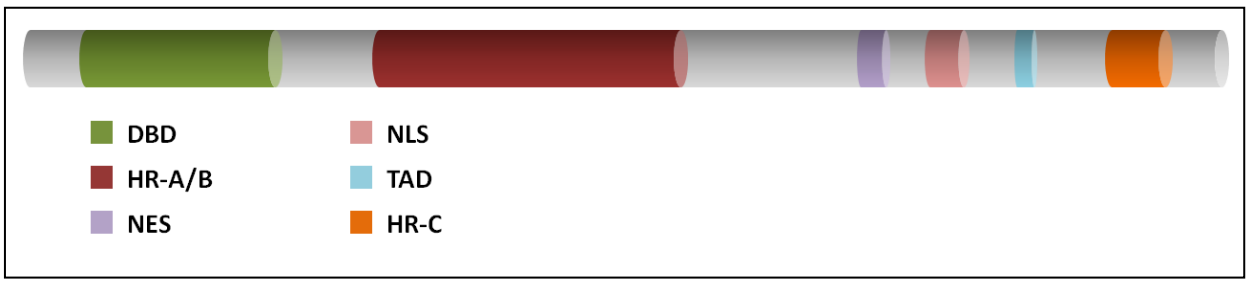


Fig.1.1. General structure of HSF monomers. Schematic diagram illustrating the functional domain in HSF monomers. HSF monomers consist of a highly conserved DNA binding domain (DBD), hydrophobic oligomerisation domain (HR-A/B), nuclear export signal (NES), nuclear localisation signal (NLS), trans-activation domain (TAD), and a second oligomerisation domain (HR-C).

There are, however, structural differences between HSFs and even between HSFs in the same species. Some HSFs do not contain TAD elements and others have been shown to possess trans-repression domains (TRD) which serve as binding region for co-repressors (Scharf et al., 2012; Miller and Mittler, 2006; Nover et al., 2001). Those HSFs act as repressors of transcription rather than activators (Scharf et al., 2012; Miller and Mittler, 2006). Another structural difference between HSFs is the presence of a second oligomerisation domain proximal to the C-terminal

end of the protein (HR-C). This functional domain is responsible for the inactivation of HSF through its interaction with the HR-A/B domain on the same monomer preventing the formation of trimers and stabilising the inactive monomeric state of HSFs (Scharf et al., 2012; Åkerfelt et al., 2010; Nover et al., 2001). However, not all HSFs contain the HR-C domain and those that do not contain the HR-C functional domain are thought to be constitutively in an active trimeric form under all conditions (Scharf et al., 2012).

The highly conserved DBDs in all HSFs allow them to bind to a highly conserved DNA consensus sequence. All known HSFs in all species bind to the pentameric consensus sequences nGAAn known as heat shock *cis*-acting regulatory elements (HSEs) (Ahn et al., 2001; Littlefield and Nelson, 1999). Due to the structural configuration of HSFs and the *in vivo* active trimeric state, it has been shown that HSFs bind to three inverted repeats of the conserved nGAAn consensus sequence in the form nGAAnnTTCnnGAAnn / nTTCnnGAAnnTTCn (Ahn et al., 2001; Liu and Thiele, 1999; Littlefield and Nelson, 1999). These observations gave more insights about the functional form of HSFs *in vivo* and clearly demonstrated that HSFs are capable of binding the DNA and activating/repressing transcription *in vivo* only in trimeric forms (Sakurai and Enoki, 2010; Ahn, et al., 2001). Although *in vitro* studies have shown that HSFs are capable of binding the DNA in a monomeric and dimeric states, the majority of studies showed that this not the case *in vivo.* The conversion of HSFs from monomeric to trimeric state increases their binding affinities to HSEs and their affinities to the NLS receptors located on the nuclear envelope which by turn allow them to translocate into the nucleus (Scharf et al., 2012; Ahn et al., 2001) (figure 1.2)

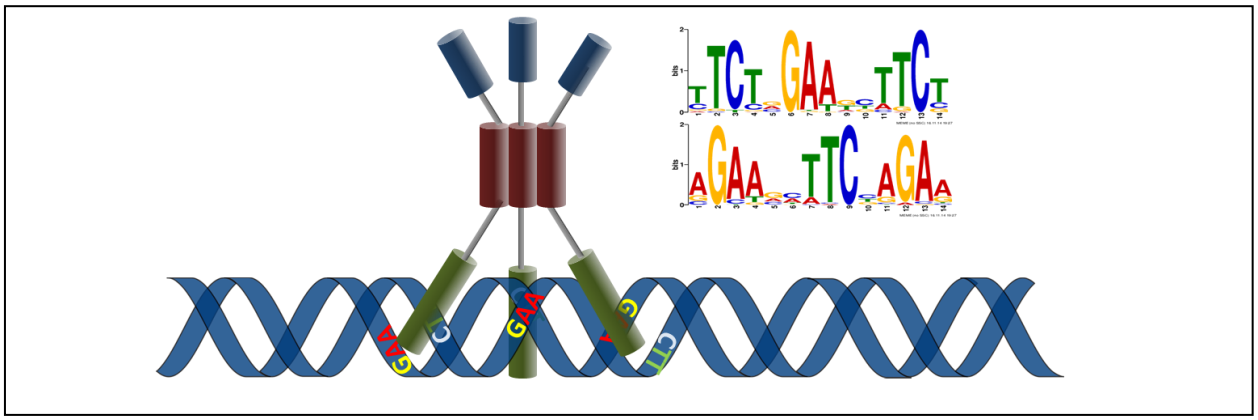


Fig.1.2. HSFs bind to HSEs *in vivo* in trimeric form. A model showing the structural state of HSFs bound to the DNA *in vivo* and the DNA consenses sequence of HSEs. Trimerisation of HSFs increases the affinity of their DBDs to the DNA binding elements (HSEs) and the NLS to their receptors on the nuclear envelope. HSEs, typically, consist of three inverted repeats of the core HSE sequence GAA with guanidine being the most important base in HSEs.

Since the discovery of HSFs and their target HSP genes up until recent times the main focus has been on the roles of HSFs as major regulators of HSR. It has been shown in many studies that HSFs transform into active trimeric forms which allow them to translocate into the nucleus then bind to HSEs on the promoters *HSPs* leading to the expression of *HSPs* in response to elevated temperature (Scharf et al., 2012; Åkerfelt et al., 2010; Nover et al., 2001) (figure 1.3).

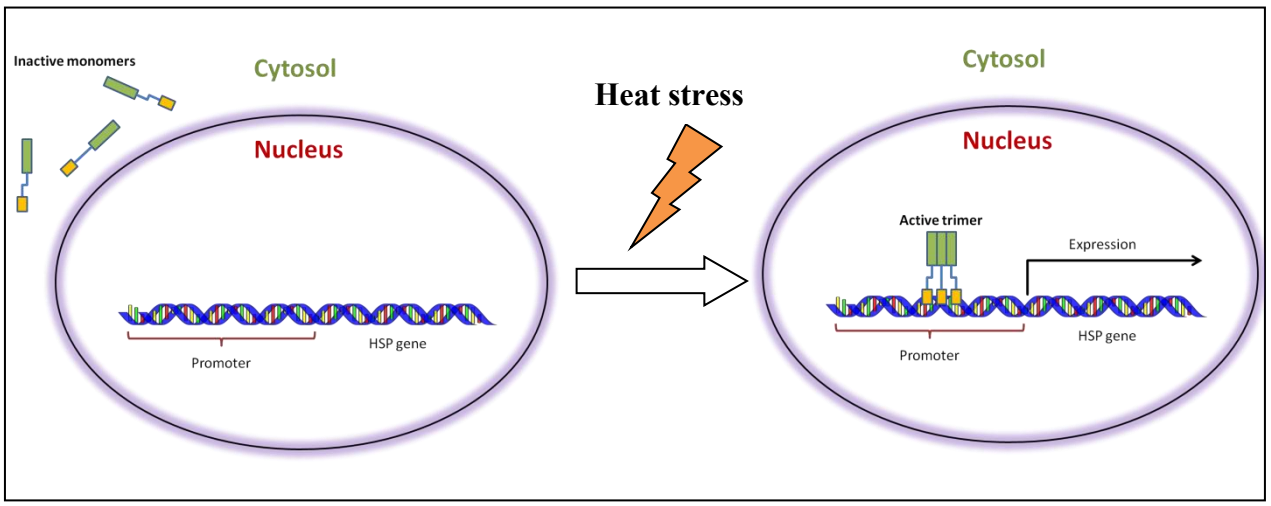


Fig.1.3. HSFs activate the expression of *HSPs* upon heat stress. Classic model showing the induction of HSP genes by HSFs in response to elevated temperature. HSFs were thought to be found in the cytosol in inactive monomeric states under normal conditions. Heat stress induces trimerisation of HSFs which allows them to translocate into the nucleus, bind to HSEs located on the promoter of HSP genes and subsequently activate the expression of *HSPs* under heat stress.

However, more recent studies have established the involvement of HSFs in the regulation of cellular response to other forms of stress. HSFs in mammals, for example, are directly involved

in the regulation of cellular response to oxidative stress, heavy metal and viral and bacterial pathogens and also implicated in cancer in the absence of the heat stresss componenet (Singh and Aballay, 2014; Zaarur et al., 2006; Jauniaux et al., 2000; Wagner et al., 1999). Moreover, it has been shown that HSFs are involved in the regulation of plants' response to not only heat stress but various other forms of abiotic and biotic stress (Pérez-Salamó et al., 2014; Hwang et al., 2013; Bechtold et al, 2013**;** Koskull-Döring et al, 2007, Miller and Mittler 2006). Based on the recent observations, HSFs have become widely known as major regulators of general stress response not only heat. These findings also revealed some important aspects about the crosstalk in stress responses in general. However, there is emerging evidence that HSFs involvements go beyond stress response to the regulation of crucial non-stress related processes.

1.1.1.1 in *Saccharomyces cerevisiae:*

Saccharomyces cerevisiae (*S. cerevisiae*) or baker's yeast possesses a single HSF. Yeast HSF (yHSF), by far, has the largest molecular mass of any known HSF in any species (Pirkkala et al., 2001) (figure 1.4). It possesses an unusual structure compared to other HSFs from in other species where it contains two distinct TADs, one located near the N-terminal end and the other is located proximal to the C-terminal end of the protein (Morano et al., 1999). Also, one of the main structural characteristics of yHSF is its lack of HR-C domain which is thought to be one of the reasons why yHSF is constitutively in an active trimeric form under all conditions (Liu and Thiele, 1999; Morimoto, 1998). Moreover, *in vitro* studies showed that the structure of yHSF is highly dynamic and it undergoes intrinsic conformational changes in response to elevated temperature and the reactive oxygen species superoxide anion. A flexible linker located between the DBD and HR-A/B domains is thought to be responsible for the flexibility of the

structure of yHSF (Erkine et al., 1999; Flick et al., 1994; Sorger 1990).

The structural flexibility of yHSF allows it to bind to various forms of HSEs other than the canonical nGAAn pentameric. It can bind to forms of extended HSEs known as gapped HSEs where the spacer between the core GAA consensus sequence can be a stretch of up to 5bp (Santoro et al., 1998).

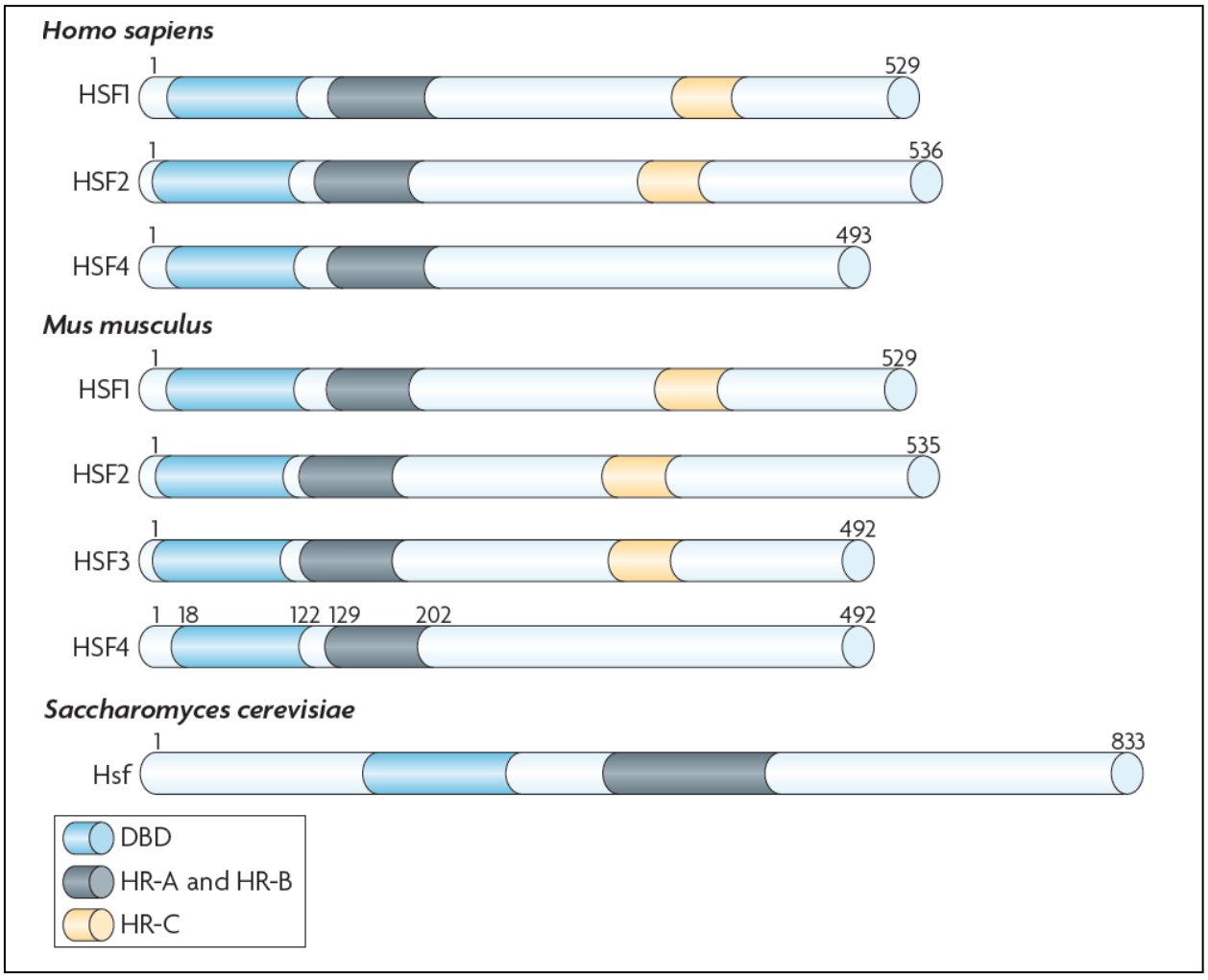


Fig.1.4. yHSF is the largest known HSF. Schematic diagram showing the size of yHSF compared to human and mouse HSFs. yHSF is the largest known HSF compared to those in other species. The diagram also shows some of the unique structural characteristics of the yHSF which are the presence of a large N-terminal activation domain, large linker between DBD and HR-A/B domains and the absence of HR-C domain. Numbers above each diagram represent the count of amino acids in each HSF monomer. The figure was adapted from (Åkerfelt et al., 2010).

The expression of HSPs in yeast is solely controlled by yHSF under normal growth condition and stress (Liu et al., 1997). However, the role of yHSF is not limited to regulation of HSR but

exceeds that to other crucial cellular functions making it an extremely multifunctional TF. Knockout studies have shown that loss of yHSF is lethal under normal growth conditions (Wiederrecht et al., 1988; Jakobson and Pelham, 1988). These results established the important role of yHSF for yeast cell viability and survival. It has been reported that yHSF is also directly involved in the regulation of cell cycle genes (Venturi et al., 2000; Smith and Yaffe, 1991). The outcome of yHSF knockout studies gave an indication that HSFs in other species might also be involved in the regulation of diverse cellular functions that go beyond just regulation of stress response.

1.1.1.2 in *Drosophila melanogaster*

Similar to yeast, fruit fly (*D. melanogaster*) possesses only one HSF (dmHSF; Åkerfelt et al., 2010). *In vitro* analysis of dmHSF showed that it is subject to intrinsic conformational changes in response to elevated temperature and oxidative chemicals. However, one of the interesting differences between yHSF and dmHSF is that the latter changes its conformation in response to hydrogen peroxide not superoxide anion (Zhong et al., 1998). Unlike yeast, dmHSF is not required for the survival of *D. melanogaster*. Loss of dmHSF leads to hyper-sensitivity to elevated temperatures but does not lead to mortality of *D. melanogaster* (Jedlicka et al., 1997). However, it has been shown that the loss of dmHSF leads to impaired growth of *D. melanogaster* larvae when exposed to elevated temperature (Fujimoto and Nakai, 2010). Further investigation also revealed that loss of dmHSF leads to defective development in *D. melanogaster* oogenesis (Fujimoto and Nakai, 2010; Jedlicka et al., 1997). Genome-wide scanning of dmHSF binding sites showed that the vast majority of its target genes under no stress are not associated with HSR only (Gonsalves et al., 2011; Guertin and Lis, 2010). These results further developed the idea that dmHSF involvement might not be limited to regulation

of stress response but also involves regulation of various crucial cellular processes under non-stress conditions, despite showing that it is not required for viability of *D. melanogaster*.

1.1.1.3 in Vertebrates:

Unlike *S. cerevisiae* and *D. melanogaster*, vertebrates contain multiple HSFs. There are 4 known HSFs in vertebrates, HSF1, HSF2, HSF3 and HSF4 (Åkerfelt et al., 2010). HSF1 and HSF2 are the most studied due to their constitutive expression patterns in all tissues and cell types (Åkerfelt et al., 2010). The expression of HSF4 is limited to eye and brain tissue, and HSF3 is only found in avian species (Pirkkala et al., 2001). The multiplicity of HSFs in vertebrates allowed for more versatility and divergence in their functions compared to their relatives in yeast and *D. melanogaster*.

Mammalian HSFs possess unique and overlapping functions. For example, HSF1, is considered to be the sole master regulator of HSR in mammals (Westerheide and Morimoto, 2005; Liu et al., 1997). Other HSFs are responsible for the regulation of other cellular processes and have little involvement in the regulation of HSR (Östling et al., 2007). While human HSF1 (hHSF1) is constitutively expressed in all tissue, it remains in an inactive monomeric form and only transforms into an active trimer in the presence of stress (Åkerfelt et al., 2010). Other members of the HSF family in human and mouse are constitutively active and their expression pattern is not responsive to stress (Åkerfelt et al., 2010). It has also been shown that hHSF2 is incapable of inducing the expression of HSPs by itself. It can only induce the expression of HSPs through interacting with hHSF1 (Östling, et al., 2007).

It has become evident that the transcriptional regulation by HSFs in vertebrates is more complex than species that contain single HSFs such as yeast and *D. melanogaster*. For instance, there are cases where certain HSFs possess dual molecular functions. For example, hHSF4 acts as an activator of gene expression; however, a splice variant of hHSF4 leads to a dramatic change in its function and converts it to a repressor (Tanabe et al., 1999). Reports have shown that functional involvement of hHSFs and mHSFs, including the sole activator of HSR, HSF1, is not limited to stress response (Westerheide and Morimoto, 2005). Knockout studies have shown that loss of mouse HSF1 (mHSF1) leads to a severe developmental impairment in mice including neurodegeneration and development of muscle atrophy (Konodo et al., 2013). Loss of mHSF2, on the other hand, results in increased embryonic lethality, mental retardation and defective spermatogenesis (Wang et al., 2003). Furthermore, it was shown that mHSF4 is required for cell differentiation in eye lens and, therefor, for proper eye development (Min et al., 2004). These examples and others that are not mentioned in this review strongly suggest that HSFs possess a developmental component beside their involvement in the regulation of the stress responses.

1.1.1.4 in Plants:

The first striking observation when looking at plants is their large HSFs families compared to other species. For instance, there are 25 HSFs in rice (*Oryza sativa*), 25 in tomato (*Solanum lycopersicum*), and 21 in Arabidopsis (*Arabidopsis thaliana*) (Scharf et al., 2012; Miller and Mittler, 2006; Nover et al., 2001). Based on their structures, plant HSFs are divided into three distinctly conserved classes (A, B, and C) (Scharf et al., 2012; Nover et al., 2001). The basic structure of plant HSFs is highly similar to those of other species. They contain highly similar functional domains to those of other HSFs in other species which include DBD, HR-A/B, NLS,

TAD/TRD, and NES (Scharf et al., 2012; Miller and Mittler, 2006; Nover et al, 2001). The structures of plant class-A HSFs differ from those in yeast, *D. melanogaster* and vertebrate mainly in the HR-A/B domain where it is considerably larger in plant class-A HSFs than those in other species (Nover et al., 2001). The large HR-A/B domain in plant HSFs is thought to provide a larger hydrophobic surface that aids the formation and stabilisation of their timeric forms. However, the HR-A/B domains in class-B and class-C plant HSFs are highly similar to those in other species (Miller and Mittler, 2006) (figure 1.5).

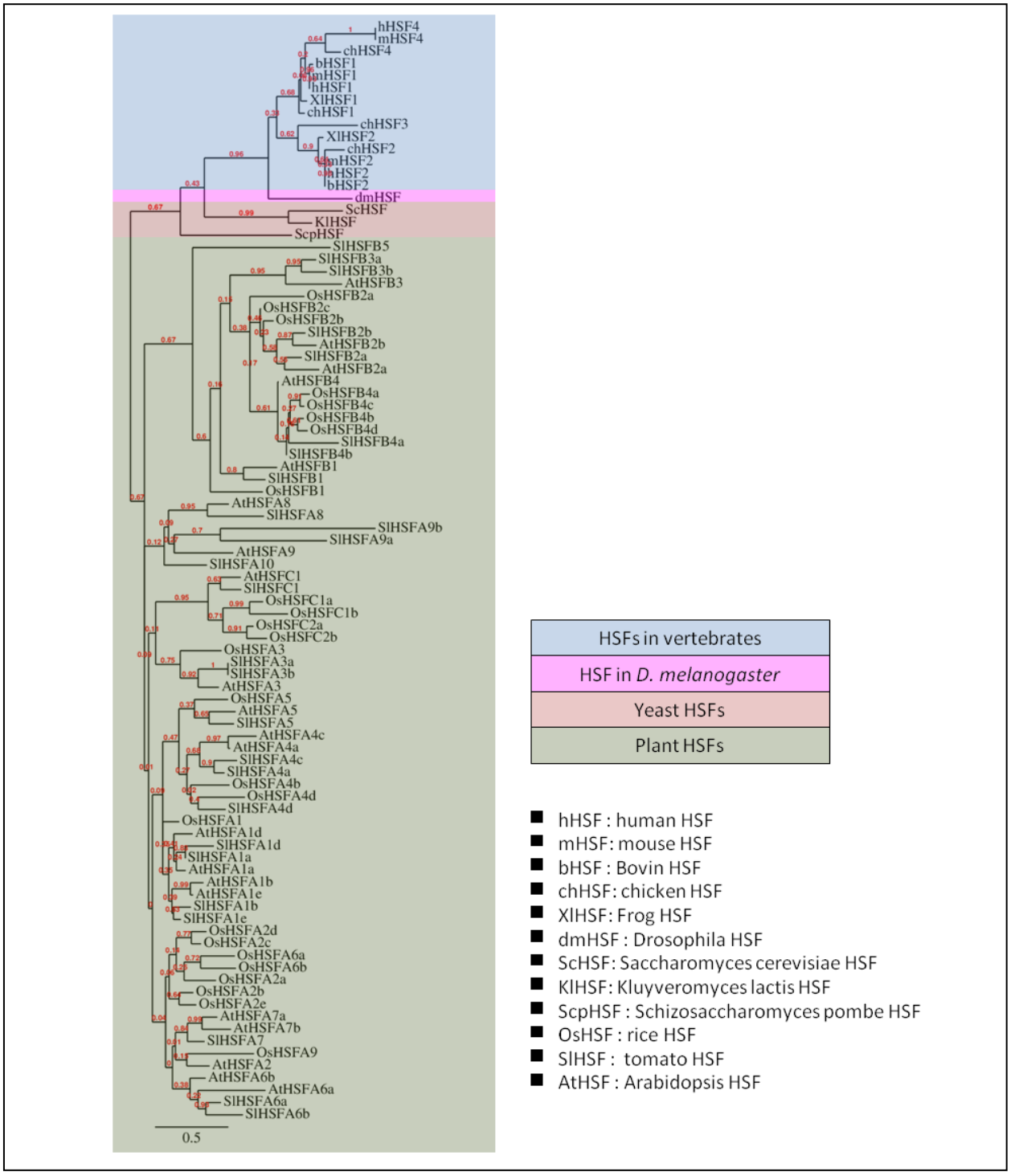


Fig.1.5. Plant HSFs compared to HSFs in other species. A phylogenetic tree showing the relationship between members of HSFs in higher eukaryotes. The tree shows that plant HSFs are structurally divergent from other HSFs in other eukaryotes. The phylogenetic tree was constructed in (http://www.phylogeny.fr) using the **Mu**ltiple **S**equence **C**omparison by **L**og-**E**xpectation (MUSLCE; Edger, 2004) and all gaps were removed from the analysis. Numbers in red show the bootstrap values from 100 bootstrap replicates carried out. HSFs from three plant species were used in the analysis (rice, tomato and Arabidopsis).

There are structural and functional diversities among plant HSFs. Class-A HSFs contain TADs and are known to be activators of transcription. Class-B lack the C-terminal TAD and thought to be repressors of transcription (Scharf et al., 2012; Miller and Mittler, 2006; Nover et al., 2001). There is still no clear evidence whether class-C HSFs act as activators or repressors despite their lack of the C-terminal TAD (Schmidt et al., 2012; Chauhan et al., 2011). Well studied plants HSFs in Arabidopsis thaliana (*A. thaliana*) and tomato (*S. lycopersicum*) recognise and bind to almost identical HSEs to the ones in yeast, *D. melanogaster* and vertebrates (Scharf, et al., 2012; Nover, et al., 2001) (figure 1.6).

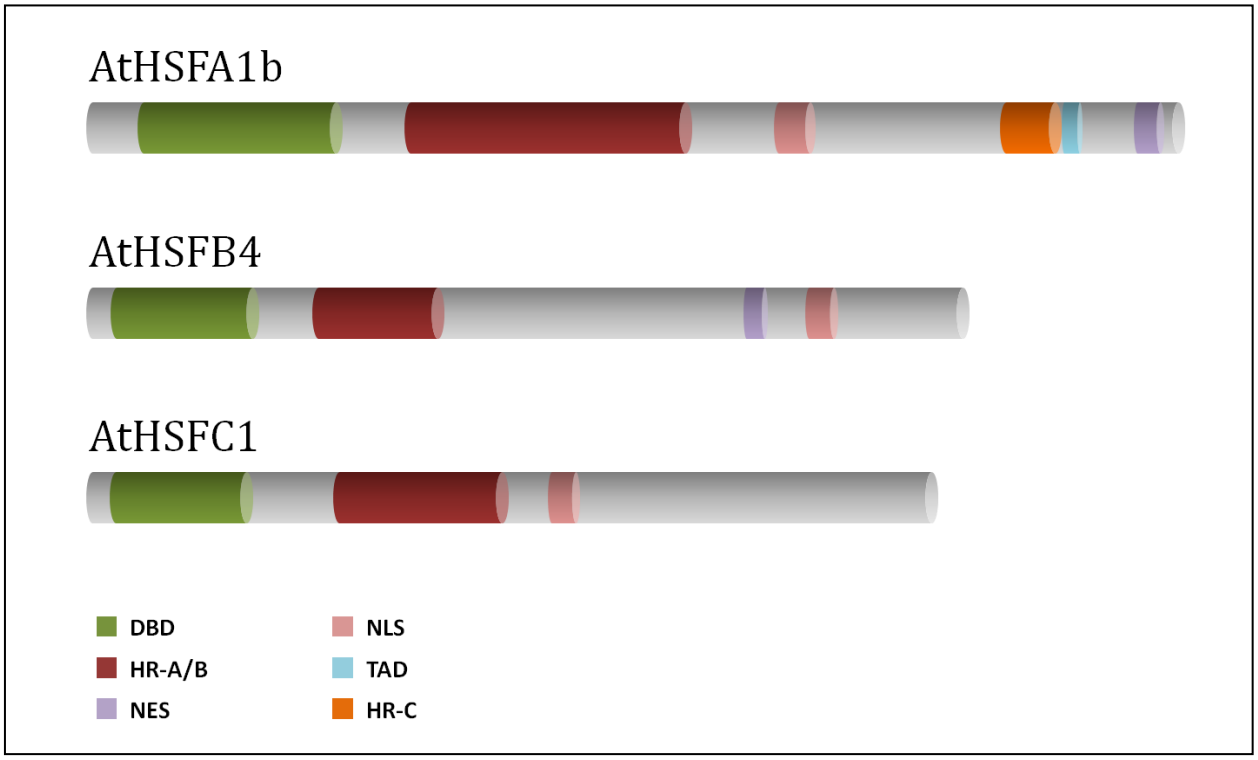


Fig.1.6. Structural differences among members of different HSF classes in plants. Schematic diagram showing the structural differences between classes A, B, and C HSFs in *A. thaliana*. Members of all classes of plant HSFs contain a highly conserved DBD proximal to the N-termina. The HR-A/B is poorly conserved among different classes of plants HSFs. HR-A/B domains in members of class-A HSFs are larger than those in classes B and C. Class-A HSFs are characterised by the presence of TAD elements allowing them to function as activators of transcription. Class-B HSFs do not contain TAD. There is no evidence whether class-C HSFs are activators or repressors of transcription despite their lack of TAD elements.

Similar to other species, HSR in plants is characterised by the fast induction and synthesis of HSP chaperones that accumulate and prevent proteins damage caused by heat and prevent the aggregation of damaged proteins in the cells under stressful conditions (Schöffl, et al., 1998)

(Section 1.2.1). It is also transcriptionally regulated by multiple HSFs that possess distinct and overlapping roles (Wang, et al., 2004). Interestingly, some HSFs in plants are expressed in a stress-dependant manner which is a process that does not exist in any of the aforementioned non-plant species (Scharf et al., 1998). The high multiplicity of plant HSFs is thought to allow for a highly flexible and more rapid response to the various permutations of changes in surrounding conditions (Miller and Mittler, 2006; Nover et al., 2001). However, there is emerging evidence that this multiplicity could also be implicated on other non-stress related processes.

The majority of studies on plant HSFs focused primarily on their direct involvement in the regulation of stress response (Schramm et al., 2006; Mishra et al., 2002; Panchuk et al., 2002; Prändl et al., 1998; Hübel et al., 1995; Lee et al., 1995). Unlike research carried out on HSFs in *S. cerevisiae*, *D. melanogaster* and vertebrates, the roles of plant HSFs in the regulation of other non-stress related cellular processes are not well explored (Liu and Charng, 2013).

1.1.1.4.1 in *Oryza sativa*

There are 25 identified HSF coding genes in *O. sativa* (*OsHSFs*) and with their duplicates the total number of genes coding OsHSFs is 38 (Chauhan et al., 2011; Hu et al., 2009; Miller and Mittler, 2006). There are 13 members in class-A OsHSF grouped into 9 groups A1 – A9, 7 members in class-B grouped into 4 groups B1 – B4, and 4 members in class-C grouped into C1 – C2 OsHSFs (Chauhan et al., 2011). Studies on basal expression patterns of OsHSFs revealed high tissue dependency (Chauhan et al., 2011). Further expression analysis of *OsHSFs* showed a large variation in their expression patterns. Some *OsHSFs*, such as *OsHSFA2a* and *OsHSFA2d* are highly induced under heat stress in both root and shoot tissues. Others are responsive to other forms of stress but not heat such as the cold responsive *OsHSFA3*. The expression pattern of *OsHSFs* is

also affected by the developmental stages of the plant. However, *OsHSFA1* seems to be the only class-A *OsHSF* that is not inducible under any form of applied stress (Chauhan et al., 2011; Hu et al., 2009). This gives an indication that *OsHSFA1* is the only HSF in rice that is not transcriptionally regulated.

Functional Characterisation of *OsHSFs* is not well established. Only a few *OsHSFs* have been functionally characterised. *OsHSFA2e* and *OsHSFA7* were cloned and expressed in *A. thaliana*. Overexpression of *OsHSFA2e* in *A. thaliana* resulted in enhanced heat and salt tolerance compared to wild type controls (Yokotani et al., 2007). On the other hand, overexpression of *OsHSFA7* in *A .thaliana* led to increased thermotolerance (Liu, et al., 2009). *OsHSFC1b* was functionally characterised in rice and it was shown that loss of *OsHSFC1b* leads to decreased tolerance to salt and osmotic stresses and high sensitivity to ABA beside overall retardation in the growth of the plant (Schmidt et al., 2012). Overexpression of *OsHSFB2b* in rice, however, resulted in impaired tolerance to drought and salt stresses. Knockout mutant of *OshsfB2b*, on the other hand, showed an opposite phenotype which suggests that OsHSFB2b acts as a negative regulator of drought and salt stress in rice (Xiang, et al., 2013).

1.1.1.4.2 in *Solanum lycopersicum*

The tomato HSF family (SlHSF) consists of 25 members (Scharf et al., 2012). A number of members of the SlHSF family were functionally characterised. SlHSFA1 is constitutively expressed and is considered as the master regulator of HSR in *S. lycopersicum* (Miller and Mittler, 2006; Mishra et al., 2002). Loss of *SlHSFA1* in tomato resulted in plants that were unable to cope with mild heat stress treatments. Overexpression of *SlHSFA1*, on the other hand, led to enhanced thermotolerance under extreme heat stress conditions. It has also been shown that

no other SlHSF can compensate for the loss of *SlHSFA1* (Mishra et al., 2002). HSR in tomato results in the accumulation of SlHSFA2 which becomes the dominant SlHSF under prolonged heat stress condition (Charng et al., 2007; Miller and Mittler, 2006; Scharf et al., 1998). However, *SlHSFA2* is transcriptionally regulated by SlHSFA1 and the loss of *SlHSFA1* results in no expression of SlHSFA2 under heat stress (Charng et al., 2007). These results give an indication that the transcriptional regulation of HSR in tomato is organised in a true hierarchical manner.

As more of are being *SlHSFs* studied, more of the complexity started to appear in the regulation of HSR. It was shown that SlHSFB1 interacts with SlHSFA1 and adds synergy to its function as an activator of transcription (Charng et al., 2007; Scharf et al., 1998). Overexpression of *SlHSFA3* in *A. thaliana* resulted in increased thermotolerance but also had a negative implication on the plant response to salt stress upon germination. From a developmental prospective, overexpression of *SlHSFA3* resulted in late flowering time in *A. thaliana*. The same study also showed that overexpression of *SlHSFA3* resulted in an increase in the transcript levels of various HSPs in *A. thaliana* (Li, et al., 2013). The ability of the SlHSFA3 to activate HSR in *A. thaliana* adds more evidence to the high conservation of HSR among plant species.

All of the studied SlHSFs showed that they are the regulators of tomato HSR. Very few studies have addressed the possible roles of SlHSFs in the regulation of plant responses to other forms of stress (Piterková, et al., 2013). However, no studies yet have examined the possible involvement of SlHSFs in the regulation of non-stress related processes and their influence on plant growth and development.

1.1.1.4.3 in *Arabidopsis thaliana*

A. thaliana possesses 21 HSFs; similar to other plant species, the AtHSF family is divided into three major classes, A, B, and C. Each AtHSF class is sub-categorised into smaller sub-groups, 9 class-A sub-groups A1-A9 which consist of 15 members, 4 class-B sub-groups B1-B4 consisting of 5 members. There is only one class-C HSF in *A. thaliana* (AtHSFC1) (Swindell et al., 2007; Miller and Mittler, 2006; Nover et al., 2001). The basic structure of all AtHSFs is highly similar to other plant and non-plant HSFs. All class-A AtHSFs contain a TAD proximal to the C-terminal end and they have been shown to be activators of transcription (Nover et al., 2001). Class-B AtHSFs, one the other hand, do not contain a TAD and act as repressors of transcription (Miller and Mittler, 2006; Nover et al., 2001). Similar to class-B, AtHSFC1 also does not contain a TAD, however, there is no evidence whether AtHSFC1 acts as a transcriptional repressor or activator despite showing weak transcriptional activation activity in yeast (Schmidt et al., 2012; Scharf et al., 2012; Kotak et al., 2004; Miller and Mittler, 2006) (figure 1.6).

There is a high degree of variability in the basal expression patterns of *AtHSFs*. Some *AtHSFs* exhibit a degree of tissue specificity; for example, *AtHSFB4* and *AtHSFC1* are highly expressed in roots compared to other tissues (Begum et al., 2013; Swindell et al., 2007; Miller and Mittler, 2006). In contrast, the expression patterns of other *AtHSFs*, such as *AtHSFA1a*, *AtHSFA1b* and *AtHSFA2*, seem to be equal in all plant tissues and cell types (Swindell et al., 2007; Miller and Mittler, 2006). Expression profiling showed that some *AtHSFs* are constitutively expressed and their expression levels do not change in response to changes in growth conditions (Swindell et al., 2007; Miller and Mittler, 2006). On the other hand, other *AtHSFs* are expressed in a stress-dependant manner. Examples of stress-responsive AtHSFs include *AtHSFA2* which is highly induced under heat stress, and *AtHSFA6a* which its expression is responsive to salt stress

(Hwang et al., 2014; Charng et al., 2007; Schramm et al., 2006).

Unlike tomato, no AtHSF has been identified as a sole master regulator of HSR in *A. thaliana* (Miller and Mittler, 2006). Single knockouts of *AtHSFs* did not impair plant response to heat stress. Double knockouts such as *AthsfA1a/Athsfa1b* showed sensitivity to heat stress (Busch et al., 2005). Furthermore, loss of both *AthsfA1d* and *AthsfA1e* have also been shown to impair plant response to heat shock (Nishizawa-Yokoi et al., 2011) Knockout of *AthsfA2*, however, showed a decrease in plant response to only prolonged heat stress treatments (Nishizawa-Yokoi et al., 2011; Charng et al., 2007; Schramm et al., 2006). These results suggest that there is a high functional overlap among AtHSFs during HSR. Moreover, it has been shown that regulation of HSR by AtHSFs is more complex than in many of the studied plants. Some of the AtHSFs, such as AtHSFA1a and AtHSFA1b, have been shown to be involved only in the regulation of immediate and early phases of HSR (Busch et al., 2005; Lohmann et al., 2004). Other AtHSFs, such as AtHSFA2 has been shown to be involved in the regulation of late and prolonged HSR (Liu et al., 2013; Charng et al., 2007). It is still not clear yet why *A. thaliana* needs more than one AtHSF to initiate early HSR.

Recent studies have established the involvement of AtHSFs in the regulation of plant response to various forms of abiotic and biotic stresses. For example, overexpression of *AtHSFA1b* showed enhanced plant response to a number of abiotic and biotic stress forms including drought and pathogen infection (Bechtold et al., 2013). Furthermore, loss of both *AthsfA1d* and *AthsfA1e* showed decrease in the activity of photosystem II under high light stress (Nishizawa-Yokoi et al., 2011). AtHSFA4a has been shown to confer plant response to salt and oxidative

stress (Pérez-Salamó et al., 2014). Furthermore, investigation of the roles of AtHSFB2a and AtHSFB2b showed that they act as negative regulators of plant resistance to the necrotrophic fungus Alternaria brassicicola (*A. brassicicola*) by repressing the expression of the defensin genes *PDF1.2a* and *PDF1.2b* (Kumar et al., 2009).

Very few studies to date have investigated the roles of plant HSFs in the regulation of cellular processes under non-stressful conditions such as signalling and plant developmental processes. Only two studies addressed that matter; one study has shown that overexpression of *AtHSFA1b* resulted in mild developmental effect manifested in stable seed yield and harvest index under no stress and drought stress conditions in *A. thaliana* and Brassica napus (*B. napus*) (Bechtold et al., 2013). Another study showed that AtHSFB2a is involved in the regulation of gametophyte development in *A. thaliana* (Wunderlich et al., 2014).

1.1.1.4.3.1 The *Arabidopsis thaliana* group-A1 HSFs

Group-A1 AtHSFs are considered by many researchers to be the master regulators of all HSFs in *A.thaliana*. This group of AtHSFs consist of 4 members, AtHSFA1a, AtHSFA1b, AtHSFA1d, and AtHSFA1e (Swindell et al., 2007; Miller and Mittler, 2006; Nover et al., 2001). The expression of all members of group-A1 *AtHSFs* does not exhibit any tissue specificity (Swindell et al., 2007; Miller and Mittler, 2006). At least two members of group-A1 AtHSFs, *AtHSFA1a* and *AtHSFA1b*, have been shown to be constitutively active and their expression is not responsive to any form of applied stress (Swindell et al., 2007; Miller and Mittler, 2006).

Numerous studies reported the direct involvement of group-A1 AtHSFs in the regulation of *A. thaliana* responses to a number of stress conditions. Overexpression of *AtHSFA1a*, results in

constitutive activation of *AtHSPs* under no stress conditions which in turn resulted in enhanced basal thermotolerance (Qian et al., 2014). *A. thaliana* plants overexpressing *AtHSFA1a* also showed enhanced tolerance to a wide range of pH changes and to hydrogen peroxide treatment (Qian et al., 2014). In a similar manner, overexpression of AtHSFA1b results in constitutive activation of *AtHSP*s and accumulation of AtHSPs under normal growth conditions. This resulted in plant high survival rate under extreme heat stress treatments compared to wild type controls (Prändl et al., 1998). This led to the conclusion that plants overexpressing *AtHSFs* phenocopy wild type plants acclimatised to heat stress which also show enhanced tolerance to heat stress (Prändl et al., 1998). However, as described in Section (1.1.2.4.3), *AthsfA1b* knockout mutants did not result in any defects in plant response to heat stress. Only double knockout mutant *AthsfA1a/AthsfA1b* resulted in impairment in plant response to heat stress (Busch et al., 2005). In a similar manner, double knockout of *AthsfA1d/AthsfA1e* resulted in a subtle impairment of plant tolerance to heat stress and excess light stress (Nishizawa-Yokoi et al., 2011).

The results obtained from analysis of knockout mutants of group-A1 AtHSFs strongly suggest that there is a high level of redundancy among members of that group. Indeed, it has been shown that the loss of function of more members of group-A1 AtHSFs results in more sensitive plants to stress. One study by Liu et al., (2011) carried out comparisons between multiple knockout mutants of members of group-A1 *AtHSFs*. The study focused on triple knockouts of group-A1 *AtHSFs* where three members were knocked out and one left functional and on a quadruple knockout (QK) mutant where the functions of all members of group-A1 *AtHSFs* were disrupted. The results of that study showed that knockout of all members of group-A1 AtHSFs resulted in plants that were unable to intiate HSR compared to wild type and triple knockout mutants (TK). The study also showed that expression of AtHSP genes under heat stress was

extremely reduced in QK plants.

Interestingly, variable phenotypes where observed in the triple knockout mutants depending on the remaining functional AtHSF with AtHSFA1e showing the weakest heat stress tolerance in mature plants that was almost identical to QK plants. The outcome of that study strongly indicated that the four members of group-A1 AtHSFs, AtHSFA1a, AtHSFA1b, and AtHSF1d, are the true master regulators of HSR in *A. thaliana*. It also shows an evidence of the high functional redundancy among members of group-A1 AtHSFs.

Strikingly, QK mutant of all group-A1 AtHSFs showed an extremely defective developmental phenotype. The loss of function of *AthsfA1a/AthsfA1b/AthsfA1d/AthsfA1e* (QK) affected the development of the plant throughout many stages of plant life cycle. Triple knockout mutants; on the other hand, showed variable developmental phenotypes depending on which group-A1 *AtHSF* is functional with a result suggesting that AtHSFA1e has the weakest influence on plant growth and development (Liu, et al., 2011). Despite showing that group-A1 AtHSFs are not essential for the viability of the plant, the study showed clear evidence that group-A1 AtHSFs play a key role in the regulation of plant developmental processes. This result is somehow in agreement with the only two studies that suggested a possible developmental component of AtHSFs (Wunderlich et al., 2014, Bechtold et al., 2013).

1.1.1.4.3.2 Transcriptional regulation of AtHSFs by group-A1 AtHSFs

The complexity of transcriptional regulation of HSR in *A. thaliana* has led to a great interest in the transcriptional regulation of *AtHSFs* themselves. A number of *AtHSFs* are classified as constitutive transcription factors (CTFs). Those AtHSFs are mainly members of group-A1 AtHSFs,

including AtHSFA1a, AtHSFA1b, AtHSFA1d, and AtHSFA1e and those are thought to be responsible for the initiation expression of the rest of AtHSFs (Liu, et al. 2011; Swindell et al., 2007; Miller and Mittler, 2006). However, there is still some controversy about that group as some studies have shown that AtHSFA1e, for instance, might be transcriptionally regulated by another class-A AtHSF (Nishizawa-Yokoi et al., 2011; Yoshida et al., 2008, Nishizawa et al., 2006).

Microarray studies have shown that overexpression of *AtHSFA1b* leads to induction of *AtHSFA7a*, *AtHSFB1*, *AtHSFB2a*, and *AtHSFB2b* (Bechtold et al., 2013)**.** Therefore, it can be speculated that AtHSFA1b controls the expression of those *AtHSFs*. Other gene expression studies showed that loss of function of both *AtHSFA1d* and *AtHSFA1e* leads to decrease in the expression levels of *AtHSFA2*, *AtHSFA3*, *AtHSFA7b*, and *AtHSFB2a* under heat and excess light stress conditions compared to wild type plants (Nishizawa-Yokoi, et al., 2011). This suggests that these *AtHSFs* might be controlled by AtHSFA1d and/or AtHSFA1e. However, the presence of *AtHSFB2a* in both studies suggests a functional overlap between AtHSFA1b, AtHSFA1d and AtHSFA1e. This shows that the transcriptional regulation of AtHSFs in *A. thaliana* might indeed be a complex process that involves a number of AtHSFs and not only one.

More microarray analysis also revealed that overexpression of *AtHSFA3* leads to upregulation of a number of AtHSFs including *AtHSFA1e*, *AtHSFA7b*, and *AtHSFB2a* (Yoshida et al., 2008). Furthermore, overexpression of *AtHSFA2* leads to induction of *AtHSFB1* and *AtHFB2a* (Ogawa, et al., 2007). However, some of the AtHSFs do not seem to have any effect on the expression of other *AtHSFs*. For example, overexpression of *AtHSFA4a* did not result in induction of any *AtHSFs* suggesting that it has little involvement in the regulation of the expression of *AtHSFs* (Pérez-Salamó, et al., 2014). The output of those studies shows that there is a large functional

overlap between AtHSFs which in turn suggests that the transcriptional regulation of AtHSFs could more dynamic than those of other plant species such as *S. lycopersicum*. This dynamic regulation of expression program of AtHSFs in *A. thaliana* might include also swaps in roles among *AtHSFs* and positive/negative feedback loops. This also could mean that the transcriptional regulation machinery by AtHSFs in *A. thaliana* is organised in a non-linear or a dynamic hierarchical manner.

It is important to state here that the conclusions drawn from those microarray studies were mainly based on the effects of loss or gain of functions of *AtHSFs* on the expression of other AtHSFs in combination with promoter motif analysis. While this type of analyses can be informative, it is not sufficient to draw definitive conclusions about the levels of control whether it is direct or indirect. For example, despite overexpression of *AtHSFA3* showed upregulation of other *AtHSFs*, it has also shown upregulation of other TF genes that might be involved in the regulation of the expression of those *AtHSFs*. Therefore, it cannot be concluded that *AtHSFA3* is a direct regulator of those *AtHSFs* induced by its overexpression.

1.1.1.4.3.2.1 The *Arabidopsis thaliana* HSFA1b

The role of AtHSFA1b in the regulation of HSR is well established. It has been shown that overexpression of *AtHSFA1b* results in constitutive activation of a large number of AtHSP genes under normal growth conditions (Prändl et al., 1998). Furthermore, there is evidence that the high induction of AtHSP genes in plants overexpressing *AtHSFA1b* is also translated into protein levels. Tests showed that large amounts of AtHSPs accumulate in plants overexpressing *AtHSFA1b* under non-stress conditions (Lohmann et al., 2004; Panchuk et al., 2002; Schöffl et al., 1998; Prändl et al., 1998). As a result, those plants showed high resistance to heat stress and

enhanced basal thermotolerance compared to wild type plants (Panchuk et al., 2002). Only the resistance of wild type plants that were acclimatised to mild heat stress condition was comparable with the performance of *AtHSFA1b* overexpressing plants (Panchuk et al., 2002; Prändl et al., 1998). The outcome of the mentioned studies, led to the conclusion that AtHSFA1b acts as one of the universal regulators of HSR in *A. thaliana* (Prändl, et al., 1998).

Loss of *AtHSFA1b* function, however, does not seem to have significant effect on plant response to heat stress compared to wild type (Busch et al., 2005). This could be a result of the partial functional redundancy among group-A1 AtHSFs (Liu et al., 2011; Section 1.2.1.4.3.1). The heat stress sensitivity appears only when the functions of more members of group-A1 AtHSFs are abolished (Liu et al., 2011; Section 1.1.2.4.3.1). This gives clear evidence that there is a large functional overlap among members of group-A1 AtHSFs in *A. thaliana*. The output of those studies showed a clear indication that a single master regulator of heat shock response in *A. thaliana* is non-existent.

Microarray analysis on plants overexpressing *AtHSFA1b* under non-stress conditions showed that this TF leads to induction of not only AtHSP genes but a number of genes that are known to be involved in the regulation of various forms of stress including abiotic and biotic stress-responsive genes (Bechtold et al., 2013). This induction of other stress response genes resulted in enhanced tolerance against various stress forms including drought and pathogen resistance (Bechtold et al., 2013). However, the induction of *AtHSPs* by AtHSFA1b might also be a contributing factor to the enhanced abiotic and biotic stress tolerance (Section 1.2.1). This result suggests that AtHSFA1b might not only be involved in the regulation of HSR; instead it may be a regulator of general stress response in *A. thaliana*. Functional analyses of TFs have

shown that some of them can be multifunctional (Farzadfard et al., 2013; Sakuma et al., 2006; Morgan, 2006; Komori, 2002). This supports the argument that AtHSFA1b could be involved in the regulation of plant response to multiple stress forms. However, it has also been shown that overexpression of *AtHSFA1b* also results in an altered developmental effect on *A. thaliana* plants in the form of enhanced seed yield and harvest index (Bechtold, et al., 2013).

None of the genes affected by overexpression of *AtHSFA1b*, however, is known to have a developmental role especially in increasing harvest index and water productivity, despite showing that overexpression of *AtHSFA1b* leads to these developmental changes (Bechtold et al., 2013). This could mean that AtHSFA1b affects the development of the plant in an alternative pathway. The involvement of HSFs in the regulation of plant development and other essential processes under normal growth conditions is an area that is not well explored. Most of the studies have focused mainly on the direct involvments of plant HSFs in the regulation of stress response. However, as it was shown in other species, HSFs play crucial roles in the developmental processes. Plant HSFs might be involved in similar processes and the recent studies by Bechtold, et al., (2013), Wunderlich, et al., (2014) and Liu et al., (2011) showed clearly that AtHSFs influence a number of non-stress related processes in *A. thaliana,* including seed yield, harvest index and water productivity and the overall developmental process. Furthermore, the high multiplicity in plant HSFs families and their expression patterns between constitutive and inducible also suggests that there might be hidden and unexplored roles of those HSFs.

1.2 Aims and objectives of this research:

Elevated temperature is a major detriment of plants' survival and productivity. The recent outputs from climate studies showed that the earth temperature is continuing to elevate. This elevation of global temperature is expected to be accompanied with a large decline in land plants including major crop plants in a world that is already suffering a major challenge in food security manifested by the rapid increase of human population and decrease in crop production (Ray et al., 2014). A major challenge in plant science nowadays is to develop plants that are not only capable of surviving but maintaining and increasing yield and productivity under increasingly detrimental weather conditions using less available resources to achieve global food security.

The role of AtHSFA1b is well established as a major regulator of plant abiotic and biotic stress responses. However, very little is known about its involvement in the regulation of other cellular processes out of this stress context. To date, only one study showed some evidence of the potential developmental influence of AtHSFA1b under non-stress and stress conditions (Bechtold et al., 2013). Furthermore, all existing models of the transcriptional regulatory networks of plant HSFs including AtHSFA1b are based solely on transcriptomics data and prediction of direct downstream targets via promoter *cis*-elements analysis. No studies to date have attempted to map the binding sites of AtHSFA1b in the *A. thaliana* genome through appropriate experimental procedures and link these to its regulated transcriptome. Therefore, the focus of this study was to explore the molecular roles of AtHSFA1b as a key regulator of non-stress related processes including its influence on plant growth and development and to examine the possible changes that may occur in its transcriptional regulatory network during transition from normal growth conditions to stress defence. The aims of this research are

summarised in the following points

- To confirm the interaction between AtHSFA1b and the promoters of its target genes predicted by Bechtold, et al., (2013) and relate those interaction to gene expression changes of those genes.
- To map the binding profile of AtHSFA1b in the *A. thaliana* genome and examine the changes that may occur in its binding profile in response to heat stress.
- To analyse the AtHSFA1b-regulated transcriptome under non-stress conditions and heat stress and relate the possible changes in the AtHSFA1b-regulated transcriptome to changes in its binding profile in response to heat stress.
- To show evidence of the molecular involvement of AtHSFA1b in the regulation of non-stress related functions including plant growth and development.
- To identify a potential HSE variant(s) that might be the preferred binding elements for AtHSFA1b.
- To show evidence of the possible hierarchical organization in the AtHSFA1b transcriptional regulatory network and how the topology of the network might change in response to heat stress.
- To generate a binary model of the AtHSFA1b transcriptional regulatory network no stress and heat stress.

CHAPTER 2

Materials and Methods

2.1 Plant material, growth conditions and stress experiments

2.1.1 Plant materials and growth conditions:

Plants used throughout this study were *A. thaliana* transgenic plants overexpressing *AtHSFA1b* tagged modified red fluorescent protein (35S-AthHSFA1b::mRFP; Bechtold et al., 2013) and wild type Col-0 plants. It was shown in a number of studies that tagging AtHSFA1b with fluorescent and other types of proteins does not affect its normal function (Bechtold et al., 2013; Yoshida et al., 2011; Prändl et al., 1998).

Seeds were stratified at 4°C for 48 hours in complete darkness. Three seeds were place in each pot to avoid crowding of seedlings then the best growing plants were selected for further experiments. All plants were then grown in soil in a controlled environment under short day conditions 8-hour-light/16-hour-dark cycle at light intensity of 130 µmol $m^{-2} s^{-1}$, temperature 22°C and 60% relative humidity. Plants were grown until the fifth week from germination which is the point where all experiments were carried out.

2.1.2 Heat stress experiment:

Heat stress treatment was carried out on both wild type and 35S-AtHSFA1b::mRFP plants in a growth cabinet (Sanyo) at 37°C for 30 minutes under the same light intensity and under 88% relative humidity to eliminate effects due to accompanying changes in vapour pressure deficit (VPD).

2.2 Chromatin Immunoprecipitation (ChIP):

2.2.1 Sample preparation:

ChIP was carried out as described in (Saleh et al., 2008), *A. thaliana* plants overexpressing *AtHSFA1b* tagged with modified red fluorescent protein (*mRFP*) (35S-AtHSFA1b::mRFP; Bechtold et al., 2013) were grown exactly as described in Section (2.1.1). Plants were then divided into two batches, one was treated with heat stress at 37°C for 30 minutes and the other was used directly for ChIP without applying any form of stress (Sections 2.1.1 and 2.1.2). Whole rosettes from each batch were submerged in 40 ml crosslinking solution (0.4M Sucrose, 10 mM Tris-HCl pH 8, 1mM EDTA, 1% (v/v) Formaldehyde and 1 mM PMSF) in a vacuum infiltrator for 15 minutes. Crosslinking was then stopped by adding 3 ml of 2M Glycine and applying vacuum for additional 5 minutes. Crosslinked plant material was then washed three times with sterile water, blotted on paper towels then immediately flash frozen in liquid nitrogen. The plant materials were then ground in liquid nitrogen using a mortar and pestle to very fine powder and immediately processed for the next steps.

Nuclei were extracted by incubating the plant powder in 25 ml of nuclei extraction buffer (0.25M Sucrose, 15 mM PIPES pH 6.8, 5 mM $MgCl_2$, 60 mM KCl, 1 mM $CaCl_2$, 0.9% (v/v) Triton X100, 1mM PMSF, 2 µg ml^{-1} Pepstatin A and 2 µg ml^{-1} Aproteinin) for 25 minutes with occasional vortexing until complete homogenisation. The homogenised mixture was then filtered through four layers of cheesecloth. The filtrate was centrifuged at 4400 x g for 35 minutes at 4°C to pellet the nuclei. Supernatant from each sample was discarded then the nuclei pellet was resuspended in 2 ml of cold nuclei lysis buffer (50 mM HEPES pH=7.5, 150 mM NaCl, 1mM EDTA, 0.1% (w/v) SDS, 0.1% (w/v) sodium deoxycholate, 1% (v/v) Triton X100, 1 µg ml^{-1} Pepstatin A and 1 µg ml^{-1} Aproteinin). Lysed nuclei from each sample were divided into

four aliquots, 500 µl each. Extracted chromatin was sheared by sonication using a bath sonicator (BioRuptor standard B01010002, Diagenode, Belgium) into ~250bp fragments. The sonication conditions were 1 hour of sonication divided into 6 cycles, each cycle was 10 minutes of which 30 seconds sonication followed by pause for 30 seconds. Tubes were cooled on ice for 3 minutes between each sonication cycle. Sonicated chromatin was then centrifuged at 18300 x g for 10 minutes at 4°C to pellet debris. The supernatant from each sample was transferred to new tubes for further processing.

Sonicated chromatin was incubated with 5 µl of anti-RFP antibody (ab62341, Abcam, Cambridge, UK) for 3 hours at 4°C with gentle rotation. 50 µl of protein A agarose beads pre-blocked with salmon sperm sheared carrier DNA (cat. no. 16-157, Millipore, Watford, UK) were added to the antibody-chromatin mix and incubated for further 2 hours at 4°C with gentle rotation. After that samples were centrifuged at 2000 x g for 3 minutes at 4°C then supernatant was discarded and beads were collected. Protein A agarose beads were washed five times with cold low salt buffer (150 mM NaCl, 20 mM Tris-HCl pH= 8, 0.2% (w/v) SDS, 0.5% (v/v) Triton X100, and 2mM EDTA) for 5 minutes with gentle rotation followed by centrifugation at 2000 x g for 2 minutes at 4°C between each wash. Protein A agarose beads were then washed once with high salt buffer (500 mM NaCl, 20 mM Tris-HCl pH=8, 0.2% (w/v) SDS, 0.5% (v/v) Triton X100, and 2mM EDTA) then once with Lithium chloride buffer (0.25M LiCl, 10 mM Tris-HCl pH=8, 1% (v/v) Nonidet P40, and 1mM EDTA). Beads were pelleted at 2000 x g for 5 minutes at 4°C then the immune-complexes were eluted from the beads by adding 250 µl of elution buffer (0.5% (w/v) SDS and 0.1M sodium bicarbonate) then incubated at room temperature for 30 minutes with gentle rotation. Samples were centrifuged at 2000 x g for 5 minutes at room temperature and the supernatant was then transferred into new 1.5 ml eppendorf tubes. A second round of

elution was carried out by adding another 250 µl of elution buffer to protein A agarose beads followed by incubation for 15 minutes at room temperature with gentle rotation then the supernatant was collected after centrifuging the samples at 2000 x g for 5 minutes. The two elutes from each sample were combined and reverse crosslinked by adding 5 µl of 5M NaCl followed by incubation at 65°C overnight. Proteins were digested by adding 1 µl of 2 µg ml^{-1} Proteinase K (Fisher Scientific, Loughborough, UK) to each sample followed by incubation at 42°C for 2 hours.

DNA fragments were extracted by adding 500 µl of water saturated Phenol (Fisher Scientific, Loughborough, UK) to each sample followed by vortexing for 30 seconds then samples were centrifuged at 18400 x g for 10 minutes. Upper aqueous phase from each sample was transferred into new eppendorf tubes. 200 µl of Chloroform was added to each sample and samples were vortexed briefly then centrifuged at 18400 x g for 10 minutes. The supernatant from each sample was transferred into new tubes and the DNA was precipitated by adding 2.5 volumes of absolute ethanol, 0.1 volume of sodium acetate and 3 µl of 20 mg ml^{-1} glycogen (Fisher Scientific, Loughborough, UK) then incubated on dry ice for 2 hours. Negative controls were samples from wild type plants treated exactly the same way as 35S-AtHSFA1b::mRFP plants with anti-RFP antibody. Positive controls, on the other hand, were extracted DNA from sheared chromatin without immunoprecipitation.

2.2.2 ChIP-PCR

PCR analyses were carried out on the three samples, positive control to confirm that the target region is amplifiable and not affected by shearing, negative control to check for possible artefacts and immunoprecipitated samples to validate the predicted interactions. Promoter

specific primers that amplify regions between 150 – 250bp around the predicted binding sites of AtHSFA1b were used in each PCR (see appendix A). All PCR reactions were carried out at a total volume of 25 µl consisting of 0.5 µM forward primer, 0.5 µM reverse primer, 0.2 mM dNTPs, 2mM $MgCl_2$, 1X Taq buffer (750 mM Tris-HCl pH 8, 200 mM (NH4)2SO4, 0.1% (v/v) Tween 20), 1 unit of Taq polymerase and 2.5 µl from each sample were used as a PCR templates and the PCR condition were, 95°C for 5 minutes, 40 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, followed by a final extension step at 72°C for 7 minutes. PCR products were then subjected to electrophoresis (105V for 45 minutes) on a 1.5% TAE agarose gel stained with ethidium bromide and visualised under UV light.

2.2.3 ChIP-SEQ

2.2.3.1 Sample preparation:

ChIP-SEQ was carried out as described in Section (2.2.1) with a few technical adjustments. Protein A/Salmon sperm beads which are blocked with salmon sperm DNA can interfere with the sequencing results (O'Green et al., 2010). Therefore, they were replaced by unblocked Protein A Sepharose beads (Sigma-Aldrich, Dorset, UK). 250 mg of lyophilised protein A Sepharose beads were swelled in 2 ml of sterile deionised water at 4°C for 4 hours. Beads were then pelleted at 2000 x g for 5 minutes at 4°C then washed three times with sterile deionised water. The beads were centrifuged again at 2000 x g for 5 minutes at 4°C then water was discarded from the beads. A 50% (v/v) protein A Sepharose beads slurry was made by adding 2 ml of storage buffer (10 mM Tris-HCl pH 8, 1 mM EDTA and 0.1% (w/v) sodium azide). Protein A Sepharose beads were blocked using ChIP lysis buffer containing 200 µg ml^{-1} BSA, 200 µg ml^{-1} Glycogen and 200 µg ml^{-1} of yeast t-RNA). Beads were blocked for 2 hours at 4°C with gentle rotation. Biological duplicates from each treatment and control samples were processed and

the samples were validated with PCR using primers that showed the best amplifications in the ChIP-PCR experiment (Section 2.2.2). Library construction and sequencing were carried out at The Genome Analysis Centre (TGAC; http://www.tgac.ac.uk/) (Norwich, UK) using TruSeq ChIP-SEQ library preparation kit (Illumina, San Diego, CA, USA). Eight ChIP-SEQ samples (2 x 35S-AtHSFA1b::mRFP no stress, 2 x 35S-AtHSFA1b::mRFP heat stress, 2 x Control no stress, 2 x Control heat stress) were multiplexed on one lane in a HiSeq2000 platform (Illumina, San Diego, CA, USA). The sequencing was carried out following the modENCODE requirements for ChIP-SEQ experiments, 100bp paired-end with minimum of 10 million reads sequencing depth (Landt, et al., 2012).

2.2.3.2 Data analysis

2.2.3.2.1 Processing ChIP-SEQ reads:

Raw short sequencing reads with their quality scores from all samples were obtained in FASTQ format. FASTQ files were quality checked and all adapter/primer contaminations from TruSeq ChIP-SEQ library preparation kit and poor quality sequences were filtered out using the programme FASTQ trimmer. Short reads were aligned and mapped to the *A. thaliana* genome using GSNAP (Wu and Nacu, 2010). Mapped reads files were checked for percentage of alignment and counts then sorted and indexed using SAMtools (Li et al., 2009) and converted into a compressed BAM format and all duplicated PCR reads were removed to generate the final sorted BAM file. ChIP-SEQ peak calling was carried out using two programs, MACS v2 (Zhang et al., 2010) and CisGenome v2 (Ji et al., 2008) were used to compare the outputs from the two programs.

2.2.3.2.2 Peak calling procedure:

In the case of MACS v2, the peaks were called directly from the sorted BAM files. Peaks were called from each replicate individually against a pooled control reads BAM file. MACS v2 generated a peak file for each individual replicate from each sample which were then used to calculate the Irreproducibility Discovery Rate (IDR) (Li et al., 2011) using the R script 'batch-consistency-analysis.r'. Three cutoffs were used in MACS v2, IDR<0.01, *p*-value < 0.0005, *q*-value < 0.01. In the case of CisGenome, BAM files were converted to alignment files with the extension (*.ALN). The alignment files were used directly in CisGenome v2 for peak calling. The cutoff used in CisGenome v2 was FDR < 0.05. Peak files were generated in COD format showing the coordinates of the peaks on the *A. thaliana* genome in addition to the score of each peak and FDR values. Peaks from the two programs were then annotated using the R package ChIPpeakAnno (Zhu et al., 2010) based on the positions of peaks relative to the nearest annotated transcriptional start site (TSS) of *A. thaliana* genes and two more annotated peak files were generated based on the position of the peaks on the 5' and 3' untranslated regions (UTRs). The annotated files were then analysed for positional enrichment of the peaks using ChIPpeakAnno package to determine the highest percentages of binding sites relative to the distance to the nearest TSS and the overall distribution of binding sites relative to the nearest genomic features. Gene ontology analysis on the final annotated ChIP-SEQ peaks was carried out using AgriGO database (http://bioinfo.cau.edu.cn/agriGO/) (Du et al., 2010) using Hypergeometric test and a Bonferroni adjusted *p*-value cutoff < 0.01.

2.2.3.2.3 Motif analysis:

ChIP-SEQ peaks were converted from coordinates on the *A. thaliana* genome in BED format to the corresponding sequences in FASTA format using the program BEDTools (Quinlan and Hall,

2009) in two ways. The first way was by extracting single nucleotides which represent the summit of ChIP-SEQ peaks in both datasets then the sequences were extended 50bp from each side of the tip of the peak. The second was by extracting the entire sequences of ChIP-SEQ peaks. Both files, the 50bp extended peak sequences and the whole peak sequences from each treatment, were used in *de novo* motif discovery using three programs, STEME (Reid et al., 2011), rGADEM (Mercier et al., 2011) and, MEME (Bailey et al., 2009). The cutoff used in all program was *p*-value < 0.0001. The outputs from all programs were evaluated and compared. The data were reanalysed using pattern matching from both experiments to validate the structures and the overall enrichment of the motifs discovered.

2.3 Quantitative real-time PCR analysis

2.3.1 RNA extraction:

35S-AtHSFA1b::mRFP plants and Col-0 wild type plants were grown as described in Section (2.1.1). Plants were divided into two batches, the first batch; RNA was extracted from three biological replicates of 35S-AtHSFA1b::mRFP plants and Col-0 plants without applying any form of stress. The second batch was three biological replicates of 35S-AtHSFA1b::mRFP plants as well as three biological of Col-0 plants were subject to short duration of heat stress treatment at 37°C for 30 minutes (Section 2.1.2). Plants were harvested then immediately flash frozen in liquid nitrogen. Plant material was ground in liquid nitrogen to fine powder using mortar and pestle. The RNA extraction was carried out according to the TRI-reagent® (Life technologies Inc., USA) instruction manual with the following modifications. Each sample was homogenised in 1 ml of TRI-reagent, vortexed briefly then kept at room temperature for 2 minutes. 200 µl of Chloroform was added to the mixture and vortexed for 30 seconds. The samples were then centrifuged at 18400 x g for 20 minutes at 4°C. 500 µl of the upper aqueous phase from each

sample was transferred into new tubes then nucleic acids were precipitated by adding one volume of isopropanol followed by incubation at room temperature for 10 minutes. Samples were then centrifuged at 18400 x g for 25 minutes at 4°C to pellet nucleic acids. Nucleic acid pellets were washed twice with ice cold 75% (v/v) ethanol, air dried for 10 minutes then resuspended in 30 µl of sterile nuclease-free water.

Genomic DNA in each sample was digested using DNA-free™ kit (Ambion Inc.) as described in the manufacturer's manual by adding 0.1 volume of 10X DNase buffer and 2 units of DNase I enzyme to each sample followed by incubating the samples at 37°C for 30 minutes in a water bath. The DNase I reaction was stopped by adding 1 volume of deactivation buffer followed by mixing and incubating the samples for 3 minutes at room temperature. Samples were then centrifuged at 18400 x g for 2 minutes then the supernatant was transferred into new tubes. RNA was re-precipitated by adding 2.5 volumes of absolute ethanol and 0.25 volume of 10 M ammonium acetate followed by incubation on dry ice for 1 hour. The RNA was then pelleted by centrifugation at 18400 x g for 25 minutes at 4°C. RNA pellet was washed three times with ice cold 75% ethanol, air dried for 10 minutes at room temperature then resuspended in 30 µl of sterile nuclease-free water. RNA from each sample was quantified using Nanodrop 1000 (Thermo Fisher Scientific Inc.) and quality checked on 1.5% (w/v) TAE agarose gel.

2.3.2 cDNA synthesis:

cDNA was synthesised using Reverse transcriptase kit (Thermo Scientific Inc.) according to the manufacturers guide. 3 µg of RNA template was mixed with 1 µl of 100 µM random hexamer primers (Thermo Fisher Scientific Inc.) and the volume was completed to 12 µl with sterile nuclease-free water. The mixture was incubated at 65°C for 10 minutes then immediately

incubated on ice for another 10 minutes. Reverse transcriptase reaction was carried out in 20 µl reaction volume containing 1X reverse transcriptase reaction buffer, 0.5 mM dNTPs, and 10 units of reverse transcriptase enzyme in addition to the 12 µl RNA-random hexamer mix. The reverse transcriptase reaction was incubated in a thermal cycler and the reaction conditions were 42°C for 1 hour followed by incubation at 72°C for 10 minutes then cooling at 12°C for 10 minutes. The cDNAs from all samples were then stored at -20°C.

2.3.3 qRT-PCR analysis:

Bioline SensiFAST™ SYBR No-ROX Kit was used for all qRT-PCR reactions. The qRT-PCR reactions were carried out as described in the instruction manual with a few modifications to the protocol. 1µl of primers was used instead of 0.8 µl and 1 µl of cDNA was used in the final qRT-PCR mix. The qRT-PCR reactions were carried out in 20 µl volumes consisting of 1X SensiMix™ SYBR® No-ROX reaction mix, 0.5 µM forward primer, 0.5 µM reverse primer, 1 µl of cDNA template and the volume was completed to 20 µl using sterile nuclease-free water. The qRT-PCR reactions were carried out in Biorad CFX96 thermal cycler (Biorad Laboratories, Inc., USA) following a three-step program, 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds, 60°C for 15 seconds, and 72°C for 15 seconds. The qRT-PCR data were analysed using excel spreadsheet template.

2.4 RNA sequencing (RNA-SEQ)

2.4.1 Sample preparation:

Growth conditions, stress treatment and total RNA extraction from 35S-AtHSFA1b::mRFP plants and Col-0 wild type plants were performed as described in Sections (2.1.1, 2.1.2 and 2.3.1). Three biological replicates were generated from 35S-AtHSFA1b::mRFP plants and wild type

plants under the same conditions as described in Sections (2.1.1 and 2.1.2). cDNA synthesis and library construction were carried out using Illumina TruSeq RNA library preparation kit (Illumina, San Diego, CA, USA) at The Genome Analysis Centre (TGAC; http://www.tgac.ac.uk/) (Norwich, UK). A total number of 12 RNA samples (3 x 35S-AtHSFA1b::mRFP no stress, 3 x 35S-AtHSFA1b::mRFP heat stress, 3 x Col-0 no stress and 3 x Col-0 heat stress) were multiplexed on one HiSeq2000 sequencing platform lane (Illumina, San Diego, CA, USA). The sequencing was carried out according to the modENCODE requirements for RNA-SEQ expression analysis with a minimum sequencing depth 50 million paired-end reads of 50bp length.

2.4.2 Data analysis:

2.4.2.1 Processing RNA-SEQ short reads:

Raw short reads files were obtained in FASTQ format. The FASTQ files were quality checked and all TruSeq primer/adapter contaminations were removed from all sample files using FASTQ trimmer tool. Short read files were then fed into TopHat v2.0.8b program (Trapnell, et al., 2012), which is based on Bowtie v2.1.0 aligner (Langmead and Salzberg, 2012). Short reads were extended to 200bp then mapped and aligned against the *A. thaliana* genome. Absolute FPKM (Fragments Per Kilobase per Exon model) expression values were generated using Cufflinks v2.1.1 (Trapnell et al., 2012). The genomic features were downloaded from Ensembl (Flicek et al., 2012) for *A. thaliana* using TAIR10 (Lamesch et al., 2011). The mean FPKM values for each gene were calculated from the three replicates in each sample. Differential expression was determined using Cufflinks sub-programme 'Cuffdiff'. Further analysis of the RNA-SEQ data and visualisation were carried out using the R package cummeRbund (http://compbio.mit.edu/cummeRbund/). The cutoffs used for significantly expressed genes were p-value < 0.0005 and q-value < 0.05 and fold change cutoff for differentially expressed

genes was 2. Gene ontology analysis on the RNA-SEQ final lists was carried out using AgriGO database (http://bioinfo.cau.edu.cn/agriGO/) (Du, et al., 2010) using Hypergeometic test and Benjamini-Hochberg FDR cutoff < 0.01.

2.4.2.2 Motif analysis:

Promoter regions of all upregulated genes from 35S-AtHSFA1b::mRFP plants under no stress and heat stressed wild type and 35S-AtHSFA1b::mRFP were used for motif analysis. 500bp upstream sequences were retrieved from the *A. thaliana* database (http://www.arabidopsis.org) using the Bulk data retrieval tool. All sequences were analysed using the same programs used in ChIP-SEQ motif analysis (MEME, STEME and rGADEM). The sequences were analysed for the presence of potential heat shock *cis*-acting regulatory elements (HSEs) as well as other promoter *cis*-acting regulatory elements with a cutoff *p*-value < 0.0001.

2.4.2.2.1 High resolution motif discovery:

Annotated ChIP-SEQ peaks were intersected with all upregulated genes in the RNA-SEQ dataset using common AGI codes. The overlapping genes were isolated and instead of examining the entire upstream regions of upregulated genes, the coordinates that represent the binding sites of AtHSFA1b from ChIP-SEQ were converted into a BED file. The coordinates were then converted into sequences files in FASTA format using the program 'BEDTools getfasta'. Peak sequences that represent the binding sequences of AtHSFA1b on upregulated genes were only analysed using the same programs used in other motif analyses (Sections 2.2.3.1.3 and 2.3.4.2.2). All motifs were saved for further analysis and comparisons. In order to enhance the resolution of HSE discovered, only the summits of the AtHSFA1b peaks on upregulated genes

were isolated and converted into a coordinate file using the method and programs described above. Peak summits are single nucleotides that represent the tip of AtHSFA1b ChIP-SEQ peaks. The peaks summits BED file was converted to a sequence file in FASTA format using the program BEDTools and the reads were extended from the peaks summits by 20bp, 30bp, 50bp, 100bp on both sides of the peaks summits. The extended sequences were then analysed using the same motif discovery programs used in the other previous motif analyses (Sections 2.2.3.1.3 and 2.3.4.2.2). Similar to all motif analyses the statistical cuttoff was *p*-value < 0.0001.

2.5 Yeast one-hybrid and functional analysis of AtHSFA1b in yeast:

2.5.1 Yeast one-hybrid:

2.5.1.1 PCR amplification of promoter fragments

The promoter regions of the genes *AtBAG6*, *AtHSFB1*, and *AtHSP90.1* were selected for Yeast one-hybrid (Y1H) assay. 1000bp upstream regions of the selected genes were PCR amplified into three overlapping promoter fragments in the size ranges of 300 to 400bp from *A. thaliana* Col-0 genomic DNA using Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific Inc.). The PCR reactions were carried out in 50 µl volumes consisting of 1X Phusion polymerase buffer, 200 µM dNTPs, 0.5 µM forward primer, 0.5 µM reverse primer, 0.1 unit of Phusion polymerase enzyme, 1 µl of *A. thaliana* genomic DNA template and the volume was completed to 50 µl using sterile water. The PCR reaction conditions were, 98°C for 30 seconds followed by 35 cycles of 98°C for 5 seconds, 60°C for 15 seconds, 72°C for 10 seconds then a final extension step 72°C for 5 minutes. An extra promoter fragment from the gene *α/β-hydrolase* was PCR amplified from the transcriptional start site to -200bp region to examine the interaction between AtHSFA1b and the promoter element HSE1b. The primers used were designed with *Eco*RI restriction site at the start of forward primers and *Mlu*I site at the end of reverse primers (see

appendix A).

2.5.1.1.1 PCR products clean-up:

All PCR products were cleaned up and purified from PCR reaction components by adding 1 volume of binding buffer (4.2M guanine hydrochloride dissolved in isopropanol) and mixed thoroughly by vortexing. The mixtures were then purified using plasmid miniprep spin columns (Thermo scientific Inc.) as described in the kit's manual.

2.5.1.2 Generating bait constructs:

2.5.1.2.1 Cloning promoter fragments into pHIS3LEU2:

Purified PCR products and the bait plasmid pHIS3LEU2 were sequentially digested using the enzymes *EcoR*I and *Mlu*I. The first digestion was using the enzyme *EcoR*I and it was carried out it a 20 µl reaction volume containing 1X of *EcoR*I unique buffer, 10 units of *EcoR*I enzyme, 2 µg of DNA then the volume was completed to 20 µl using sterile water followed by incubation at 37°C for an overnight. The digested PCR products and the bait vector were then cleaned up and purified as described in Section (2.5.1.1.1). The second restriction digestion round was also carried out in 20 µl reaction volume containing 1X of *Mlu*I appropriate buffer, 20 units of *Mlu*I enzyme, 1 µg of DNA then the volume was brought to 20 µl using sterile water. The restriction digestion reaction was incubated at 37°C in a water bath for an overnight. The digested PCR products and the bait vector were then purified as described in (2.5.1.1.1).

Each PCR product was then ligated into the linearised bait vector pHIS3LEU2 using T4 DNA ligase kit (Thermo Scientific Inc.). The ligation reaction was carried out in a 20 µl reaction volume consisting of 50 ng of linear pHIS3LEU2 vector, 1:1 molar ratio of PCR product over

vector, 1X of T4 DNA ligase buffer, 1 unit of T4 DNA ligase enzyme and the volume was completed to 20 µl using sterile water. The reaction was incubated at 22°C in a thermal cycler for 2 hours. The ligation mix was then transformed into One Shot® OmniMAX™ 2 T1^{R} *E. coli* cells (Life technologies, USA) using the heat shock method: 2 µl of the ligation reactions were added to 50 µl of cells and incubated on ice for 20 minutes. The ligation-cells mixture was then heat shocked at 42°C for 30 seconds in a water bath then immediately incubated on ice for 5 minutes. 200 µl of Luria-Bertani (LB) broth was added to the cells followed by incubation at 37°C for 1 hour in an orbital shaker. Cells were then plated on LB-agar media containing 50 µg ml^{-1} of kanamycin.

2.5.1.2.1.1 Confirming the presence of the promoter fragments in pHIS3LEU2:

Transfomants were verified for the presence of the promoter fragments inserts using two methods. The first was by colony PCR using Taq polymerase (Thermo scientific Inc.) and insert specific primers (appendix A). Each colony PCR was carried out in a 20 µl volume as described in Section (2.5.1.1). The PCR reaction conditions were 95°C for 5 minutes followed by 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute, then followed by a final extension step 72°C for 7 minutes. Positive transfomants were then inoculated in 5 ml of LB broth containing 50 µg ml^{-1} of kanamycin and incubate overnight in an orbital shaker at 37°C with shaking at 250 RPM. Bacterial cells were pelleted at 8000 x g for 5 minutes and the plasmid DNA was purified using GeneJET plasmid miniprep kit (Thermo scientific Inc.) as described in the instruction manual. Purified plasmids were then checked again for the presence of the inserts using restriction digestion with the enzyme *Pvu*II (Thermo scientific Inc.). The reaction was carried out in a total volume of 20 µl containing 1X of *Pvu*II appropriate buffer, 5 units of *Pvu*II restriction enzyme, 0.5 µg of DNA and the volume was completed to 20 µl using sterile

water. The restriction digestion reaction was incubated at 37°C for 1 hour in a water bath. The digestion was then visualised on 1.5% (w/v) TAE agarose gel. Promoter fragments orientation and sequences in the bait vector were then confirmed by Sanger sequencing at GATC biotech (http://www.gatc-biotech.com/en/index.html).

2.5.1.3 Yeast media

2.5.1.3.1 YPDA media:

The rich media yeast extract, peptone, glucose, adenine (YPDA) consisted of 20 g L^{-1} peptone, 10 g L^{-1} yeast extract, 20 g L^{-1} agar (for plates only) and 100 mg L^{-1} adenine hemisulfate then the volume was completed to 950 ml using water. The media was autoclaved then 50 ml of 40% (w/v) glucose solution was added to the media using syringe and filter to a final concentration of glucose to 2% (v/v).

2.5.1.3.2 Synthetic dropout media:

Synthetic dropout (SD) media were prepared containing 6.7 g L^{-1} yeast nitrogen base without amino acids, 20 g L^{-1} agar (for plates only) then the volume was brought to 850 ml with water then the mixture was autoclaved. Dropout solution is a solution that contains all the essential amino acids for yeast excluding the amino acids used for selection. 10X dropout solution for Y1H experiment was prepared by combining the following amino acids:

200 mg L^{-1} Arginine monohydrochloride

300 mg L^{-1} Isoleucine

300 mg L^{-1} Lysine monohydrochloride

200 mg L^{-1} Methionine

500 mg L^{-1} Phenylalanine

2000 mg L^{-1} Threonine

300 mg L^{-1} Tyrosine

200 mg L^{-1} Uracil

1500 mg L^{-1} Valine

200 mg L^{-1} Histidine monohydrochloride monohydrate

200 mg L^{-1} Tryptophan

Then the volume was brought to 1000 ml with sterile water. 50 ml of the dropout solution, 50 ml of 40% (w/v) glucose and 50 ml of 10X adenine hemisulfate were added to the SD media using syringes and filters brining the final concentrations to 1X of dropout media, 2% (v/v) of glucose and 1X of adenine hemisulfate.

2.5.1.4 Yeast transformation:

The bait constructs were transformed into the yeast strain Y187 (*MATα, ura3-52, his3-200, ade2-101, trp1-901,leu2-3, 112, gal4Δ, met–, gal80Δ, MEL1,URA3::GAL1UAS -GAL1TATA-lacZ*) (Harper, et al., 1993) using the yeast lithium acetate transformation method. The yeast strain Y187 was grown overnight in 10 ml YPDA liquid media in an orbital shaker at 28°C shaking at 200 RPM. Yeast cells from 1 ml aliquot were harvested by centrifugation at 2000 x g for 5 minutes. The media was discarded and the pelleted cells were washed twice with 1 ml of sterile and freshly made 100 mM lithium acetate solution with a centrifugation step at 2000 x g between each washing step. Cells were then resuspended in 1 ml of 100 mM lithium acetate and incubated at 30°C in a water bath for 1 hour. The DNA mix was prepared consisting of 1 µg of plasmid DNA, 50 µg of single-stranded salmon sperm carried DNA (Sigma-Aldrich), and 290 µl of 50% (w/v) PEG4000 (Sigma-Aldrich) and preheated at 30°C. 120 µl of the yeast cells

suspension was added to each DNA mix followed by gentle mixing using a sterile pipette tip then the DNA-cells mix was incubated at 30°C in a water bath for one hour. The DNA-cells mix was then heat shocked in a water bath at 42°C for 25 minutes. Cells were then pelleted by centrifugation at 2000 x g for 10 minutes and the supernatant was discarded. Cells were then resuspended in 120 µl of sterile water then plated on SD-L solid media and incubated at 28°C for 2 days until colonies formed.

2.5.1.5 Yeast one-hybrid screening

Yeast one-hybrid assay was carried out in the School of Life Sciences at the University of Warwick (Warwick, UK). The promoter fragments of the genes *AtBAG6*, *AtHSFB*1, and *AtHSP90.1* were screened against a library of 1500 *A. thaliana* TFs cloned into pDEST22 (Life Technologies, USA) and transformed into the yeast strain AH109 (*MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL1UAS-GAL1TATA-HIS3, GAL2UAS-GAL2TATA-ADE2, URA3::MEL1UAS-MEL1 TATA-lacZ*) (Mitsuda et al., 2010; James, et al., 1996). The yeast strain Y187 carrying the constructs containing the promoter fragments was grown in SD-L at 28°C for 48 hours. Colonies from each plate were inoculated and grown in liquid SD-L for and overnight at 28°C. At the same time, two pools of the yeast strain AH109 containing the Arabidopsis transcription factors library (T1B and T2B) were grown from glycerol stocks in liquid SD-media lacking Tryptophan (SD-W).

3 µl from each promoter strain was spotted on two solid onto solid YPDA mediain the form 96 spots. The cell spots were kept at room temperature to dry then the AH109 pools carrying the TFs library were spotted on top of the Y187 spots then kept for a further 10 minutes to dry. Plates were then incubated at 30°C for 24 hours. Diploid yeast cells from YPDA plates were

transferred onto a sterile velvet cloth. The transferred colonies were then printed from the velvet cloth to SD-media plates without Leucine and Tryptophan (SD-LW) and SD-media plates without Leucine, Tryptophan and Histidine (SD-LWH) then plates were incubated at 30°C for four days. To prevent auto-activation of *HIS3* reporter genes, three more sets of SD-LWH plates containing 3-amino-1,2,4-trizol (3AT), which is an inhibitor of the *HIS3* gene product (imidazoleglycerol-phosphate dehydratase) were prepared. Each set of SD-LWH plates contained a different concentration of 3AT, set 1 contained 25 mM of 3AT, set 2 contained 50 mM of 3AT, and set 3 contained 100 mM of 3AT. Yeast colonies from each SD-LWH plate were transferred onto a velvet cloth then printed on the SD-LWH plates containing 3AT. Colonies that tolerated 3AT were then streaked on fresh SD-LWH for further analysis.

All positive yeast colonies streaked on plates were grown at 28°C for three days. Half colony from each plate was picked using sterile pipette tip and placed into a PCR tube containing 10 µl of 20 mM NaOH. The mix was then heated at 99°C for 10 minutes and used as template for PCR analysis. PCR was carried out as described in Section (2.5.1.2.1.1) using the pDEST22 vector specific primers SABR1 and SABR2 (see appendix D) flanking the coding sequence of the TFs. PCR reaction conditions were 95°C for 5 minutes, 55 cycles of 95°C for 30 seconds, 55°C for 30 second, 72°C for 4 minutes, followed by a final extension step 72°C for 10 minutes. The PCR products were then visualised on 1.5% agarose gel to confirm results. After that, the PCR products were purified using plasmid miniprep spin columns then sequenced using Sanger sequencing method at GATC biotech (http://www.gatc-biotech.com/en/index.html) to identify the TFs that interact with the promoter fragments tested.

2.5.2 Functional analysis of AtHSFA1b in yeast

2.5.2.1 PCR amplification of AtHSFA1b coding sequence:

AtHSFA1b coding sequences was PCR amplified from wild type *A. thaliana* Col-0 cDNA using Phusion® Hot start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific Inc.). The PCR reaction were carried out in a volumes of 50 µl consisting of 1X Phusion polymerase buffer, 200 µM dNTPs, 0.5 µM forward primer including the 4 nucleotide gateway sequence, 0.5 µM reverse primer without stop codon, 0.1 unit of Phusion polymerase enzyme, 1 µl of *A. thaliana* col-0 cDNA template and the volume was completed to 50 µl using sterile water. The PCR reaction conditions were 98°C for 30 seconds, followed by 35 cycles of 98°C for 15 seconds, 60°C for 30 seconds, 72°C for 50 seconds, followed by a final extension step 72°C for 7 minutes. PCR cleaned-up and purified as described in Section (2.5.1.1.1).

2.5.2.2 Generating *AtHSFA1b* yeast expression clone

2.5.2.2.1 Cloning *AtHSFA1b* into pENTR D/TOPO gateway plasmid:

The purified PCR product was cloned into Gateway® entry vector pENTR using pENTR™/D-TOPO® cloning kit (Life Technologies Inc.). The cloning reaction was carried out in 6 µl volume consisting of 20 ng of PCR product, 1 µl of salt solution, 1 µl of TOPO vector and the volume was brought to 6 µl with sterile water. The reaction was incubated at 25°C in a thermal cycler for 2 hours. The DNA mix was then transformed into One Shot® OmniMAX™ 2 $T1^R$ *E. coli* cells (Life technologies, USA) using the heat shock method: 2 µl of the ligation reactions were added to 50 µl of cells and incubated on ice for 20 minutes. The ligation-cells mixture was then heat shocked at 42°C for 30 seconds in a water bath then immediately incubated on ice for 5 minutes. 200 µl of LB broth was added to the cells followed by incubation at 37°C for 1 hour in an orbital shaker. Cells were then plated on LB-agar media containing 50 µg ml^{-1} of kanamycin.

Transfomants were verified for the presence of the *AtHSFA1b* coding sequence inserts using colony PCR using Taq polymerase (Thermo Scientific Inc.) and insert specific primers (appendix A). Each colony PCR was carried out in a 20 µl volume as described in Section (2.5.1.2.1.1). The PCR reaction conditions were 95°C for 5 minutes followed by 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 2 minutes, then followed by a final extension step at 72°C for 10 minutes. Positive transfomants were then inoculated in 5 ml of LB broth containing 50 µg ml^{-1} of kanamycin and incubate overnight in an orbital shaker at 37°C with shaking at 250 RPM. Bacterial cells were harvested at 8000 x g for 5 minutes and the plasmid DNA was purified using GeneJET plasmid miniprep kit (Thermo scientific Inc.). The *AtHSFA1b* coding sequence and orientation in in pENTR D/TOPO were then verified by Sanger sequencing at GATC biotech (http://www.gatc-biotech.com/en/index.html).

2.5.2.2.2 Subcloning *AtHSFA1b* into the yeast expression vector pAG424GPD-ccdB-eYFP:

The *AtHSFA1b* coding sequences was transferred from pENTR D/TOPO plasmid into the yeast expression vector pAG424GPD-ccdB-eYFP (Addgene, USA) using using Gateway® LR Clonase® II enzyme mix (Life Technologies Inc.). The subcloning reaction was carried out 8 µl consisting of 100 ng of pENTR D/TOPT containing the *AtHSFA1b* coding sequence, 15 ng of pAG424GPD-ccdB-eYFP, 1 µl of LR clonase II enzyme mix and the volume was completed to 8 µl using TE buffer pH=8.0. The reaction was incubated at 25°C in a thermal cycler for 2 hours and 1 µl of 2 µg $µl^{-1}$ Proteinase K was added to the mix and incubated at 37°C for 10 minutes in a thermal cycler. The DNA mix was then transformed into One Shot® OmniMAX™ 2 $T1^{R}$ *E. coli* cells (Life Technologies Inc.) using the heat shock method: 2 µl of the ligation reactions were added to 50 µl of cells and incubated on ice for 20 minutes. The ligation-cells mixture was then heat shocked at 42°C for 30 seconds in a water bath then immediately incubated on ice for 5 minutes. 200 µl of LB media

was added to the cells followed by incubation at 37°C for 1 hour in an orbital shaker. Cells were then plated on LB-agar media containing 100 µg ml^{-1} carbenicillin. Verification of positive transformants and plasmid purification were carried out as described in Section (2.5.1.2.1.1) with the exception that the LB media in this case were containing 100 µg ml^{-1} carbenicillin. An expression construct containing *yHSF* coding sequence was also generated to be used as a control. The generation of the expression vector harbouring the *yHSF* coding sequence was carried out exactly as described above.

The expression constructs were then transformed into the yeast strain PS145 (Liu et al., 1997) (*ade2-1 trp1, can1-100, leu2-3,-112, his3-11,-15, ura3, Δhsf::LEU2, YCpGAL1-yHSF*) using the yeast lithium acetate transformation method as described in Section (2.5.1.4). Transformed yeast cells were plated on solid SD-media without tryptophan and containing D-glucose (Dextrose) (SDDEX-W) as a carbon source then were grown at 28°C for 3 days.

CHAPTER 3

Preliminary Analysis of Plants overexpressing *AtHSFA1b*

3.1 Introduction:

Unlike plants grown in controlled conditions, field grown plants suffering from heat stress are normally subject to other forms of abiotic stresses combined such as dehydration and osmotic stress (Mittler, 2005). Furthermore, there is also evidence that abiotic stress is generally associated with a decrease in plant defence against biotic stress (Wang, et al., 2009; Noel, et al., 2007). However, plant stress responses also come at the expense of other crucial processes. It has been shown in a number of studies that plant stress defence is generally accompanied with a degree of inhibition in plant growth and development (Pajerowska-Mukhtar, 2012; Tian et al., 2003). Therefore, plants switch from normal growth conditions to stress response is a complex process that involves crosstalk between various factors and allocation of resources.

Numerous studies have established the role of AtHSFA1b as a major regulator of HSR in *A. thaliana* plants grown in controlled environments. However, very little has been done to study the role AtHSFA1b in the regulation of other forms of stress (Section 1.2.1.4.3.1). A recent study by Bechtold et al., (2013) addressed that matter and examined the role of AtHSFA1b in the regulation of plant response to various forms of stress other than heat. The study has shown that overexpression of *AtHSFA1b* results in enhanced tolerance to drought stress and pathogen infections in *A. thaliana* compared to both wild type plants and *AthsfA1a/AthsfA1b* double knockout plants. Furthermore, the study found that overexpression of *AtHSFA1b* also resulted in a developmental effect on plants. *A. thaliana* plants overexpressing *AtHSFA1b* showed increased harvest index and seed yield (Bechtold et al., 2013). A similar yield phenotype was observed when *AtHSFA1b* was overexpressed in *Brassica napus* (oilseed rape). This result showed clues that besides being a regulator of plant stress response; AtHSFA1b might also have a developmental component.

The study also included microarray analysis on plants overexpressing *AtHSFA1b*. The microarray part of the study showed that overexpression of *AtHSFA1b* resulted in altering expression of 509 genes, 352 of which were upregulated and 157 downregulated. The majority of the genes affected by overexpression *AtHSFA1b* were mainly genes involved in plant response to abiotic and biotic stress. However, none of the genes altered in expression are known to be involved in regulation of developmental processes such as genes that might affect harvest index and seed yield. Motif analysis showed enrichment for a variant of HSE, named HSE1b, in the promoters of 54 upregulated genes which the study predicted that they might be direct downstream targets of AtHSFA1b. Based on that, Bechtold et al., (2013) predicted that AtHSFA1b might be operating a multilayer transcriptional network where it directly controls the expression of 54 genes that by turn control the expression of indirect targets of AtHSFA1b.

This part of the research contributed to the aforementioned published work. The main aims of this part were to validate the output of the microarray analysis by confirming the expression levels of selected genes from the microarray dataset including the 7 TFs that were predicted to be direct downstream targets of AtHSFA1b. The second aim was to examine the interaction between AtHSFA1b and the promoters of predicted downstream targets using ChIP-PCR (Section 2.2.2). However, this part of the study was further developed to examine the DNA binding activity of AtHSFA1b to the promoter of its target genes under heat stress conditions and compare that to the binding behaviour under normal growth conditions.

3.2 Results

3.2.1 Validating the microarray data with qRT-PCR:

The expression levels of *AtHSFA1b* in three independent lines of plants overexpressing *AtHSFA1b::mRFP* were measured. Furthermore, selected genes from the microarray dataset were also tested using qRT-PCR. The majority of the genes tested with qRT-PCR, including the 7 TFs that were proposed to be directly regulated by AtHSFA1b, showed increase in transcript levels compared to wild type col-0 plants. However, other genes that showed upregulation in the microarray dataset did not show any change in their expression in the qRT-PCR experiment. Those genes that did not show any change in expression include, *GA2OX6*, *FAD-binding Berberine 1* (*FBB-1*) and *WUSCHEL* (*AtWOX*) (figure 3.1)

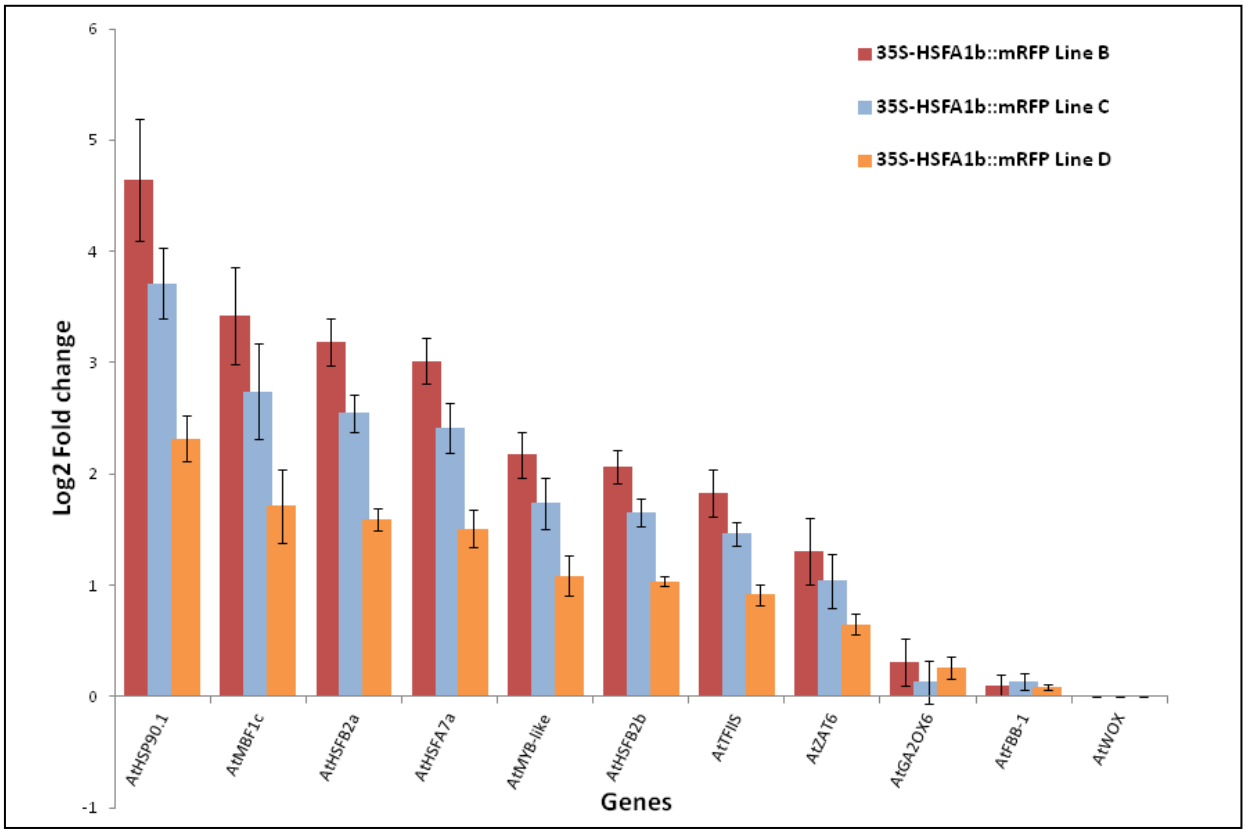


Fig.3.1. Validating microarray data with qRT-PCR. Overexpression of *AtHSFA1b* induces the expression of a number of genes including the 7 predicted TFs without applying any form of stress. Three 5-week-old *A. thaliana* independent lines (B, C and D) overexpressing *HSFA1b::mRFP* at different levels (165 folds, 95 folds and 80 folds respectively) compared to wild type plants were subjected to qRT-PCR. *AtACT2* was used for data normalisation. The bar plot shows the mean of Log2 expression in three biological replicates from each line ± standard deviations of the three independent experiments. The level of induction of target genes correlates with the level of overexpression of *AtHSFA1b* in a dose-dependent manner.

3.2.2 Confirming AtHSFA1b predicted downstream targets:

A number of the predicted downstream targets containing the HSE1b element in their promoters were tested with ChIP-PCR. The results showed that AtHSFA1b interacts with promoters of the selected genes including 7 TFs (figure 3.2)

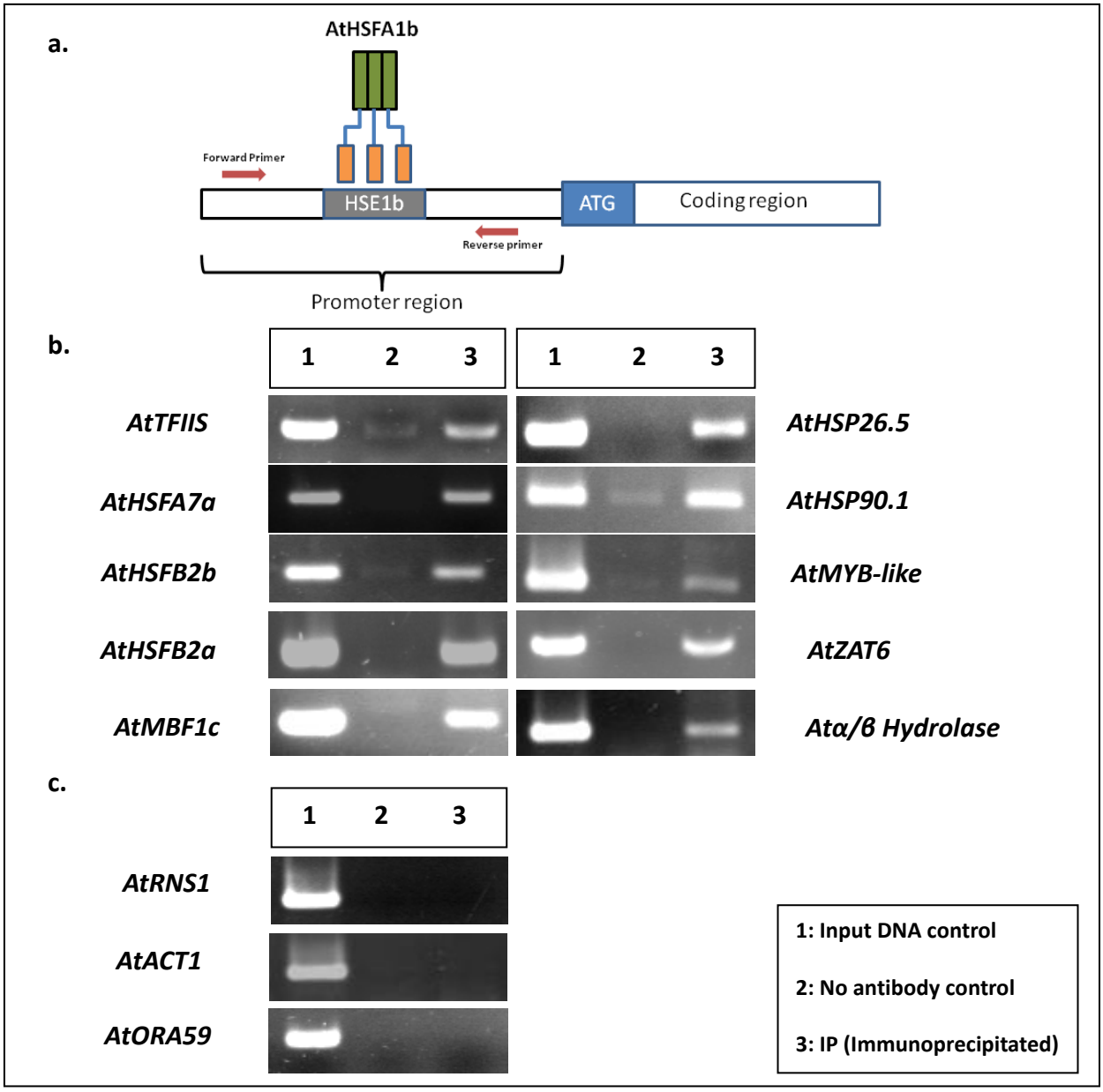


Fig.3.2. AtHSFA1b interacts with genes containing HSE1b element on their promoters. (a.) diagram showing the PCR amplification strategy, red arrows show the position of the primers used in PCR analysis, gray box shows the predicted position of HSE1b on the promoter region of the target genes. **(b.)** Gel electrophoresis images of the output of ChIP-PCR showing interaction between AtHSFA1b in *35S-AtHSFA1b::mRFP* plant with the promoter regions of its direct downstream targets including the 7 TFs predicted by Bechtold et al., (2013). Chromtin fragments of the size ~300 bp were prepared from three biological replicates of unstressed 5-week-old Arabidopsis plants overexpressing *AtHSFA1b::mRFP* (line B), immunoprecipitated by antibody to RFP tag. Purified immunoprecipitated chromatin fragments were subject to PCR analysis using primers spanning 150 - 200 bp around the predicted sites where HSE1b elements are located in the promoters of the predicted targets. **(c.)** negative control, genes known not to be interacting with AtHSFA1b.

3.2.3 AtHSFA1b releases some of its targets under heat stress:

Binding of AtHSFA1b to the promoters of its target genes was tested under heat stress and compared to its binding behaviour under normal growth conditions using ChIP-PCR. The results showed that AtHSFA1b releases some of the promoters of its target genes under heat stress despite showing positive signals under normal growth conditions (Figure 3.3).

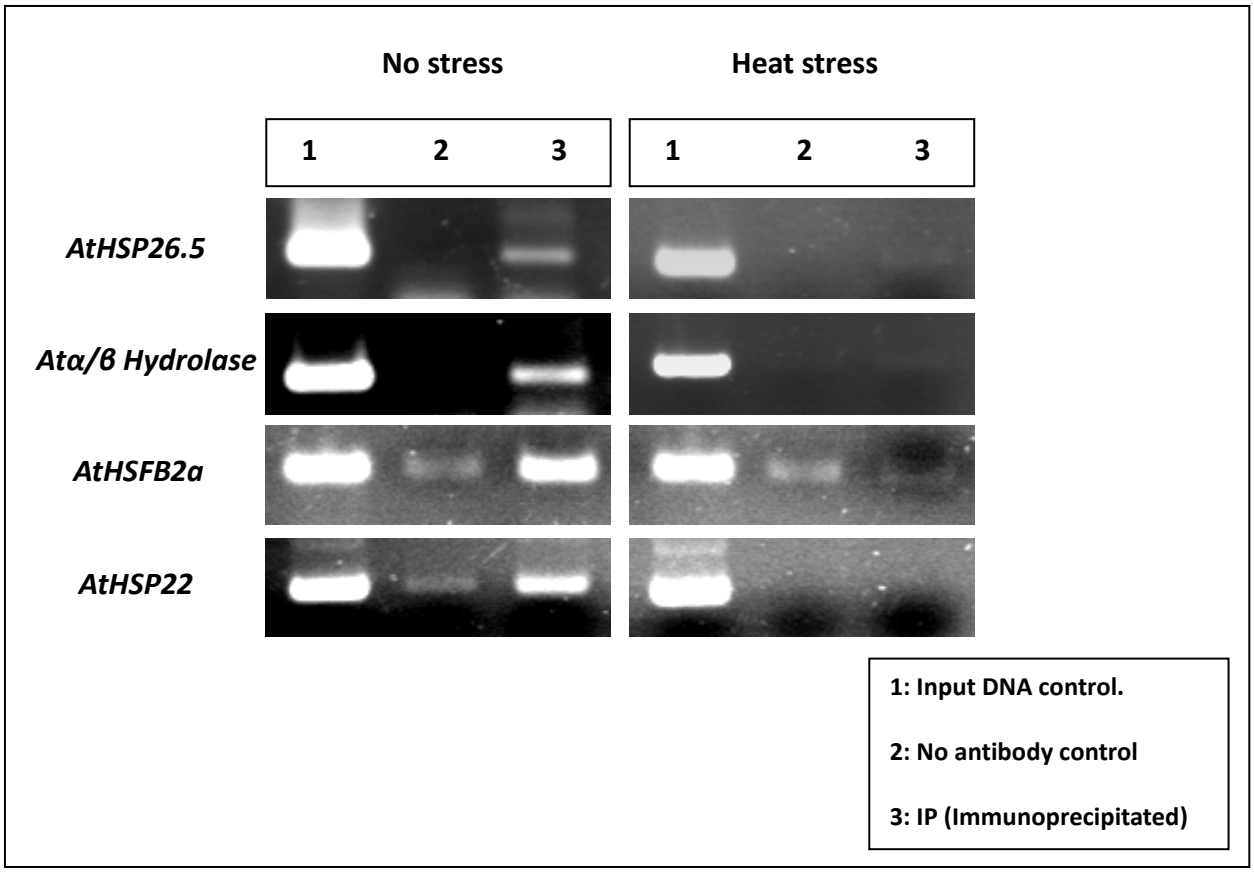


Fig.3.3. AtHSFA1b releases some of its promoter targets under heat stress. Gel electrophoresis images showing the output of the ChIP-PCR experiment of the comparison between the interaction of AtHSFA1b and its target promoters under no stress and heat stress conditions. Chromatin fragments were isolated immunoprecipitated from unstressed and heat stressed (37°C , 30 minutes) 35S-AtHSFA1b::mRFP plants, purified chromatin fragments were subjected to PCR analysis to test the binding of AtHSFA1b to HSE1b in the promoters of its target genes under non-stress and heat stressed conditions.

3.3 Discussion:

3.3.1 Confirming the results of the microarray experiment:

The qRT-PCR analysis clearly showed that overexpression of *AtHSFA1b* induces the expression of the selected genes from the microarray dataset. This added an extra validation and credibility to the microarray study done on *AtHSFA1b* overexpressing plants (Bechtold, et al., 2013). However, some of genes that the microarray showed as upregulated did not show any change in expression in the qRT-PCR experiment. It is important to note that the cutoff for the microarray study was *q*-value ≤ 0.05. Those genes that were shown to be induced as a result of overexpression of *AtHSFA1b* in the microarray dataset but did not show any change in expression in qRT-PCR experiment might be false positives which is a common issue in microarray experiments (Pawitan, et al., 2005). Therefore, it can be suggested that to ensure for expression studies is to validate the output of microarray by qRT-PCR (Dallas, et al., 2007). The seven TFs that were suggested to be directly controlled by AtHSFA1b also showed increase in expression levels in qRT-PCR experiment which indicates that there might be a possible involvement of those TFs in the regulation of some of the indirect target genes of AtHSFA1b.

3.3.2 AtHSFA1b directly controls the expression of 7 TFs

The ChIP-PCR experiment, one the other hand, showed positive interaction between AtHSFA1b with some of the predicted downstream targets of AtHSFA1b including those 7 TFs suggested to be directly regulated by AtHSFA1b. This suggested that the model that predicted direct targets of AtHSFA1b was accurate to a large extent. This also gave a strong indication that AtHSFA1b might be directly contributing the change of expression of those 7 TF genes. This result is of great importance because it gave early clues about the levels of control in the AtHSFA1b network. It showed hints of a possible hierarchical organisation in the AtHSFA1b network where

AtHSFA1b might be directly influencing the expression of those 7 TFs which by turn might be regulating the expression of some of AtHSFA1b indirect targets. However, it was not possible to predict how the AtHSFA1b transcriptional regulatory network operates from just the outputs of ChIP-PCR and qRT-PCR.

3.3.3 Unusual binding pattern of AtHSFA1b under heat stress

Comparing the binding of AtHSFA1b to the promoters of target genes under no stress and heat stress showed an interesting binding behaviour of AtHSFA1b under the two different conditions. It showed that AtHSFA1b no longer binds to the promoters of some of its target genes under heat stress (Figure 3.2). This result showed glimpses of evidence that the AtHSFA1b regulatory network might undergo changes in response to changes in growth conditions. Although the ChIP-PCR results from heat stressed and unstressed plants gave clue about possible changes in the AtHSFA1b network, it cannot be taken as a firm evidence of the changes that might occur in the AtHSFA1b network in response to changes in growth conditions. Further analysis is required to confirm the changes in the AtHSFA1b transcriptional regulatory network. An appropriate experiment to examine those changes in binding behaviour of AtHSFA1b to the promoters of its target genes would be a genome-wide survey of AtHSFA1b binding sites under the same growth conditions used in this experiment.

CHAPTER 4

Genome-wide mapping of AtHSFA1b binding profile

4.1 Introduction

AtHSFA1b and possibly all group-A1 AtHSFs are constitutively active and expressed in all *A. thaliana* tissues and their expression levels do not change in response to any applied stress, despite the controversy about the inducibility of *AtHSFA1e* (Bechtold, et al., 2013; Swindell et al., 2007; Miller and Mittler, 2006). This has led to the conclusion that group-A1 AtHSFs including AtHSFA1b act as master regulators of the HSR in *A. thaliana* plants (Scharf et al., 2012; Liu et al., 2011; Miller and Mittler, 2006; Nover et al., 2001). However, it has also been suggested that AtHSFA1a and AtHSFA1b are only involved in the immediate early heat stress response (Li et al., 2010; Busch et al., 2005; Lohmann et al., 2004).

Despite the extensive studies on the molecular roles of plant HSFs, little has been done to elucidate their regulatory networks. A number of studies attempted to resolve the transcriptional regulatory networks of HSFs through genome-wide analysis of their binding profiles and analysis of their transcriptomes in a number of species including human, *D. melanogaster* and yeast (Mendillo et al., 2012; Guertin and Lis, 2010; Hahn et al,. 1999). The data generated from those studies gave extremely valuable information about the broad functions of HSFs in those species and their involvement in other cellular processes beside stress response.

However, the majority of research done in plants relied mainly on transcriptomics data as the sole tool to understand the HSFs' regulatory networks (Pérez-Salamó et al., 2014; Li et al., 2013; Bechtold et al., 2013; Busch et al., 2005). While, studying the transcriptome of overexpressed or knocked out TFs is important to obtain a general overview of the genes that respond to the gain or loss of function of TF genes, it does not provide sufficient information about the complexity

and depth of transcriptional networks. The reason is that transcriptomics data cannot give a clear distinction between direct targets of TFs and indirectly regulated genes (Hull et al., 2013). Furthermore, it is more challenging to understand the structure of transcriptional networks from only transcriptomics data when the TF of interest is part of a large gene family whose members have a wide overlap in their roles such as the HSF family in *A. thaliana* and in other plant species (Wunderlich et al., 2014). Therefore, the first step to understand the hierarchical structure of transcriptional networks is to identify genes that are specifically targeted by the TF of interest.

A number of approaches have emerged in the last few years to resolve large scale identification of the binding activities of TFs with their cognate DNA *cis*-acting regulatory elements (Massie and Mills, 2008; Johnson et al., 2007). However, one method that seems to be able to show that distinction between direct and indirect targets of TFs *in vivo* is ChIP followed by high throughput methods (Collas, 2010). This method has been rapidly adopted as a standard global experiment to map binding profiles of TFs *in vivo* (Ho et al., 2011). The method is basically identical to that described in Section 2.2. However, the main difference, beside a few technical adjustments, is that in this case ChIP is coupled with a high throughput method such as tiling array (ChIP-chip) or high throughput sequencing (ChIP-SEQ) (Ho, et al., 2011). This allows for large scale identification of TFs binding sites instead of examining the interaction between a TF and selected individual targets (Johnson, et al., 2007). Studying genome-wide binding profiles of TFs is a crucial step to understand transcriptional regulatory networks structures and can give more insights about the diverse functional roles of TFs (Kaufmann, et al., 2009).

The success in validating some of the predicted downstream targets of AtHSFA1b and the peculiar binding behaviour of AtHSFA1b to its targets under heat stress observed in the ChIP-PCR experiment (Sections 3.2.2 and 3.2.3) was encouraging to carry out genome-wide mapping of AtHSFA1b binding sites. The loss of binding of AtHSFA1b to the promoters of its target genes observed in the ChIP-PCR experiment gave an indication that the AtHSFA1b regulatory network in *A. thaliana* might be subject to a reconfiguration during the transition for normal growth conditions to a heat stress response. However, it was not possible to come up with a definitive conclusion about the changes in the AtHSFA1b transcriptional network that take place in response to changes in growth conditions from a few tests on individual target genes. In order to examine this hypothesis, a genome-wide mapping approach of AtHSFA1b binding sites (ChIP-SEQ) was adopted using the same plants tested with ChIP-PCR under two conditions, no stress conditions and a short duration of heat stress (Section 2.1.1 and 2.1.2). The aim of this part of the study was to validate what was observed in the ChIP-PCR experiment and to have an in depth view of the changes in the AtHSFA1b transcriptional network in response to changes in growth conditions. Another aim was to explore the broad functional involvement of AtHSFA1b in other cellular processes and how that involvement changes with the change in growth conditions. The output of this experiment can be of great importance as the HSF families are highly conserved not only in plants but among all eukaryotes (Scharf et al., 2012; Åkerfelt et al., 2010; Nover et al., 2001).

4.2 Results:

4.2.1 The influence of peak calling algorithms on ChIP-SEQ output:

The ChIP-SEQ data showed that AtHSFA1b binds to thousands of targets in the *A. thaliana* genome under no stress condition. However, the number of AtHSFA1b binding sites drops dramatically after 30 minutes of heat stress. The number of ChIP-SEQ peaks obtained from the peak callers MACS v2 and CisGenome v2 (detailed comparison between MACS v2 and CisGenome v2 is described in section 4.3.1) under both conditions with the statistical cutoffs are summarised in table 4.1.

	MACS v2.0 (p-value ≤ 10^{-5}, q-value ≤ 0.05, IDR ≤ 0.01)	CisGenome v2.0 (FDR ≤ 0.05)
35S-AtHSFA1b::mRFP No stress	**7284**	**4792**
35S-AtHSFA1b::mRFP Heat stress	**654**	**237**

Table 4.1. AtHSFA1b binds more targets under no stress. The table shows the outputs of two different peak callers with the statistical cutoffs used. Peaks were called from sequencing dataset using two different programs (MACS v2 and CisGenome v2) to examin the effect of the peak calling procedure on the final out of the ChIP-SEQ experiment. Both programs showed that AtHSFA1b binds more targets in the A. thaliana genome under no stress.

An overlap between the outputs of the two programs was observed. However, a number of peaks were unique to each program. The total number of significant peaks called by CisGenome v2 was 34% less than those called by MACS v2 under no stress and was 64.8% less than the total number of peaks called by MACS v2 under heat stress. Furthermore, a poor overlap between the peaks called by the two programs was observed from the heat stress experiment (figure 4.1).

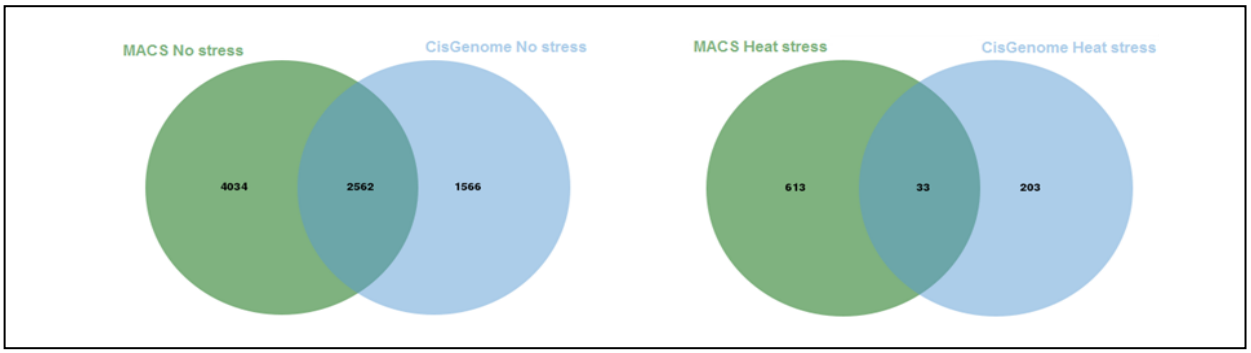


Fig.4.1. Overlap between MACS v2 and CisGenome v2 outputs under both conditions. Venn diagrams showing the total numbers of peaks called by each program and the numbers of unique and overlapping peaks under both conditions.

Positional analysis revealed that the highest number of binding events of AtHSFA1b *in vivo* occurs upstream of target genomic features and the number of binding sites spikes at very close proximities to the annotated transcriptional start sites (TSS) of target genes under both conditions (figure 4.2a). However, AtHSFA1b binding events were not exclusive to upstream promoter regions. The ChIP-SEQ data under both conditions showed that AtHSFA1b also binds to 5' and 3' untranslated regions (UTR), inside, and downstream of target genes but at less frequencies compared to binding events that occur upstream of target genes (figure 4.2b).

It is also worth mentioning that positional binding results were heavily dependent on the type peak caller program used. Common peaks called by the two programs were given different positions from the TSSs of target genes. CisGenome v2 showed that AtHSFA1b tends to bind upstream with a spike in the frequency of binding events at a very close distance to the TSSs of target genes under both conditions. Despite showing that the vast majority of AtHSFA1b bindings in the genome occur upstream of target genes, CisGenome v2 output also showed that AtHSFA1b binding activity is not exclusive to upstream regions (figure 4.2b). The output showed that AtHSFA1b can also bind inside and downstream of target genes. There is a slight decrease in the percentage of upstream binding events in favour of bindings inside target genes under

heat stress condition (figure 4.2).

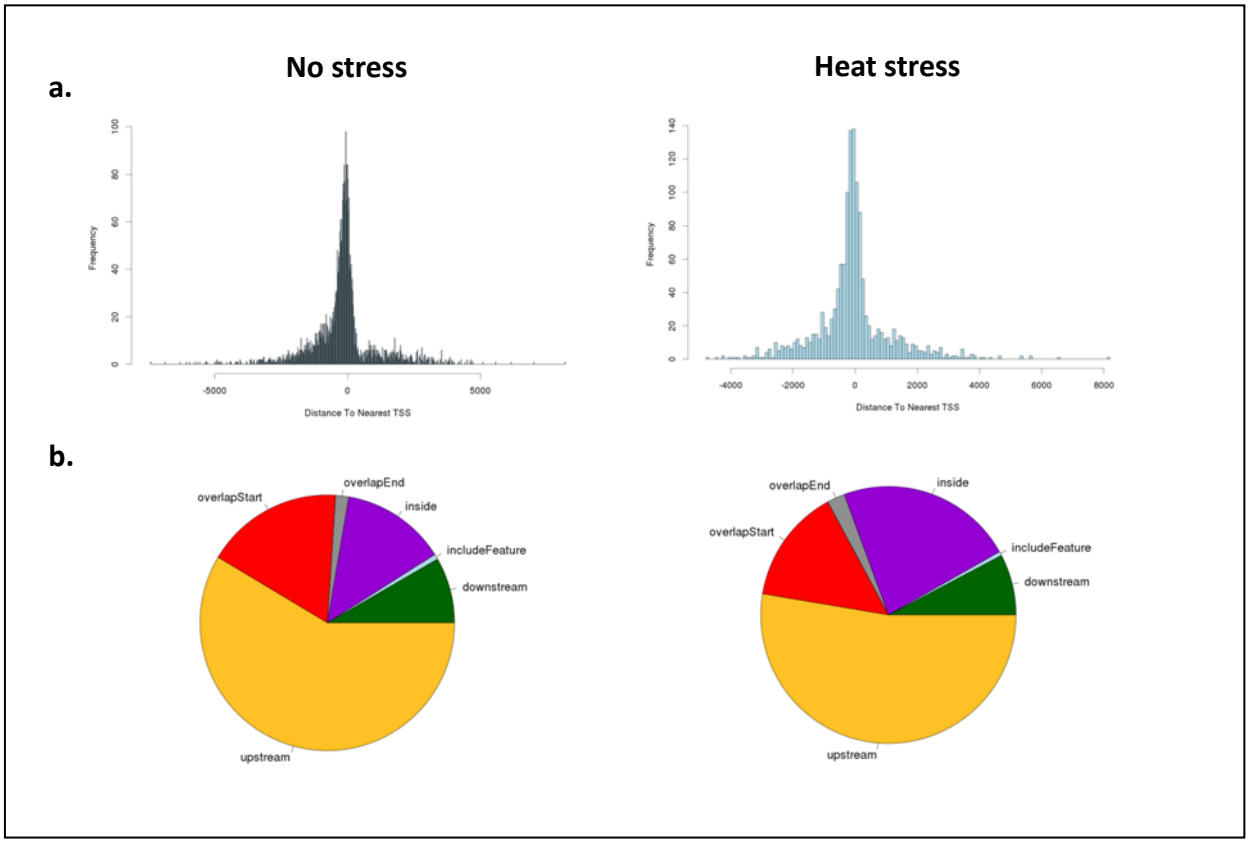


Fig.4.2. **Summary of CisGenome v2 output of AtHSFA1b genome-wide binding profile under both conditions. (a)** Histogram of the frequency of binding events relative to the distance from annotated TSSs of target genes. **(b)** Pie charts showing the distribution of AtHSFA1b binding events.

MACS v2 output showed two spikes of binding frequencies around the TSSs under both conditions. It showed that the highest number of bindings overlapped with TSS without giving clear distinction whether those binding events are inside or upstream of genomic features. No major difference in the distribution of binding events was observed between the two conditions apart from a slight decrease in the percentage of binding events that overlap with the TSS in favour of bindings that occur inside and overlap with the end of genomic features (figure 4.3).

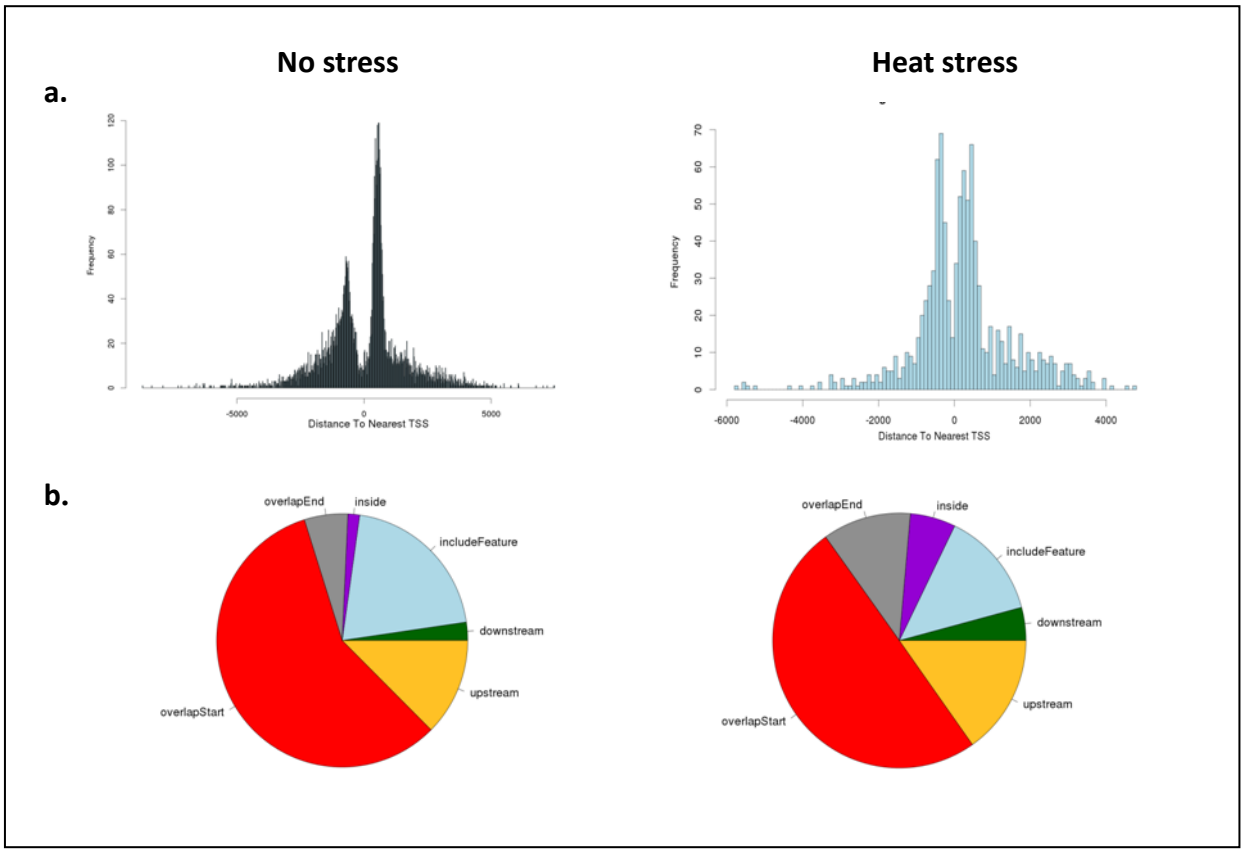


Fig.4.3. **Summary of MACS v2 output of AtHSFA1b genome-wide binding profile under both conditions. (a.)** Histogram of the frequency of bindings of AtHSFA1b relative to the distance from TSSs of target genes. **(b.)** Pie charts showing the distribution of AtHSFA1b binding events.

Despite the different outputs of the two programs in terms of position and distribution of binding events, they both seem to be in agreement that the majority of the binding events occur at very close proximity of the annotated TSSs of target genes. Moreover, they both show a similar shift in the binding distribution of AtHSFA1b during transition from normal growth conditions to heat stress.

4.2.2 The final output of the ChIP-SEQ experiment:

All of the statistically significant peaks from the two peak caller programs were merged into one dataset and used for further analysis. The overall outputs of the ChIP-SEQ experiments are summarised in table 2

		Total number of binding sites in the genome **MACS (p-value<10^{-5}, q-value<0.05, IDR<0.01)** **CisGenome (FDR<=0.05)**
35S-AtHSFA1b::mRFP	**No stress**	**9117**
35S-AtHSFA1b::mRFP	**Heat stress**	**857**

Table 4.2. The total number of AtHSFA1b binding sites in the *A. thaliana* genome under the two experimental conditions. The table shows the total number of AtHSFA1b under both conditions after merging the two datasets from MACS v2 and CisGenome v2. The statistical cutoffs used in both peak callers before merging the datasets are shown in the first row.

What was surprising is the extremely reduced number of significant binding sites under heat stress condition. Over 90.6% of the binding sites were lost after a 30-minute heat stress treatment at 37°C see figure 4.4.

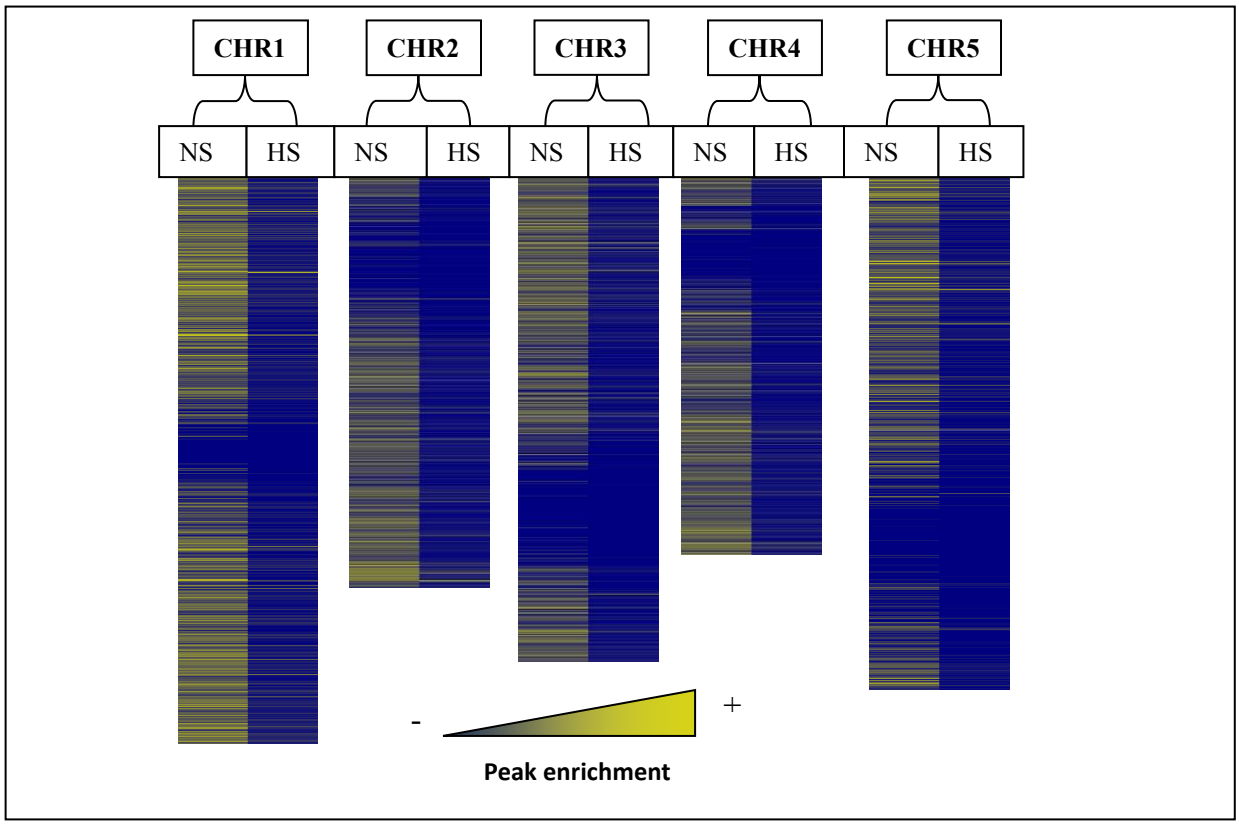


Fig.4.4. Overview of AtHSFA1b binding patterns on each chromosome under both conditions. NS indicates 'no stress' condition and HS is 'heat stress'. Yellow lines show regions in the *A. thaliana* genome that are bound by AtHSFA1b. The heat map was generated using Integrated Genome Browser (IGB) (Nicol, et al., 2009).

A degree of overlap between the binding sites under both conditions was observed. However, there were a few new unique significant binding sites gained under heat stress. The number of

sites retained and gained under heat stress condition was very small compared to the number of sites lost under the same treatment. There were 608 binding sites common to both treatments. However, 241 new binding sites were gained under heat stress treatment (figure 4.5).

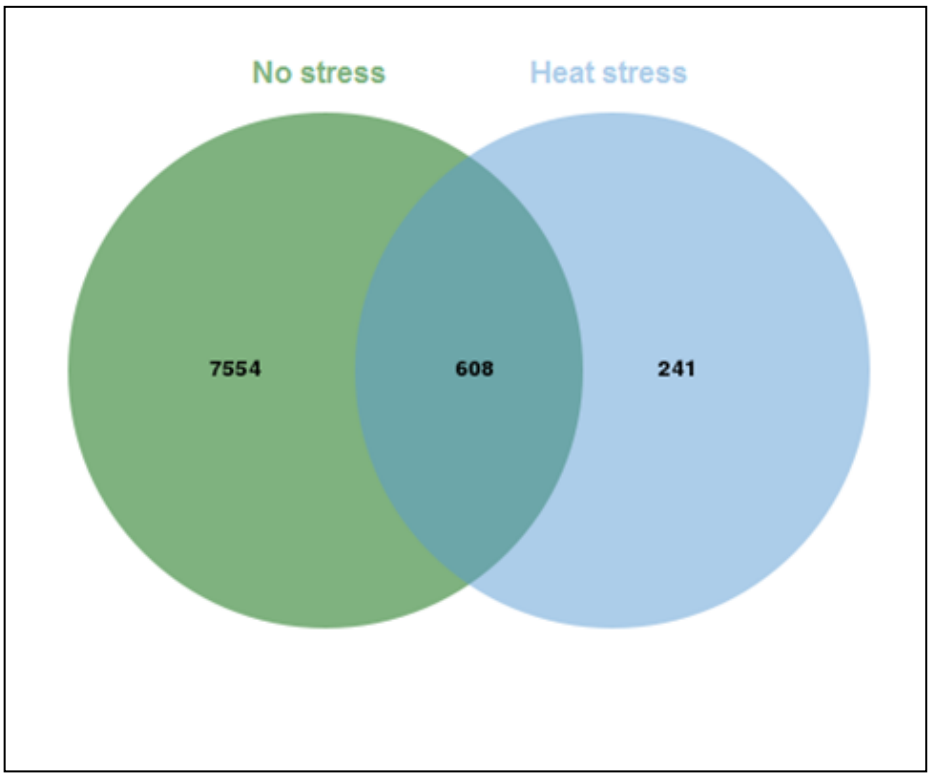


Fig.4.5. Overlap between AtHSFA1b binding sites under both conditions. Venn diagram showing the total numbers, unique and overlapping binding sites under both conditions. The numbers shown in the Venn diagram are less than the actual number of binding sites due to removal of duplicated AGI codes in each dataset by the Venn diagram generator.

4.2.3 Positional analysis of AtHSFA1b binding events:

The merged datasets were reanalysed for position and distribution of binding events. The vast majority of *in vivo* bindings occur upstream and overlap with the TSS of target genes which indicates that the preferred *in vivo* binding activity of AtHSFA1b occurs upstream of target genes at a very close distance from their TSSs. The final output is shown in figure 4.6.

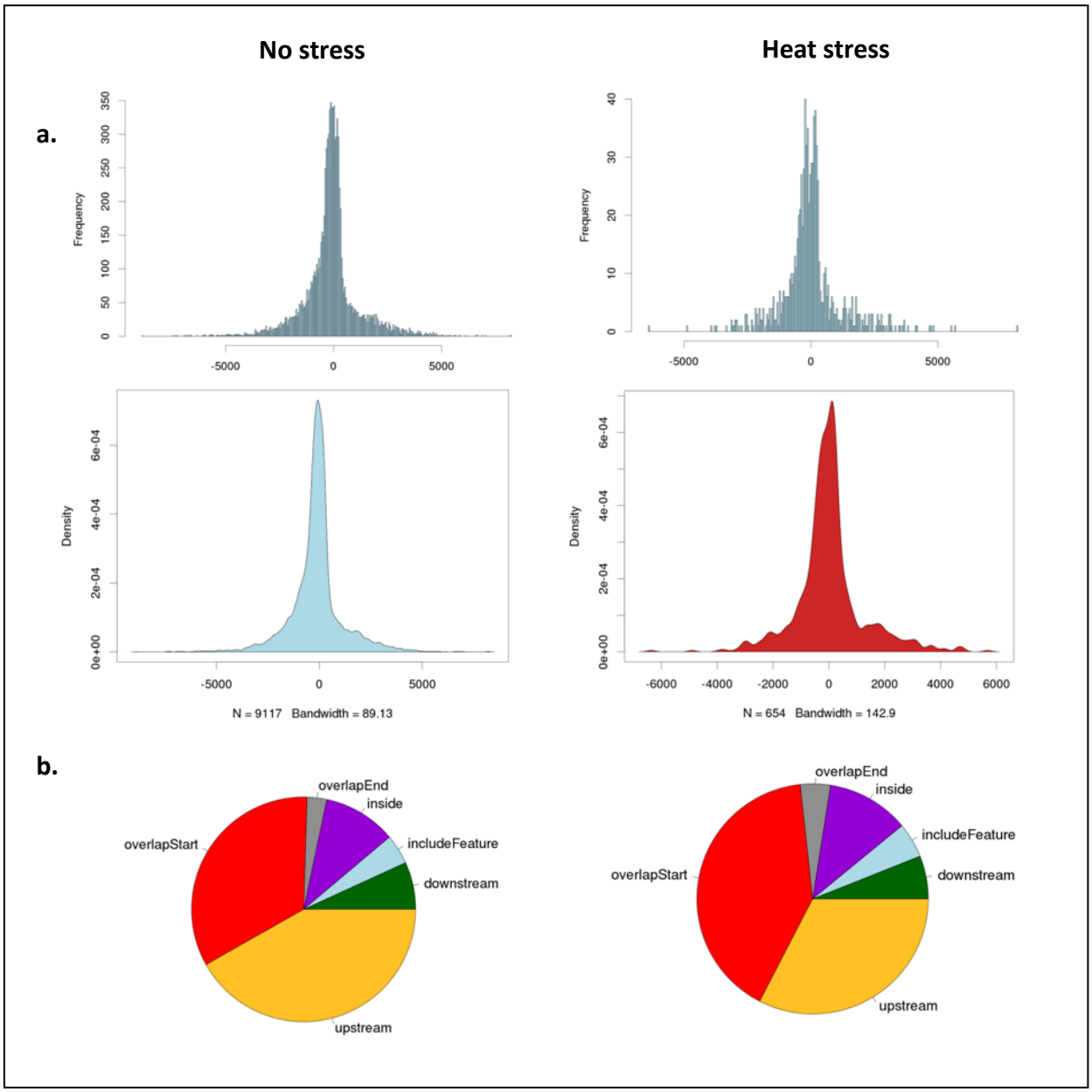


Fig.4.6. Summary of the output of the final merged data. (**a**) Histograms show the frequency of AtHSFA1b binding events relative to the distance from the nearest annotated TSSs of target genes. Below these are the corresponding density plots showing the density of probability of binding positions relative to the distance from the TSS of target genes. (**b**) Pie charts showing the distribution of the binding events around the target genes.

4.2.4 Confirming the loss of AtHSFA1b bindings under heat stress:

The ChIP-SEQ output is consistent with what was observed in the ChIP-PCR experiment (Section 3.2.2 and 3.2.3) on some of the target genes of HSFA1b under no stress and heat stress conditions where they showed complete disappearance of PCR signal under heat stress (figure 7)

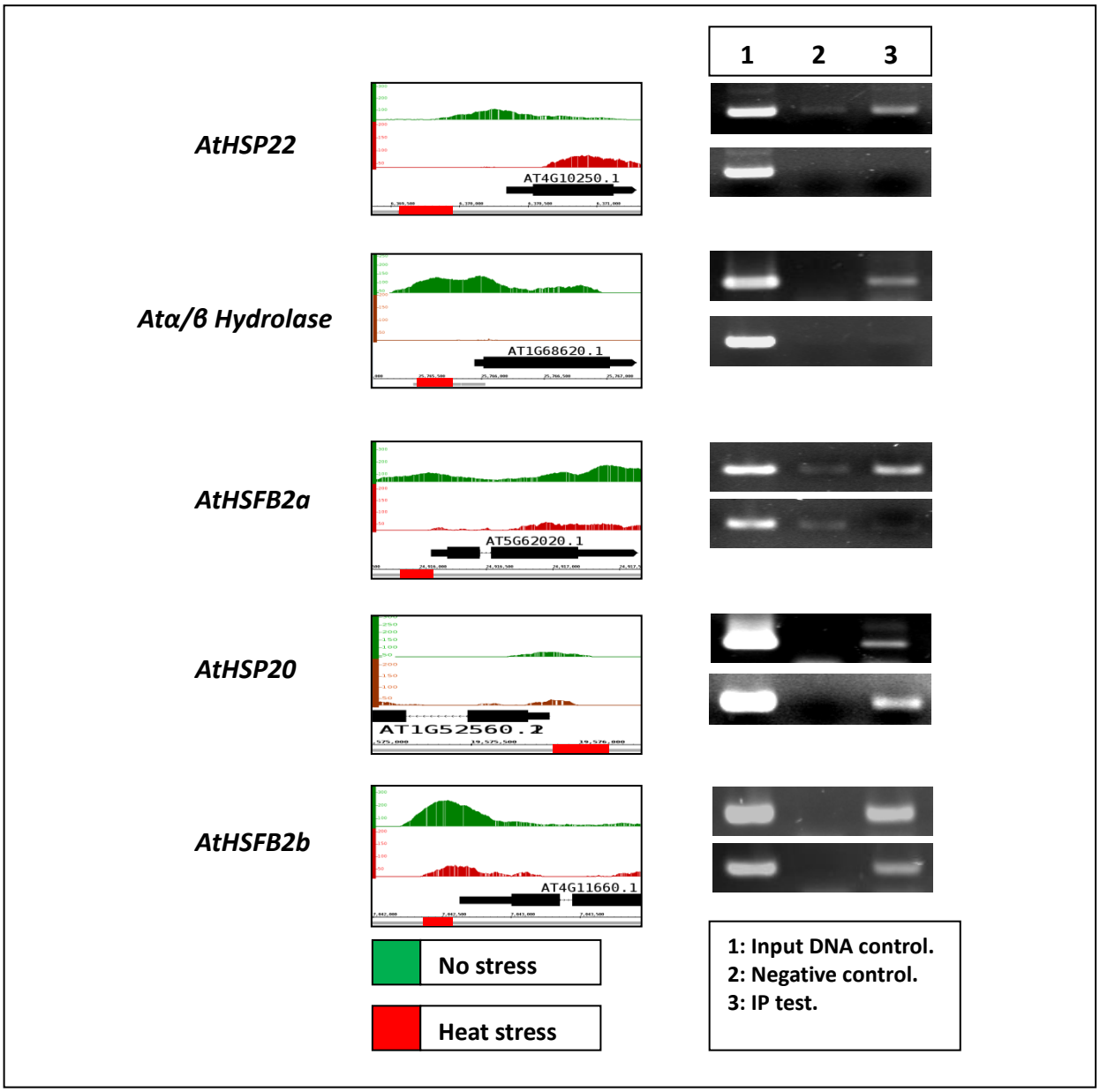


Fig.4.7. Confirming the loss of AtHSFA1b binding under heat stress. Snapshots from integrated genome browser IGB (Nicol et al., 2009) showing enriched peaks on the promoters of selected target genes (left panel) with their corresponding ChIP-PCR results (right panel) under both conditions. The red rectangle underneath the peaks and the target genes show the regions targeted by PCR. ChIP-SEQ and ChIP-PCR were carried out on 5-week-old *A. thaliana* plants overexpressing *AtHSFA1b::mRFP* under non-stress and heat stress (37°C, 30 minutes) conditions. Negative control in both experiments were wild type plants treated exactly the same way as the transgenic plants and using anti-RFP antibody to eliminate non-specific interactions between the antibody and other proteins.

4.2.5 The AtHSFA1b binding motif:

ChIP-SEQ peaks sequences from both conditions were extracted and used for *de novo* motif discovery analysis. Motif analysis was carried out using three different programs MEME (Bailey et al., 2009), STEME (Reid et al., 2011), and rGADEM (Mercier et al., 2011). The output of the three programs was in agreement and showed very high enrichment for a form of heat shock

cis-regulatory element (HSE). The analysis showed that AtHSFA1b tends to bind to **GA/TC** rich regions in the *A. thaliana* genome. The main feature of the motif discovered is that it consists of three repeats of the hexamer sequences **AGARRR/YYYTCT**. Therefore, the general structure of AtHSFA1b binding element is in the form **AGA[R]$_3$AGA[R]$_3$AGA / TCT[Y]$_3$TCT[Y]$_3$TCT.** No difference in the structure of HSE was observed between no stress condition and heat stress see figures 4.8 and 4.9.

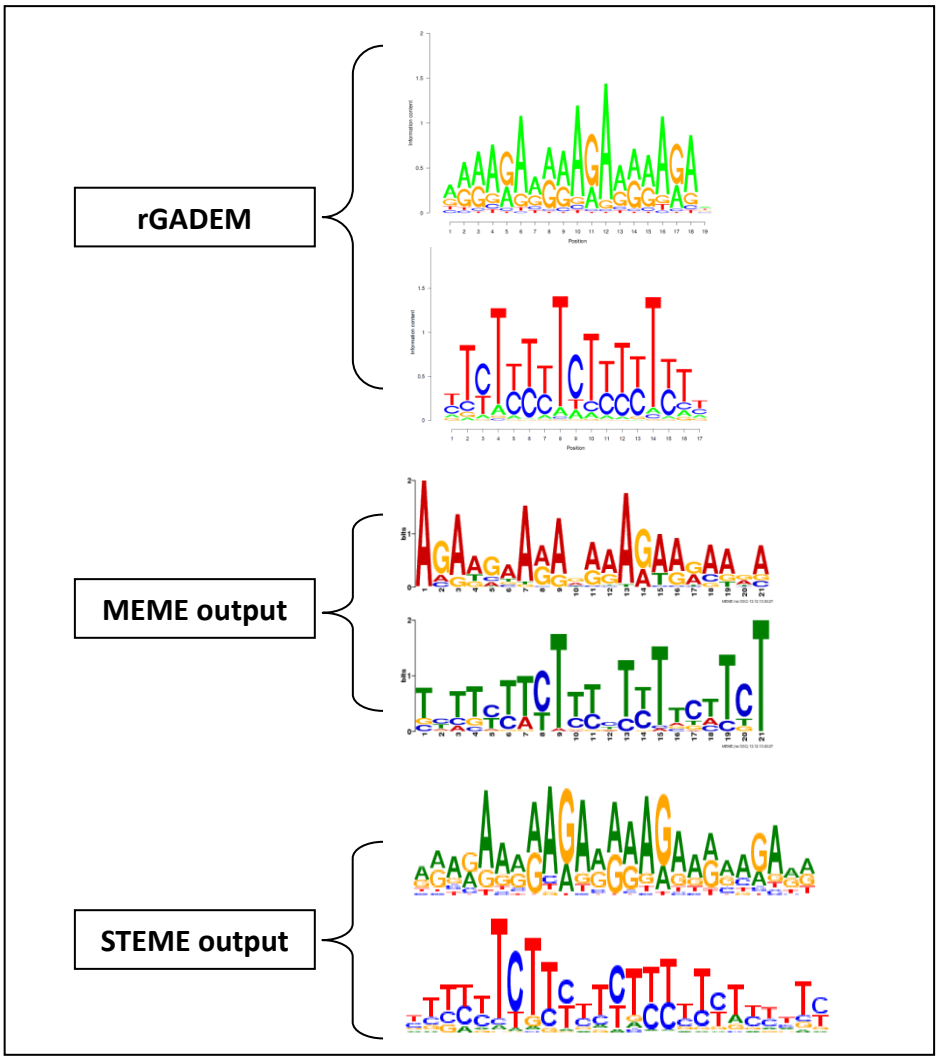


Fig.4.8. **Structure AtHSFA1b DNA binding element.** The figure shows the general structure of HSEs discovered from *de novo* motif analysis on the sequences of AtHSFA1b binding regions from the ChIP-SEQ dataset under both conditions (FDR<0.05). All binding regions sequences of AtHSFA1b from both treatments were analysed using *de novo* motif analysis. The *p*-value cutoff used in the motif analysis was *p*-value < 0.0001.

All possible forms of HSE were re-analysed using pattern matching programs and it was observed that any change in the **AGA/TCT** core sequences results in significant decrease in the enrichment of the motif within the datasets. The results demonstrate that there is only one predominant form of HSE in both datasets (**GAAGAAGAAG/CTTCTTCTTC**). Both HSEs are identical with equal frequency of occurrence under both conditions (figure 4.9)

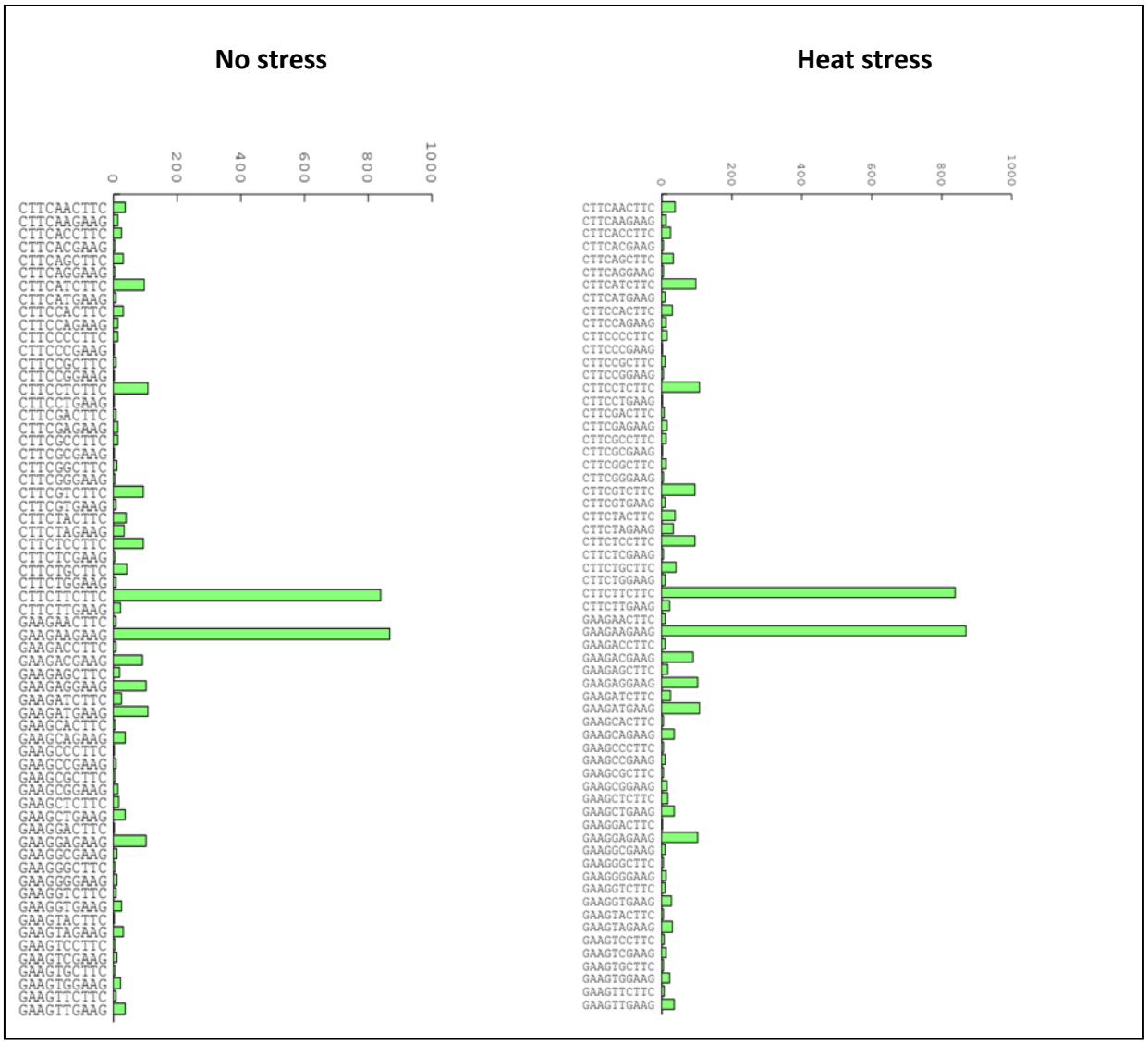


Fig.4.9. Frequency of the occurrence of different forms of HSE within AtHSFA1b target sequences. Histogram showing frequency of occurrences of different forms of HSEs within the ChIP-SEQ peaks sequences under both conditions. All different permutations on HSEs were reanalysed using pattern matching, each possible form of HSE was ranked based on its overall frequency of occurrence in the ChIP-SEQ peak sequences.

Further motif discovery analysis revealed two more known *cis*-regulatory elements that were significantly enriched (p-value < 0.0001) along with the HSE discovered in the peak sequences

but with less frequencies compared to HSE. The two other known cis-elements discovered with HSE were G-box/E-Box element (CACGTG), which is known as a binding site for bZIP and bHLH TFs (Qiu et al., 2009; Siberil et al., 2001; Williams et al., 1992), and Site II binding element (TGGGC[C/T]) which is known as a binding site for TCP transcription factors in *A. thaliana* (Giraud et al., 2010; Welchen and Gonzalez, 2006; Trémousaygue et al., 2003). Interestingly, G-box/E-Box was only present in ChIP-SEQ dataset under no stress condition and did not show up as a significant co-occurring *cis*-element under heat stress treatment. Site II binding element, on the other hand, was highly enriched along with HSE under both conditions. Other motifs from both experiments were less significant than G-Box/E-Box and Site II element (table 4.3 and figure 4.10)

No stress		Heat stress	
Consensus	**Annotation**	**Consensus**	**Annotation**
	G-Box/E-Box		Site II element
	Site II element		GAGA-box
	Unknown		Unknown
	Unknown		UP2 element (Tatematsu, et al., 2005)

Table 4.3. **Co-occurring *cis*-elements with HSE in AtHSFA1b target regions.** The table shows all significantly enriched co-occurring *cis*-elements discovered in the sequences of regions occupied by AtHSFA1b under both conditions (p-value < 0.0001) ranked from highest enriched (top of the table) motif to least enriched motif (bottom of the table).

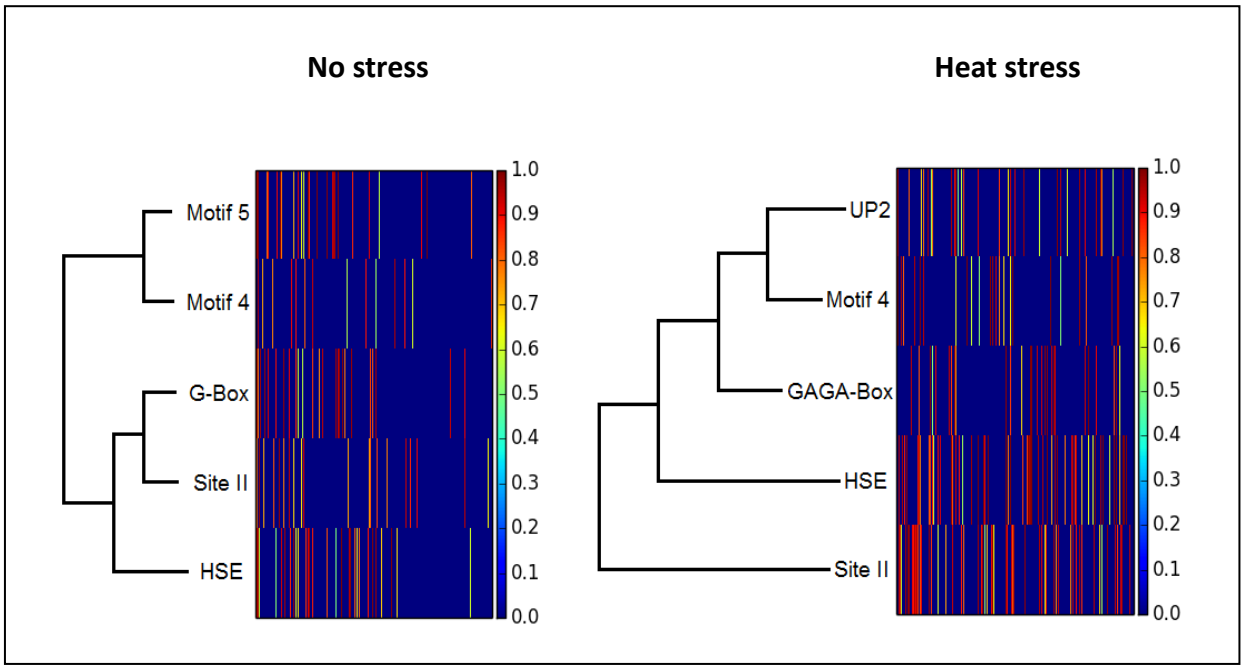


Fig.4.10. Enrichment of the HSE and other co-occurring cis-elements in AtHSFA1b binding regions. Two heat maps showing the best hit of each motif on each ChIP-SEQ positive sequence at (FDR < 0.05). The y-axis shows hierarchical clustering of the motifs discovered based on which sequences the *cis*-elements show the strongest presence. The x-axis is the ChIP-SEQ positive sequences (FDR<0.05) analysed in the motif discovery program.

4.2.6 Functional analysis of AtHSFA1b target genes:

The first observation from functional analysis on AtHSFA1b target genes revealed that AtHSFA1b does not target all HSFs in *A. thaliana* plants. Even when the overexpression of *AtHSFA1b* is up to 165 fold compared to wild type under non-stress conditions (Section 3.2.1) it still maintains specificity. Only 8 AtHSFs out of the 21 in *A. thaliana* plants were targeted by AtHSFA1b. The majority of the AtHSFs targeted by AtHSFA1b were bound in their promoter regions except AtHSFA4a and AtHSFA3 which had AtHSFA1b bound inside their genomic features (table 4.4).

AGI code	Name	Binding region(s)	Distance from TSS (bp)	Condition
AT4G11660	*AtHSFB2b*	Upstream Upstream	-180 -330	No stress and heat stress
AT4G18880	*AtHSFA4a*	Inside	640	No stress
AT3G51910	*AtHSFA7a*	Upstream	-94	No stress
AT5G62020	*AtHSFB2a*	Upstream	-293	No stress
AT1G46264	*AtHSFB4*	Upstream	-4599	No stress
AT3G24520	*AtHSFC1*	Upstream	-2	No stress
AT3G02990	*AtHSFA1e*	Upstream	-1249	No stress
AT5G03720	*AtHSFA3*	Inside	418	No stress

Table 4.4. **List of *A. thaliana* HSFs targeted by AtHSFA1b**. The table shows all *AtHSFs* targeted by AtHSFA1b with their AGI codes, position of binding, distance from their TSS and under which condition they were bound by AtHSFA1b.

Gene ontology analysis on AtHSFA1b target genes showed high enrichment for groups of genes that are involved in various biological processes. The analysis showed high enrichment for groups of genes that are involved in abiotic and biotic stress response including, heat (GO:0009408), light stimuli (GO:0009416), osmotic stress (GO:0006970), and response to bacterium (GO:0009617). Moreover, the output of the analysis showed that AtHSFA1b targets groups of genes that are responsive to endogenous and exogenous chemical stimuli such as, hormone signalling (GO:0009725) and metal ions (GO:0010038). Those groups of genes represent what would be expected to be controlled by an HSF in general (Pérez-Salamó et al., 2014; Bechtold et al., 2013; Nishizawa-Yokoi et al., 2011; Voyer and Heikkila, 2008; Miller and Mittler, 2006). However, two significantly enriched groups of genes targeted by AtHSFA1b were unexpectedly highly enriched. Those two groups are genes that are involved in developmental processes (GO:0032502) and transcription factors (GO:0006350) (figure 4.11).

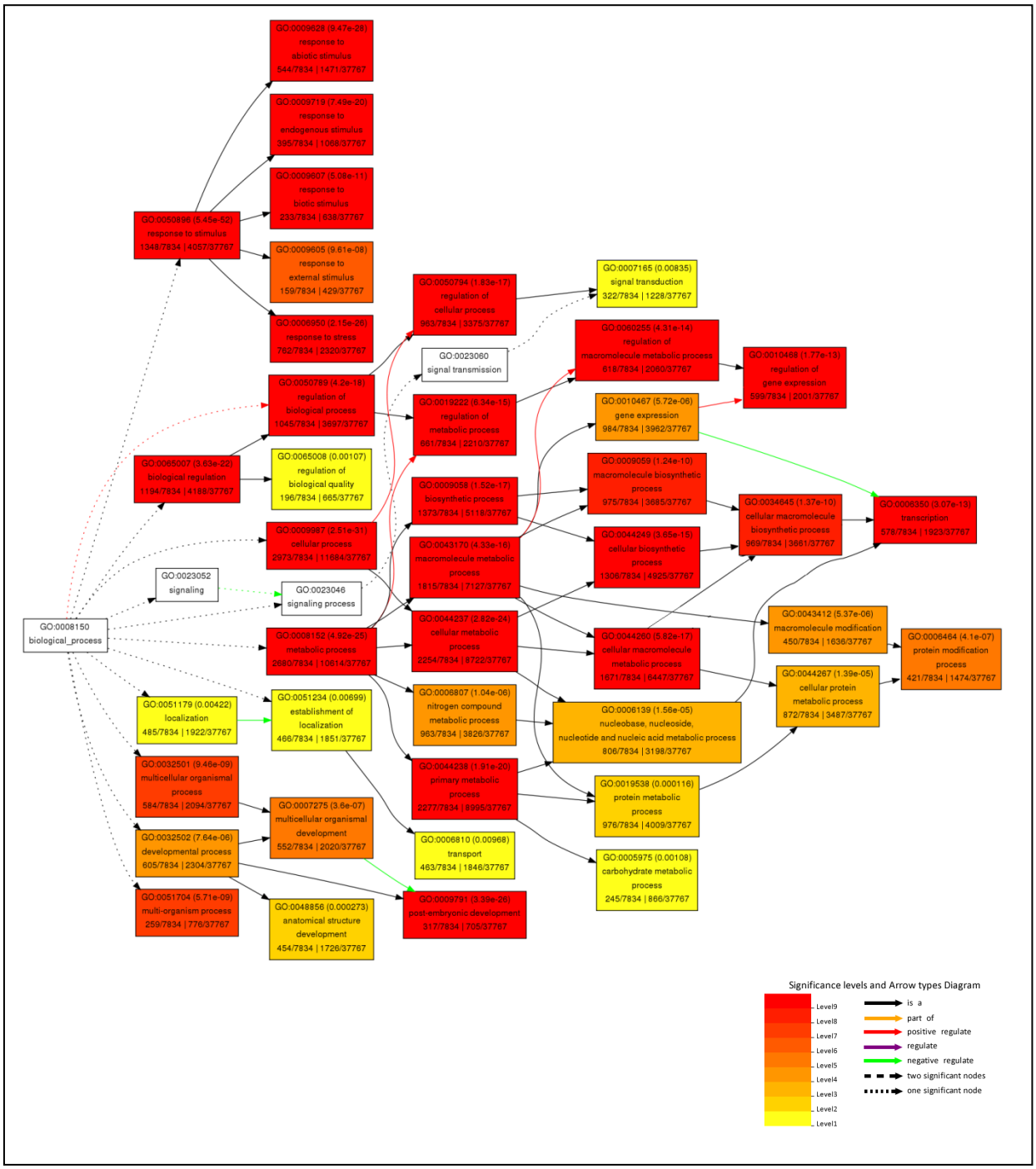


Fig.4.11. Functional enrichment of groups of genes targeted by AtHSFA1b under non-stress condition. Gene ontology enrichment map of the biological functions of AtHSFA1b target genes under no stress condition. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Bonferroni *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis of groups of genes targeted by AtHSFA1b was Bonferroni adjusted *p-value* ≤ 0.01.

Similarly, AtHSFA1b targeted groups of genes that are involved in various cellular functions under heat stress. However, the overall enrichment and the statistical significance of those groups of genes were far less than the enriched groups of genes targeted by AtHSFA1b under no

stress due to the small number of genes targeted by AtHSFA1b under heat stress compared to that under no stress (figure 4.12).

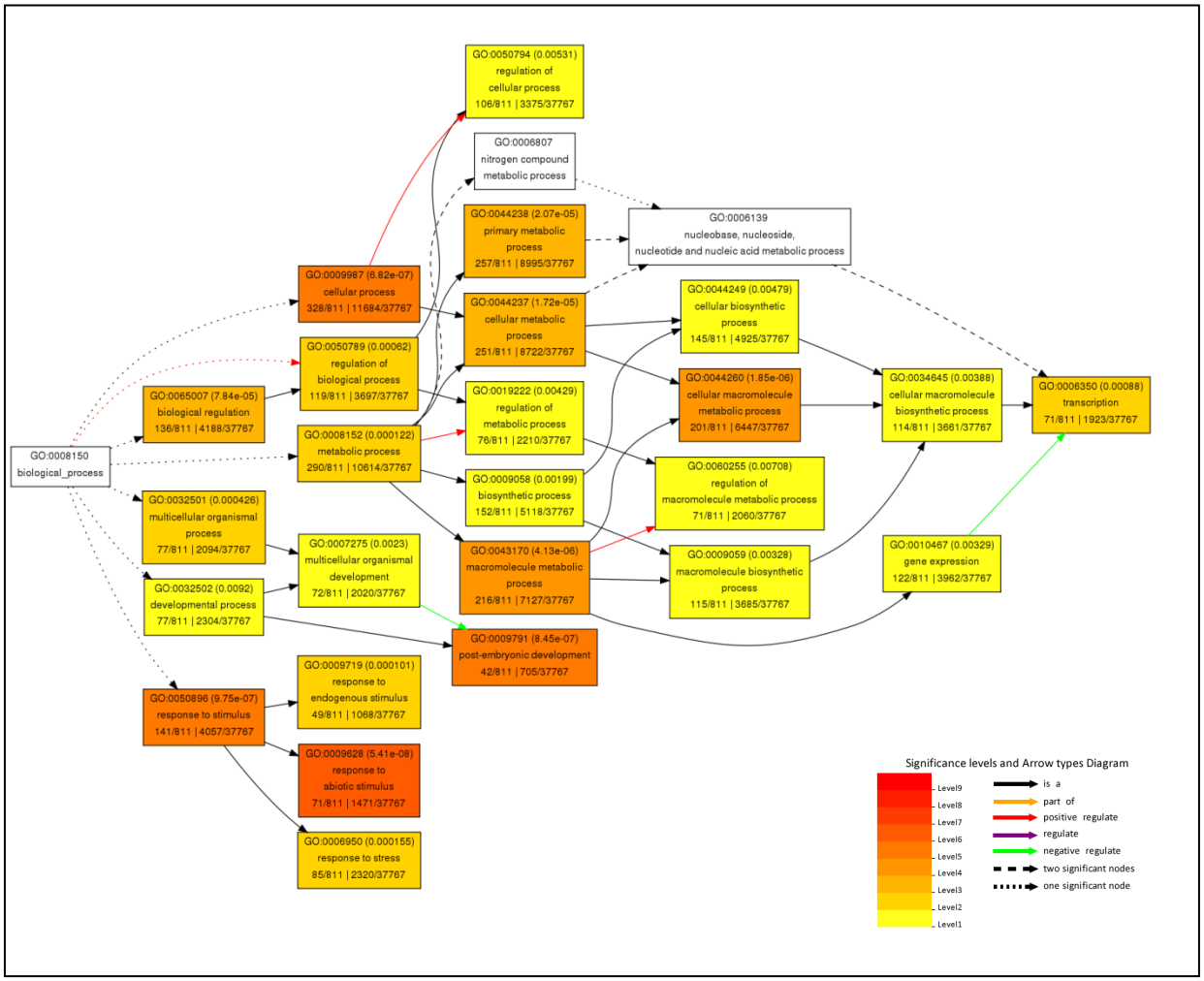


Fig 4.12. Functional enrichment of groups of genes targeted by AtHSFA1b under heat stress condition. Gene ontology enrichment map of the biological functions of AtHSFA1b target genes under heat stress condition. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Bonferroni *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis of groups of genes targeted by AtHSFA1b was Bonferroni adjusted *p-value* ≤ 0.01.

In terms of molecular functions, AtHSFA1b targeted mainly three distinct groups. The most significant group of genes targeted by AtHSFA1b was 'binding' (GO:0005488) including 'DNA-binding proteins' (GO:0003677) and 'protein-binding proteins' (GO:0005515). The second was a group of genes that code for enzymes involved in 'kinase activity' (GO:0004672). The third, was group of genes that code for enzymes with 'glycosyl transferase activity' (GO:0016757) (figure 4.13)

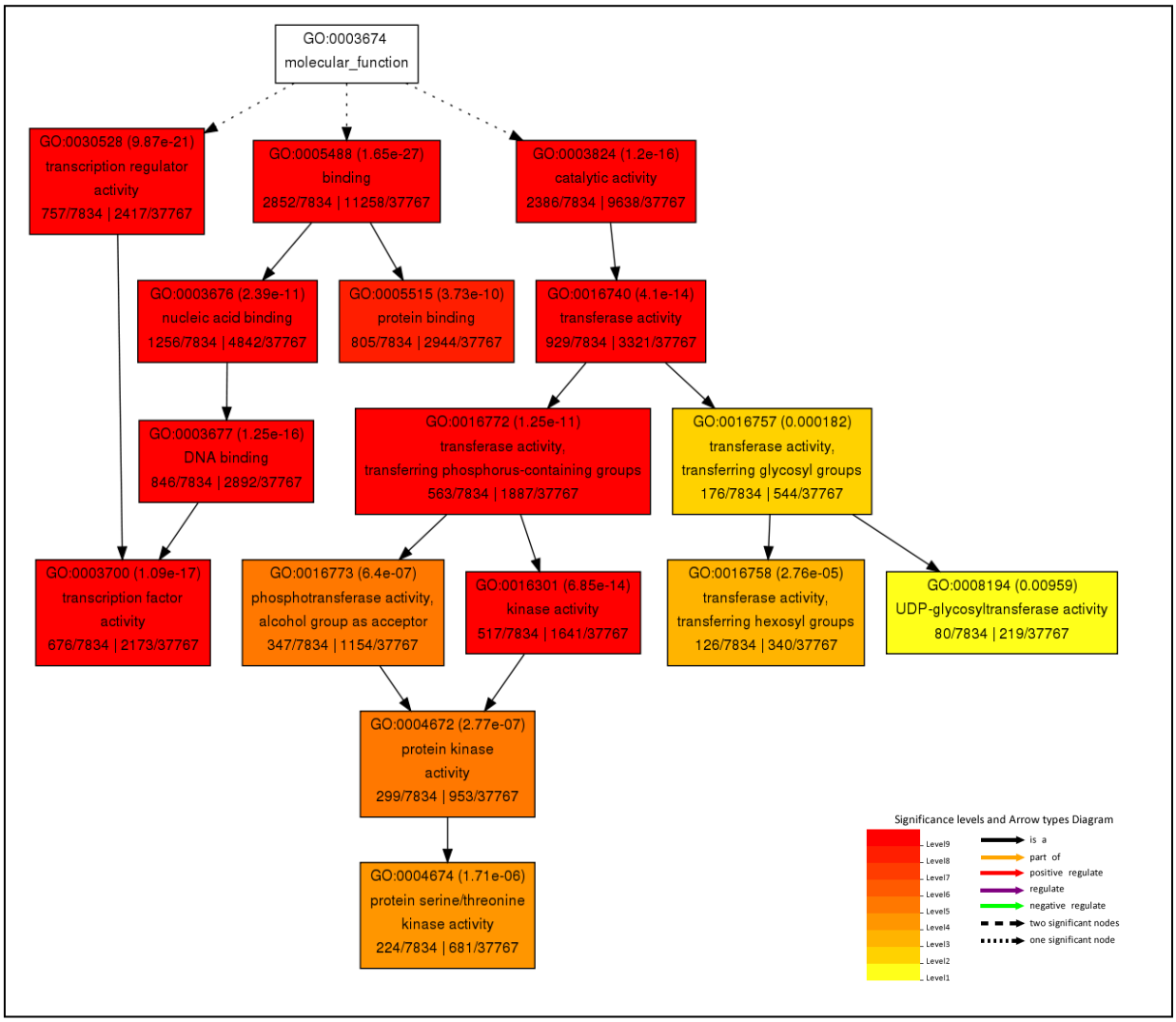


Fig.4.13. Molecular function enrichment of groups of genes targeted by AtHSFA1b under non-stress condition. Gene ontology enrichment map of the molecular functions of AtHSFA1b target genes under no stress condition. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Bonferroni *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis of groups of genes targeted by AtHSFA1b was Bonferroni adjusted *p-value* ≤ 0.01.

Heat stress treatment, on the other hand, showed complete loss of the kinase and glycosyl transferase groups. Only the TFs (GO:0003700) and protein-protein (GO:0005515) interaction groups remained highly enriched under heat stress (figure 4.14)

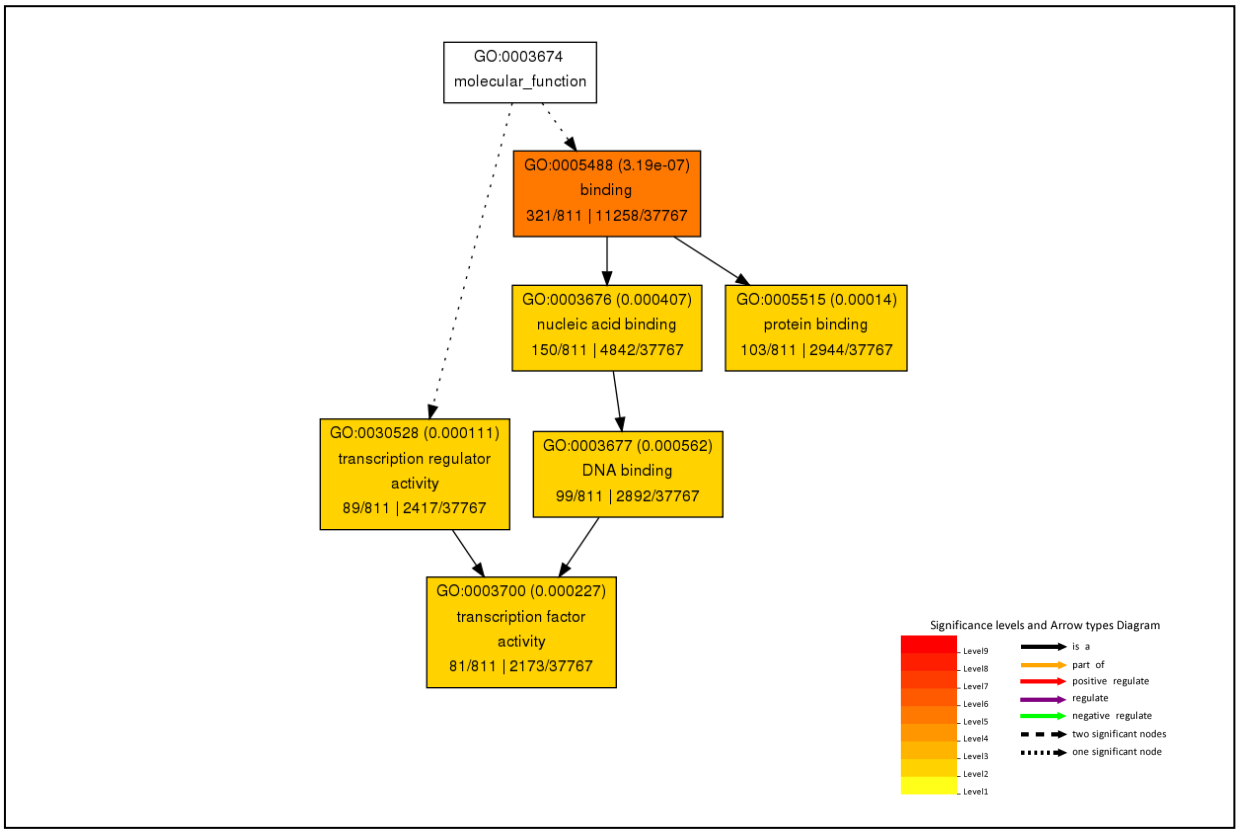


Fig.4.14. Molecular function enrichment of groups of genes targeted by AtHSFA1b under heat stress condition. Gene ontology enrichment map of the molecular functions of AtHSFA1b target genes under heat stress condition. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Bonferroni *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis of groups of genes targeted by AtHSFA1b was Bonferroni adjusted *p-value* ≤ 0.01.

The groups of genes that were targeted by AtHSFA1b under no stress and showed high enrichment for molecular functions were isolated and re-analysed. Surprisingly there was extremely little overlap between the biological functions of those groups. TFs showed high enrichment for two major biological processes, developmental processes (GO:0032502), and response to stimulus (GO:0050896). Whereas the group of genes that code for proteins involved in kinase activity showed high biological process enrichment for recognition and hormone signalling processes. The third group of genes that code for protein involved in glycosylation showed only enrichment for biological processes related to plant cell wall modification (See appendix C).

A large number of TFs targeted by AtHSFA1b are well-characterised and their roles in the regulation of developmental processes and stress response were experimentally validated. Examples of experimentally validated developmental TFs targeted by AtHSFA1b include, APETELLA1 (Mandel et al., 1992), ARF19 (Okushima et al., 2007), BEL1 (Ray et al., 1994), and SVP (Lee et al., 2007; Hartmann et al., 2000). The ChIP-SEQ data showed that AtHSFA1b targets a large number of TFs. The analysis showed, that the TFs targeted by AtHSFA1b show high enrichment for two distinct biological processes. The first is related to plant developmental processes. The second is mainly related to abiotic and biotic stress response including, temperature stimulus (GO:0009266), response to light stimulus (GO:0009416), osmotic stress (GO:0009670) and response to bacterium (GO:0009617). Further gene ontology analysis on the TFs targeted by AtHSFA1b showed that they are involved in numerous developmental processes in *A. thaliana* (figure 4.15). The TFs group targeted by AtHSFA1b under heat stress showed an overlap to a great extent with those enriched under no stress. However, the number of those TFs targeted under heat stress was less than those targeted under no stress condition which resulted in no significant enrichment of any biological process.

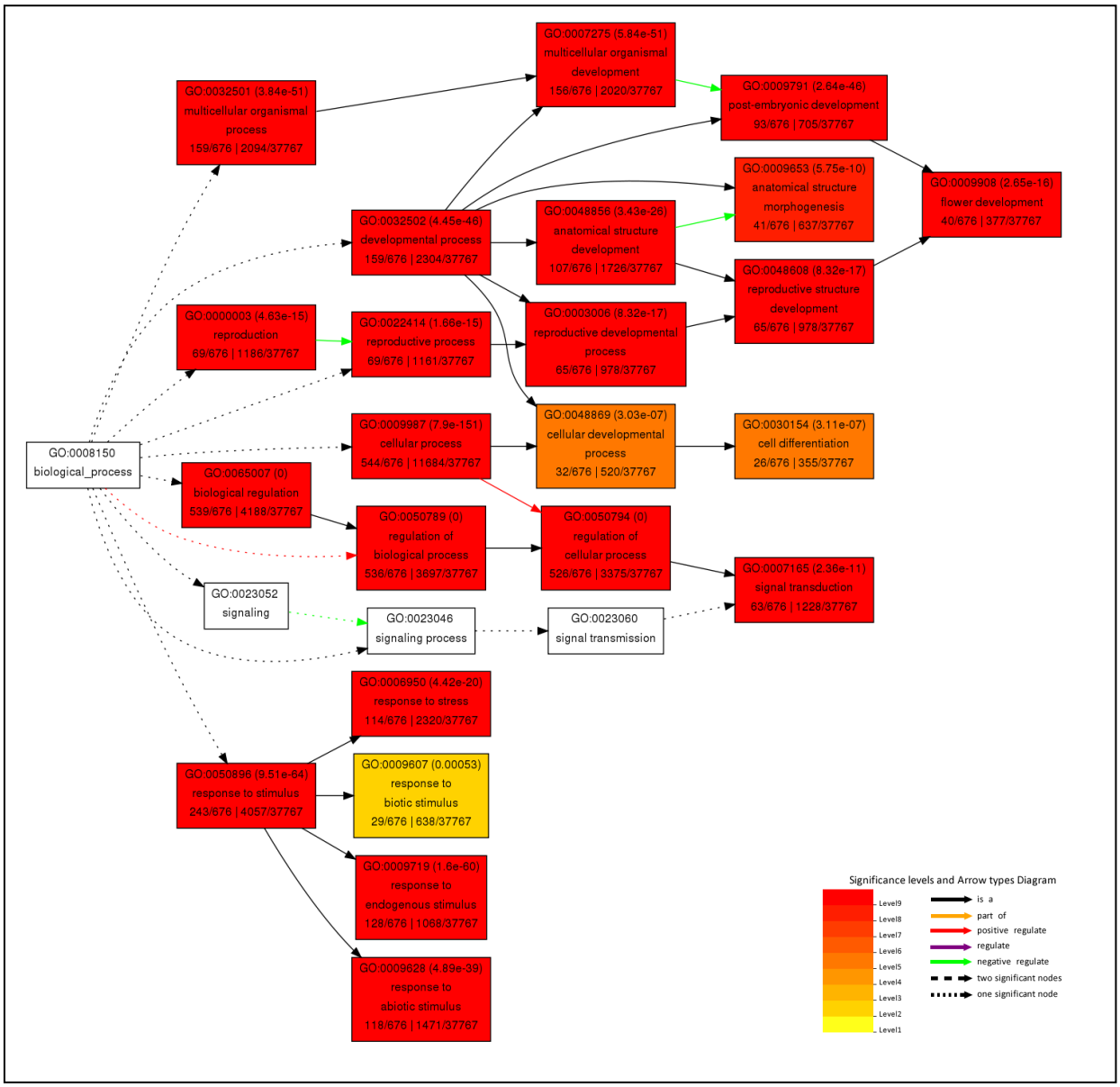


Fig.4.15. Gene ontology enrichment map of the biological processes of TFs targeted by AtHSFA1b under no stress. The map was adjusted by removing the enrichment for the group genes involved in 'expression activity' to eliminate bias in the analysis.

4.3 Discussion:

4.3.1 The ChIP-SEQ output is highly influenced by peak callers algorithms:

The ChIP-SEQ analysis showed that the general output relies to a large extent on the type of peak caller program used. Different programs with different algorithms tend to give different outputs depending on parameters used to call peaks (Landt, et al., 2012). Other peak caller programs other than MACS v2, CisGenome v2 and SIPeS (Wang, et al., 2010; no included in the analysis) do not distinguish between paired-end and single-end reads; therefore they were discarded from the analysis. The two programs used in this study, MACS v2 and CisGenome v2, use similar algorithms and statistical models in determining positive peaks. The main difference between the two peak callers used is some small technical details and how final positive peaks are called and shaped from ChIP-SEQ sequences. However, those differences seem to have an impact on the overall output of the analysis.

Significant peaks are called by MACS v2 in the presence of control samples by applying dynamic Poisson distribution. Then it compares the ChIP sample to negative control with a pre-set statistical parameter p-value < 10^{-5}. The q-value is calculated by dividing the tags counts of control peaks over the tags counts of ChIP peaks. CisGenome, on the other hand, scans the genome with a default minimum window size 100bp (can be changed by the user) and locates regions with enriched read counts. In the presence of negative control, CisGenome v2 applies a binomial model to determine the significance of peaks by comparing read enrichments in ChIP samples to negative control. Finally, the FDR is then calculated by determining the ratio between read enrichments in the control to the observed in ChIP samples. The directionality of reads is considered by CisGenome v2 and by default the program refines peak boundaries to a specific range (minimum 100bp, maximum 750bp) depending on the Watson and Crick tags of

peaks in each sample (Landt et al., 2012; Laajala et al., 2009). MACS v2 benefits from its compatibility with IDR analysis which measures the consistency between replicates in ChIP-SEQ experiments. This IDR parameter adds an extra validation and credibility to the final output. The IDR analysis discards peaks that are not present in the both replicates. However, there are drawbacks in the IDR analysis. The main issue with that type of analysis in ChIP-SEQ experiments is that it automatically discards peaks that are not in complete overlap in position even if they are significant based on other statistical measures and if the same peaks and present in replicates.

In terms of volume of output MACS v2 gave significantly more peaks than CisGenome v2 (table 4.1). The widths of peaks called by MACS v2 were more variable than those called by CisGenome v2. Peaks called by MACS v2 showed a wide range from 150 up to 7300bp in width. Whereas, peaks called by CisGenome v2 seemed to be more uniform with reasonable width range from 100 to 750bp. The differences in the natures of the peaks called by the two programs led to differences in the final output. The variability in the widths of peaks called MACS v2 affected its final positional output of the binding sites. Positional analysis of AtHSFA1b binding sites based on MACS v2 output showed a high degree of uncertainty of the binding events as it showed that the highest percentage of binding events were either overlapping with TSS or including an entire genomic feature (figure 4.3). In the case of CisGenome v2, the positional analysis of the frequency of binding events was more consistent as it showed that the vast majority of binding events occur upstream of genomic features at very close proximity to the TSSs (figure 4.2), which is consistent with what is widely known about transcriptional regulation in plants (Shamimuzzaman and Vodkin, 2013; Yant, et al., 2010; Kaufmann, et al., 2009; Zheng, et al., 2009; Berendzen, et al., 2006). However, due to its strict built-in statistical

model, CisGenome v2 tends to produce a high rate of false negatives which can result in an underestimate of the output of ChIP-SEQ and consequently losing valuable real peaks. Therefore, it can be concluded that the best practice is to merge the outputs of the two peak callers into one dataset for further analysis to allow for more precise analysis of position and distribution of binding events and to reduce the rate of false negative in the data.

4.3.2 Overexpression of *AtHSFA1b* does not change its *in vivo* binding behaviour:

The ChIP-SEQ experiment showed that AtHSFA1b binds to thousands of targets in the *A. thaliana* genome under non-stress conditions (table 4.2). This result is consistent with the majority of results from genome-wide mapping of TFs binding profiles in plants and other species where it has been shown that TFs bind to thousands and in some instances tens of thousands of binding sites in genomes (Savic et al., 2013; Mendillo, et al., 2012; Wang, et al., 2012; Yant, et al., 2010; Kaufmann, et al., 2009; Valouev, et al., 2008). Moreover, the analysis showed that AtHSFA1b *in vivo* binding events spike at upstream regions of target genes at very close proximity to the annotated TSSs. This is also in agreement with the majority of published work about the transcriptional regulation machinery in plants which showed that TFs interactions occur mainly on upstream regions of target genes at very close distances from the annotated TSSs (Shamimuzzaman and Vodkin, 2013; Yant, et al., 2010; Kaufmann, et al., 2009; Berendzen, et al., 2006).

One interesting observation is that AtHSFA1b bindings are not exclusive to upstream regions. The output of ChIP-SEQ experiment showed that AtHSFA1b also binds to target regions inside and downstream of genomic features but with lower frequencies than those occurring upstream of target genes. It has been shown that those binding events in human cells,

particularly bindings of TFs on exons, can play a role in re-defining DNA codons which can lead to changes protein amino acid sequences (Stergachis, et al., 2013). However, the impact of those binding events on gene expression and the biological significance of such events are widely unknown in the plant realm. There is a possibility that those binding sites might be occuring on alternative promoters where there might exist an alternative TSS for other target genes. The other possibility, however, is that those binding sites might just be accessible sites for AtHSFA1b and it remains bound on those sites without having any significant functional effect. It is not possible to come up with a clear answer about the consequence of those binding events from this experiment. However, this still remains an interesting observation and it agrees with all of the ChIP-SEQ experiments done on plants where there has been evidence that TFs can bind to regions other than promoters (Shamimuzzaman and Vodkin, 2013; Yant, et al., 2010; Kaufmann, et al., 2009)

4.3.3 AtHSFA1b binds a unique form of HSE:

The form of HSE discovered from the ChIP-SEQ datasets shows that it consists of three repeats of AGA/TCT sequences. Motif analysis showed that the structure of HSE is either **AGA**RRR**AGA**RRR**AGA** or **TCT**YYY**TCT**YYY**TCT** and the reason for this is because the final processed ChIP-SEQ peak sequences are in the form of single-stranded sequence. However, it has been shown in numerous studies that HSFs from all species belong to a group of proteins known as winged helix-turn-helix (WHTH) DNA binding proteins (Sakurai and Enoki, 2010; Aravind et al., 2005; Ahn et al., 2001; Liu and Thiele, 1999; Littlefield and Nelson, 1999). Structural analyses on those proteins have shown that they bind to inverted repeats of DNA sequences (Åkerfelt et al., 2010; Mishra et al., 2002; Ahn et al., 2001; Liu and Thiele, 1999). Based on that, it can be concluded that the DNA binding element of AtHSFA1b is arranged in

inverted repeats of the core motif sequences AGARRR/YYYTCT. Therefore, the general structure of HSE is in the form **TCT**NNN**AGA**NNN**TCT** and **AGA**NNN**TCT**NNN**AGA**. This structure of HSE discovered from ChIP-SEQ peak sequences is unusual when compared to the widely known pentameric repeats of HSEs (GAAnn/nnTTC) in other species. The structure of the HSE discovered from the ChIP-SEQ datasets gives a strong indication that AtHSFA1b binds to its target HSE in trimeric form *in vivo* where each AGA/TCT unit serves as a core binding motif for each of AtHSFA1b monomer (figure 4.16)

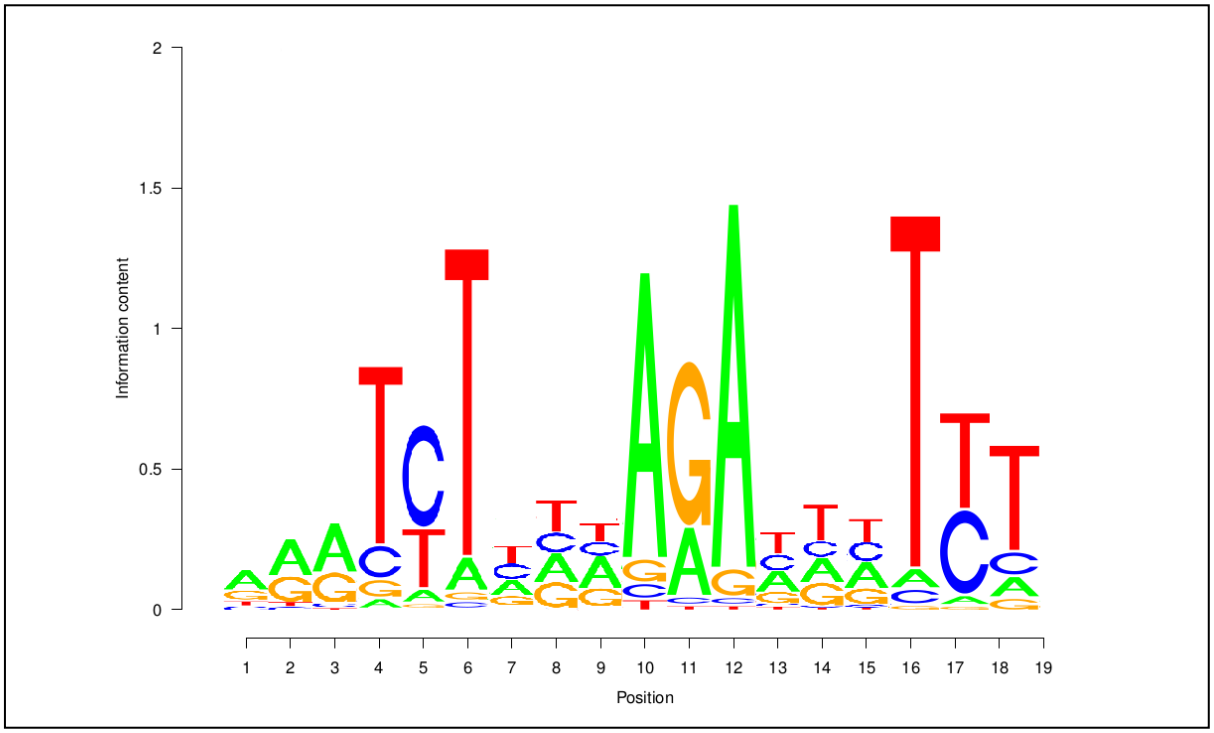


Fig.4.16. The final suggested form of the HSE recognised by AtHSFA1b. The suggested HSE motif was based on the output of *de novo* motif analysis, pattern matching and the structure and binding behaviour of closely related HSFs from other species.

The presence of HSEs on the promoters of target genes varies depending on the target genomic feature itself. Further investigation on the promoters of target genes of AtHSFA1b revealed that some of those targets contain extremely GA/TC- rich repeats in their promoter regions. Interestingly, target genes that code for HSPs fall into that category. This suggests that the binding instances of AtHSFA1b on those promoters might be higher than those on promoters of other target genes with less GA/TC content. This phenomenon is known as homotypic clustering where a single promoter can possess several binding sites for one TF (Whitefield et al., 2012; Gotea et al., 2010). It has been shown that homotypic clusters have a strong influence on the levels and the noise of gene expression of target genes (Ezer et al., 2014).

4.3.4 Co-occurring *cis*-elements in AtHSFA1b target sequences:

The analysis showed high co-occurrence of G-Box/E-Box and Site II binding elements along with HSE under no stress condition. G-box/E-Box is a binding motif for bHLH and bZIP TFs (Qiu et al., 2009; Siberil et al., 2001; Williams et al., 1992), while Site II element is known to be a binding motif for TCP TFs (Giraud et al., 2010; Welchen and Gonzalez, 2006; Trémousaygue et al., 2003). This co-occurrence of those *cis*-elements suggests that AtHSFA1b might function cooperatively with TFs that bind to those elements under normal growth conditions. However, under heat stress, the analysis showed that G-Box/E-Box does not co-occur with HSE which indicates that there may be no involvement of TFs that bind to G-Box/E-Box under heat stress. Site II binding element remained strongly enriched under heat stress which indicates that AtHSFA1b could work cooperatively with TFs that bind to Site II element (AtTCPs) under both conditions (table 4.3). The impact of the presence of those *cis*-elements with HSEs on gene expression remains known as ChIP-SEQ does not provide any information about expression levels of target genes. However, this result suggests that G-Box/E-Box and Site II binding TFs might be part of the

AtHSFA1b transcriptional machinery under no stress condition. Under heat stress there might be no involvement of G-Box/E-Box binding TFs but Site II binding proteins may remain essential in the AtHSFA1b transcriptional regulation under both conditions (Niu et al., 2014; Oh et al., 2012; Kaufmann et al., 2009).

A question that may arise from the outcome of the ChIP-SEQ experiment is; do all of those binding events lead to changes in expression levels of target genes? Due to the nature of ChIP-SEQ experiment, it is not possible to give an answer to that question. A number of studies showed that not all binding events of TFs result in changes in expression levels of target genes. For instance, it has been shown that the TF AtHY5 binds to 9000 targets in the *A. thaliana* genome; however, combining binding data from ChIP-chip with expression data from microarray showed that AtHY5 was unable to cause any changes in the expression levels of the majority of its target genes (Zhang et al., 2011). Furthermore, it was also shown that expression levels of target genes of AtHY5 were variable depending on the experimental conditions used in the study (Zhang et al., 2011). Therefore, one can predict that not all of the binding events of AtHSFA1b will lead to changes in expression of the target genes for a number of reasons. The first is that gene expression is a complex process and it does not rely solely on one TF binding to an element on the promoter of target genes. A number of factors contribute to changes in gene expression, including epigenetic changes (Jaenisch and Bird, 2003), other TFs bound on the same promoters whether they are activators or repressors (Lee and Young, 2013; Juven-Gershon and Kadonaga, 2010). Co-activators and co-repressors also play a central role in determining gene expression levels (Lee and Young, 2013; McKenna and O'Malley, 2010).

It was reported by Santoro et al., (1998) that the architecture of HSE itself can have a direct influence on gene activation upon binding of yHSF in yeast. They showed that the longer the space between the three core consensus sequences (GAA/TTC) the less the chance for gene activation. However, this observation was based on a simple artificial reporter gene system that does not take into account the complexity of eukaryotic DNA structure and the TF studied was yHSF is considerably different from AtHSFA1b in structure.

4.3.5 AtHSFA1b might be more than just an activator of HSR:

The output of the ChIP-SEQ experiment showed that AtHSFA1b targets not only *AtHSP* and stress response genes but it targets genes that are involved in a variety of cellular functions under normal growth conditions. This is consistent with what has been published about the involvement of HSFs in the other cellular processes besides controlling the expression of HSPs in yeast, *D. melanogaster* and mammals (Mendillo et al., 2012; Guertin and Lis, 2010; Hahn et al,. 1999). Although a large number of AtHSFA1b target genes are stress responsive genes, the data shows that those are not the only targets of AtHSFA1b. There is high functional enrichment for groups of genes that are involved in various cellular processes including, developmental processes, primary and secondary metabolism. The results suggest that AtHSFA1b might be involved in the regulation of numerous cellular processes not only HSR. This is consistent with what has been published about the broad functions of HSFs in other organisms such as yeast, *D. melanogaster* and mammals beyond their involvement in the regulation of stress response (Sections 1.1.2.1, 1.1.2.2, and 1.1.2.3).

Strikingly, gene ontology analysis showed that AtHSFA1b targets three groups of genes with distinct molecular functions. Those three groups of genes were TFs, genes that code for

enzymes with kinase activity and genes that code for enzymes with glycosyl transferase activity. Surprisingly, a poor overlap in the biological functions of those groups of genes was observed. The TFs group targeted by AtHSFA1b showed enrichment for developmental processes as well as abiotic and biotic stress response. Examples of TFs involved in the regulation of developmental processes are shown in table 4.5.

AGI code	TF name	Description	References
AT3G50060	*AtMYB77*	Encodes a member of the R2R3 transcription factor gene family. Expressed in response to potassium deprivation and auxin. Involved in lateral root development. Interacts with ARF7 and regulates the expression of some auxin responsive genes.	Shin et al., 2007
AT1G79840	*AtGL2*	Glabra 2, a homeodomain protein affects epidermal cell identity including trichomes, root hairs, and seed coat. It also down-regulates seed oil content. Expressed in atrichoblasts and required to suppress root hair development. Also expressed abundantly during early seed development. Directly regulated by WER.	Rerie et al., 1994
AT5G10510	*AtAIL6*	Encodes an AP2-domain transcription factor involved in root stem cell identity and root development. It is also required to maintain high levels of PIN1 expression at the periphery of the meristem and modulate local auxin production in the central region of the SAM which underlies phyllotactic transitions.	Krizek, 2007
AT1G69120	*AtAGL7*	Floral homeotic gene encoding a MADS domain protein homologous to SRF transcription factors. Specifies floral meristem and sepal identity. Required for the transcriptional activation of AGAMOUS. Interacts with LEAFY.Binds to promoter and regulates the expression of flowering time genes SVP, SOC1 and AGL24.	Mandel et al., 1992
AT1G53230	*AtTCP3*	Encodes a member of a recently identified plant transcription factor family that includes Teosinte branched 1, Cycloidea 1, and proliferating cell nuclear antigen (PCNA) factors, PCF1 and 2. Regulated by miR319. Involved in heterochronic regulation of leaf differentiation.	Cubas et al., 2002
AT1G68480	*AtJAG*	Encodes a putative zinc finger transcription factor that is necessary for proper lateral organ shape and is sufficient to induce the proliferation of lateral organ tissue. Together with NUB, it is involved in stamen and carpel development.	Ohno et al., 2004

Table 4.5. Examples of experimentally characterised TFs involved in plant development that are targeted by AtHSFA1b. Annotations were obtained from the *A. thaliana* database (TAIR; http://www.arabidopsis.org).

Whereas, the kinase group was mainly enriched for biological processes such as, hormone signalling and recognitions (table 4.6).

AGI code	Gene name	Description	References
AT5G16590	*AtLRR1*	Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, transmembrane receptor protein tyrosine kinase signaling pathway, response to symbiotic fungus	Choi et al., 2012. Zhou et al., 2009.
AT2G02220	*AtPSKR1*	Encodes a protein interacting with phytosulfokine, a five amino acid sulfated peptide (YIYTQ). Contains dual guanylate cyclase and kinase catalytic activities that operate in vivo.	Loivamäki et al., 2010.
AT5G46330	*AtFLS2*	Encodes a leucine-rich repeat serine/threonine protein kinase that is expressed ubiquitously. FLS2 is involved in MAP kinase signalling relay involved in innate immunity. Essential in the perception of flagellin, a potent elicitor of the defense response. FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB.	Gómez-Gómez et al., 2001. Gómez-Gómez and Boller, 2000.
AT2G18790	*AtPHYB*	Red/far-red photoreceptor involved in the regulation of de-etiolation. Exists in two inter-convertible forms: Pr and Pfr (active). Involved in the light-promotion of seed germination and in the shade avoidance response. Promotes seedling etiolation in both the presence and absence of phytochrome A. Overexpression results in etiolation under far-red light. Accumulates in the nucleus after exposure to far red light. The phosphorylation state of the Ser-86 residue of the phytochrome B molecule alters dark reversion of the molecule.	Sharrock and Quail, 1989. Johansson et al., 2014. Reddy and Finlayson, 2014.

Table 4.6. Examples of experimentally characterised genes involved and kinase activity targeted by AtHSFA1b. Annotations were obtained from the *A. thaliana* database (TAIR; http://www.arabidopsis.org).

Finally, the glycosyl transferase group of genes was enriched for cell wall processes. Examples of genes that code for proteins involved in glcosyl-transferase activity are shown in table 4.7.

AGI code	Gene name	Description	References
AT4G03210	*AtXTH9*	encodes a member of xyloglucan endotransglucosylase/hydrolases (XTHs) that catalyze the cleavage and molecular grafting of xyloglucan chains function in loosening and rearrangement of the cell wall. Gene is expressed in shoot apex region, flower buds, flower stalks and internodes bearing flowers.	Hyodo et al., 2003.
AT1G78580	*AtTPS1*	Encodes an enzyme putatively involved in trehalose biosynthesis. The protein has a trehalose synthase (TPS)-like domain but no trehalose phosphatase (TPP)-like domain. ATTPS1 is able to complement yeast tps1 mutants in vivo. The gene product modulates cell growth but not cell differentiation by determining cell wall deposition and cell division.	Wahl et al., 2013. Chary et al., 2008.
AT4G26850	*AtVTC2*	Encodes a novel protein involved in ascorbate biosynthesis, which was shown to catalyze the transfer of GMP from GDP-galactose to a variety of hexose-1-phosphate acceptors. Recessive mutation has a reduced amount of vitamin C, lower level of non-photochemical quenching, and reduced rate of conversion of violaxanthin to zeaxanthin in high light.	Linster et al., 2007.
AT3G02230	*AtRGP1*	RGP1 is a UDP-arabinose mutase that catalyzes the interconversion between the pyranose and furanose forms of UDP-L-arabinose. It appears to be required for proper cell wall formation. rgp1/rgp2 (at5g15650) double mutants have a male gametophyte lethal phenotype. RGP1 fusion proteins can be found in the cytosol and peripherally associated with the Golgi apparatus.	Delgado et al., 1998. Dhugga et al., 1997.
AT4G32410	*AtCESA1*	Encodes a cellulose synthase isomer. CESA1 mutants have cellulose defect in the primary cell wall. Multiple lines of evidence suggest that CESA1, along with CESA3 and CESA6 are present in the same plasma membrane complex for cellulose biosynthesis. lasma membrane complex for cellulose biosynthesis. As inferred from the null role of secondary wall-type CesAs, included in a set of five primary wall-type CesAs that may support trichome cell wall thickening.	Persson et al., 2007. Desprez et al., 2007.

Table 4.7. Examples of experimentally validated genes that code for protein involved in glycosyl-transferase activity that are targeted by AtHSFA1b. Annotations were obtained from the *A. thaliana* database (TAIR; http://www.arabidopsis.org).

However, this output does not mean that those processes or pathways are separate. There might be an integration of all of those processes and the end result of that integration might form the overall phenotype of plants overexpressing *AtHSFA1b*.

Plant cells undergo modifications on the structure of cell wall in response to various conditions including, growth and development, preparation for cell differentiation and during stress response. Glycosylation is major component of cell wall modification process. It was shown that enzymes the possess glycosyltransferase activity play a key role in the synthesis glycosidic linkages in the cell wall (Cosgrove, 2005). From a developmental point of view, it was shown that glycosylation of plant cell wall is a key component of cell differentiation and growth of plants (Lerouxel et al., 2005). However, the implication of cell wall modifications is not limited to plant cell differentiation and development of the plant. It is widely known that plant cell wall is the first defence line. This means that this part of plant cell is perhaps under continuous modifications in response to changes in growth conditions and physiological states. There is evidence that cell wall modifications occur under water limited conditions. These changes include decrease in the expression of genes that promote cell wall expansion (Bray, 2004). Therefore, those modifications on the cell wall controlled by AtHSFA1b could be a key component in dictating the overall phenotype of plants overexpressing *AtHSFA1b*.

What is unusual here is the loss of AtHSFA1b binding activity after 30 minutes of heat stress which gives the immediate impression that AtHSFA1b has little implication in the regulation of prolonged heat stress response in *A. thaliana*. However, this unusual binding behaviour of AtHSFA1b under heat stress supports the study that showed that AtHSFA1b is only involved in the immediate and very early phases of HSR (Lohmann et al., 2004). There are 21 HSFs in *A.*

thaliana and the expression levels of some of those HSFs, for example AtHSFA2, AtHSFA7a, AtHSFA7b, are highly responsive to heat stress (Yu et al., 2012; Sugio et al., 2009; Rizhsky et al., 2004) (Section 1.1.2.4.3). Therefore, it can be predicted that heat responsive AtHSFs might compete with AtHSFA1b on its binding sites and might eventually displace AtHSFA1b from its binding regions assuming that all HSFs in *A. thaliana* recognise and bind the same HSEs. This result might also explain why plants have large families of HSFs compared to other species. Therefore, there is an indication that AtHSFA1b may be more than just a regulator of HSR and it might has broader functions than those that are highly inducible in response to stress.

Microarray studies on plant overexpressing stress-inducible AtHSFs have shown that the majority of them induce small numbers of genes compared to the number of genes induced as a result of overexpression of *AtHSFA1b*. For example, overexpression of *AtHSFA2* resulted in upregulation 59 genes involved stress response (Bechtold et al., 2013; Ogawa et al., 2007). Similarly, overexpression of *AtHSFA3* led to the upregulation of 118 genes involved only in the stress response (Yoshida et al., 2008). The genes induced by overexpression of both *AtHSFA2* and *AtHSFA3* seem to be subsets of genes induced by overexpression of *AtHSFA1b* (figure 4.17). This shows that AtHSFA1b might be involved in broader functions than those stress-inducible AtHSFs. Furthermore, the constitutive expression nature of AtHSFA1b in wild type plants gives further evidence that it might be required for the regulation of non-stress related processes under normal growth conditions and its function under heat stress is perhaps limited to triggering the expression of some *AtHSFs* and other *AtHSP* genes at the early phases of heat stress response. The expression of those genes might then be maintained by other factors under prolonged stress. This suggests there might be shifts in roles among AtHSFs during stress. Therefore, it can be hypothesised that stress response triggers the expression of highly

specialised AtHSFs that target subgroups of AtHSFA1b and possibly other group-A1 AtHSF target genes. The ChIP-SEQ result showed some clues that there might be a reallocation of roles among the AtHSF family members in response to different growth conditions. However, this remains as a pure prediction based on the information provided from the ChIP-SEQ experiment. Only appropriate experiments can prove or disprove this theory.

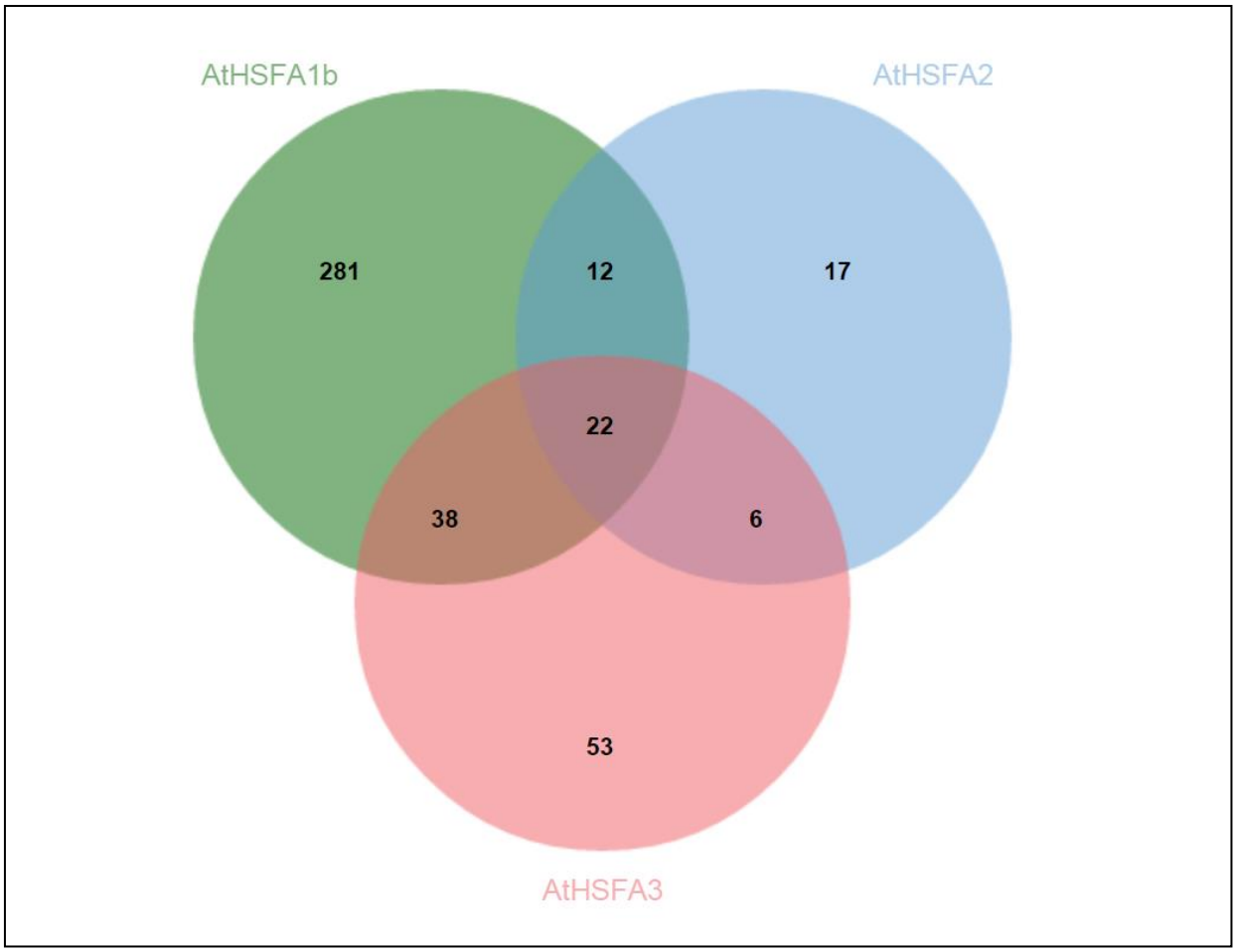


Fig.4.17. Overlap between genes controlled by AtHSFA1b, AtHSFA2, and AtHSFA3. Venn diagram showing the total number of genes controlled by each AtHSF, overlapping, and unique to each AtHSF. The figure was generated from published microarray datasets, AHSFA1b (Bechtold et al., 2013), AtHSFA2 (Bechtold et al., 2013; Ogawa et al., 2007), and AtHSFA3 (Yoshida et al., 2008).

CHAPTER 5

Analysis of the AtHSFA1b-regulated transcriptome

5.1 Introduction:

Binding of TFs to *cis*-elements on the promoters of target genes is only one step in the transcriptional regulation machinery. However, regulation of transcription in eukaryotic cells, including plant cells, is a complex process that depends on a number of factors such as the overall state of the chromatin, the accessibility of promoter regions to TFs, recruitment of RNA polymerase II complex to the TSS of target genes and the presence of activators and repressors on the same promoter of a target gene and the nature of binding of TF on the promoter of target gene whether it is transient or prolonged (Kaufmann et al., 2009; Gao et al., 2004). There are three possible scenarios that may result from binding of TFs to *cis*-elements on the promoters of its target genes. However, the importance of binding events of TFs on the promoters of their target genes generally comes when the binding results in changes in the rate of expression of target genes (Tate and Bird, 1993). Changes in expression levels of target genes give indication of possible changes in the amount of proteins coded from target genes which in turn may influence the overall biology (Fabian et al., 2010; Pavelka et al., 2010).

The availability of high throughput methods has allowed for massive simultaneous scanning of changes in transcript levels of thousands of genes (Morozova et al., 2010; Rita, et al., 2008). The benefit of such methods is that they allow for examining changes in the total transcriptome in an organism which by turn could allow for a more accurate prediction of the changes in the overall amounts of functional proteins and subsequently relate that to the overall physiology of an organism (Reddy et al., 2014; Tacchi et al., 2012; Folsom et al., 2010; Desikan et al., 2001). The development of high throughput sequencing methods, in particular, allowed for a more precise and unbiased monitoring of changes in transcriptomes compared to the aging array methods (Fu et al., 2010; Wang et al., 2009; Marioni et al., 2008).

The mapping of the AtHSFA1b binding profile in the *A. thaliana* genome revealed valuable information about the potential involvement of AtHSFA1b in non-stress related processes including developmental processes and the possible reconfiguration of its transcriptional regulatory network in response to heat stress (Chapter 4). However, the output of ChIP-SEQ experiment raised a number of questions equal to the number of answers it provided. One of the peculiar observations was the loss of AtHSFA1b bindings to target regions in the *A. thaliana* genome including the promoters of target genes (Chapter 4). Do all of the non-stress associated genes bound by AtHSFA1b show change in expression levels? What is the implication of the loss of AtHSFA1b binding on the expression levels of target genes? The answers to these questions are beyond the scope of an experiment such as ChIP-SEQ as it does not provide any information about changes in expression levels of target genes. So far, identifying AtHSFA1b binding sites in the genome provided only part of the AtHSFA1b regulatory network picture.

This part of the study was set out to analyse the AtHSFA1b-regulated transcriptome under two conditions, no stress and heat stress. The aim of this part is to examine the changes that occur in the *A. thaliana* transcriptome as a result of *AtHSFA1b* overexpression and analyse the transcriptomic changes in *AtHSFA1b* overexpressing plants compared to wild type plants in response to heat stress. Furthermore, this part of the research is aimed to examine the relationship between binding of AtHSFA1b to target genes and subsequent effects on the rate of expression of those genes and the effect of the loss of binding of AtHSFA1b to target genes under heat stress on the expression levels.

5.2 Results

5.2.1 Overview of 35S-AtHSFA1b::mRFP differentially expressed genes:

Overexpression of *AtHSFA1b* in 35S-AtHSFA1b::mRFP rosettes resulted in 896 differentially expressed genes (DEGs) compared to wild type under no stress. However, heat stress treatment on both 35S-AtHSFA1b::mRFP and Col-0 wild type plants altered the expression of 3380 genes and 2883 respectively (table 5.1)

Plant	Total no. DEGs	Upregulated	Downregulated
35S-AtHSFA1b::mRFP no stress	896	798	98
35S-AtHSFA1b::mRFP heat stress	3380	1664	1716
Wild type heat stress	2883	1765	1118

Table 5.1. Summary of the numbers of DEGs in each treatment compared to wild type under no stress. Three biological replicates of each plant were subject to RNA-SEQ analysis. The table shows the total numbers of DEGs compared to unstressed wild type control in each plant under no stress and heat stress (37°C for 30 minutes) conditions and a breakdown of the numbers of DEGs to upregulated and downregulated. The cutoffs used to determine DEGs were, the mean of 2-fold change from each experiment, *p*-value ≤ 5^{-3} and *q*-value ≤ 0.05.

Interestingly a large overlap was observed between upregulated genes in plants overexpressing *AtHSFA1b* under no stress conditions and those that are upregulated by heat stress treatment in *AtHSFA1b* overexpressing and wild type plants (figure 5.1). Furthermore, a significant overlap in downregulated genes was observed between 35S-AtHSFA1b::mRFP and wild type plants under heat stress. Furthermore, 49% of downregulated genes in 35S-AtHSFA1b::mRFP plants under no stress overlap with downregulated genes in both wild type and 35S-AtHSFA1b::mRFP plants under heat stress (figure 5.2).

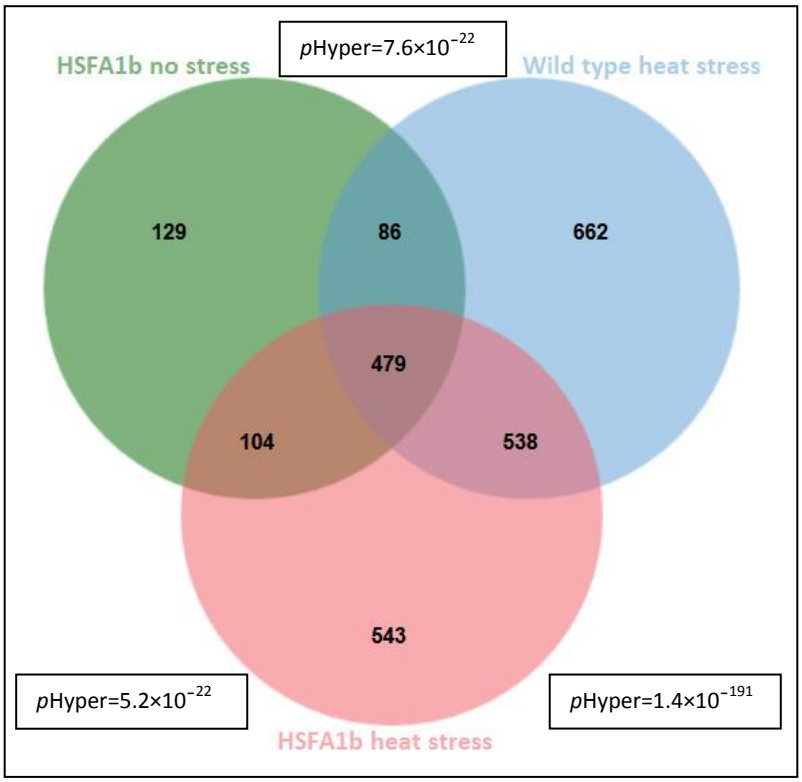


Fig.5.1. Overlap between upregulated genes in all plants tested under both conditions. Venn diagram showing the total numbers of upregulated, overlapping and unique genes in each treatment with the corresponding hypergeometric *p*-values.

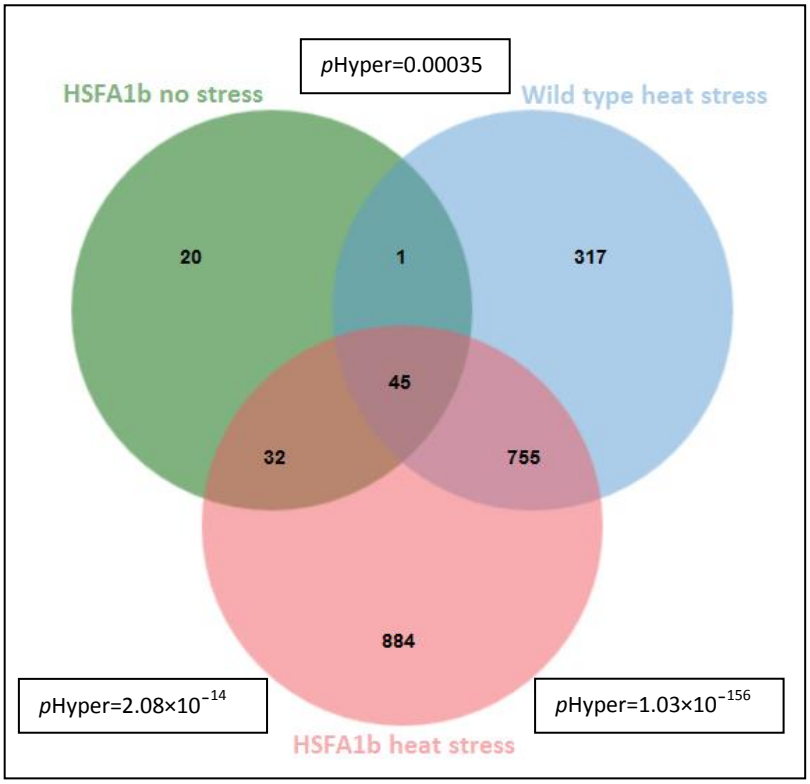


Fig.5.2. Degree of overlap between downregulated genes in all plants tested under both conditions. Venn diagram showing the total numbers of downregulated, overlapping and unique genes in each treatment with the corresponding hypergeometric *p*-values.

5.2.2 *AtHSFA1b* overexpressing plants show a partial heat stress transcriptome:

Principal component analysis (PCA) and multidimensional scaling (MDS) analysis on the total transcriptomes of plants tested showed that heat stressed 35S-AtHSFA1b::mRFP plants have almost identical expression profile to wild type plants under the same condition. Under no stress condition, however, the transcriptome of 35S-AtHSFA1b::mRFP plants showed a partial heat stress expression profile (figure 5.3).

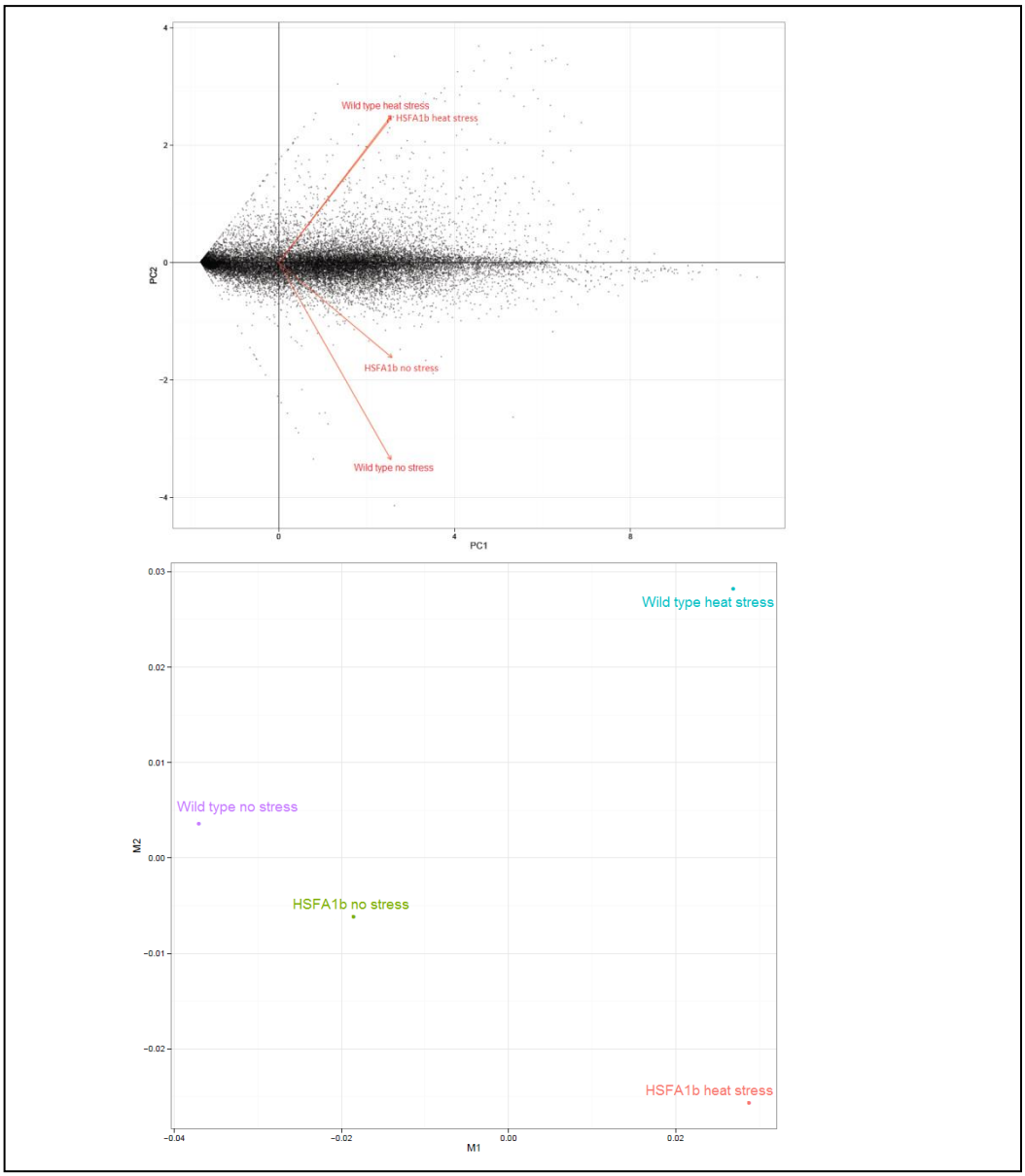


Fig.5.3. Overexpression of AtHSFA1b leads results is partial heat stress expression profile under non-stress conditions. Principal component analysis (PCA; top panel) and Multidimensional scaling (MDS; bottom panel) plots showing the correlation and the variation between the total transcriptome in wild type and 35S-AtHSFA1b::mRFP plants under no stress and wild type and 35S-AtHSFA1b::mRFP under heat stress.

PCA and MDS analyses only give very general overview of the degrees of correlation and variability between the datasets. Further analysis using Pearson's correlation method (figure 5.4) on the datasets from both treatments revealed that the expression profile of plants overexpressing *AtHSFA1b* under no stress fall in an area in between unstressed and heat stressed plants. There is a 92% correlation between the expression profiles of 35S-AtHSFA1b::mRFP plants and wild type under no stress conditions. However, there is also a high correlation between the expression profile of 35S-AtHSFA1b::mRFP plants under no stress, wild type and 35S-AtHSFA1b::mRFP plants under heat stress. The correlation in the expression profile between wild type under no stress and wild type and 35S-AtHSFA1b::mRFP plants under heat stress is 66% and 64% respectively. The data also showed 10% variation between 35S-AtHSFA1b::mRFP and wild type expression profile under heat stress.

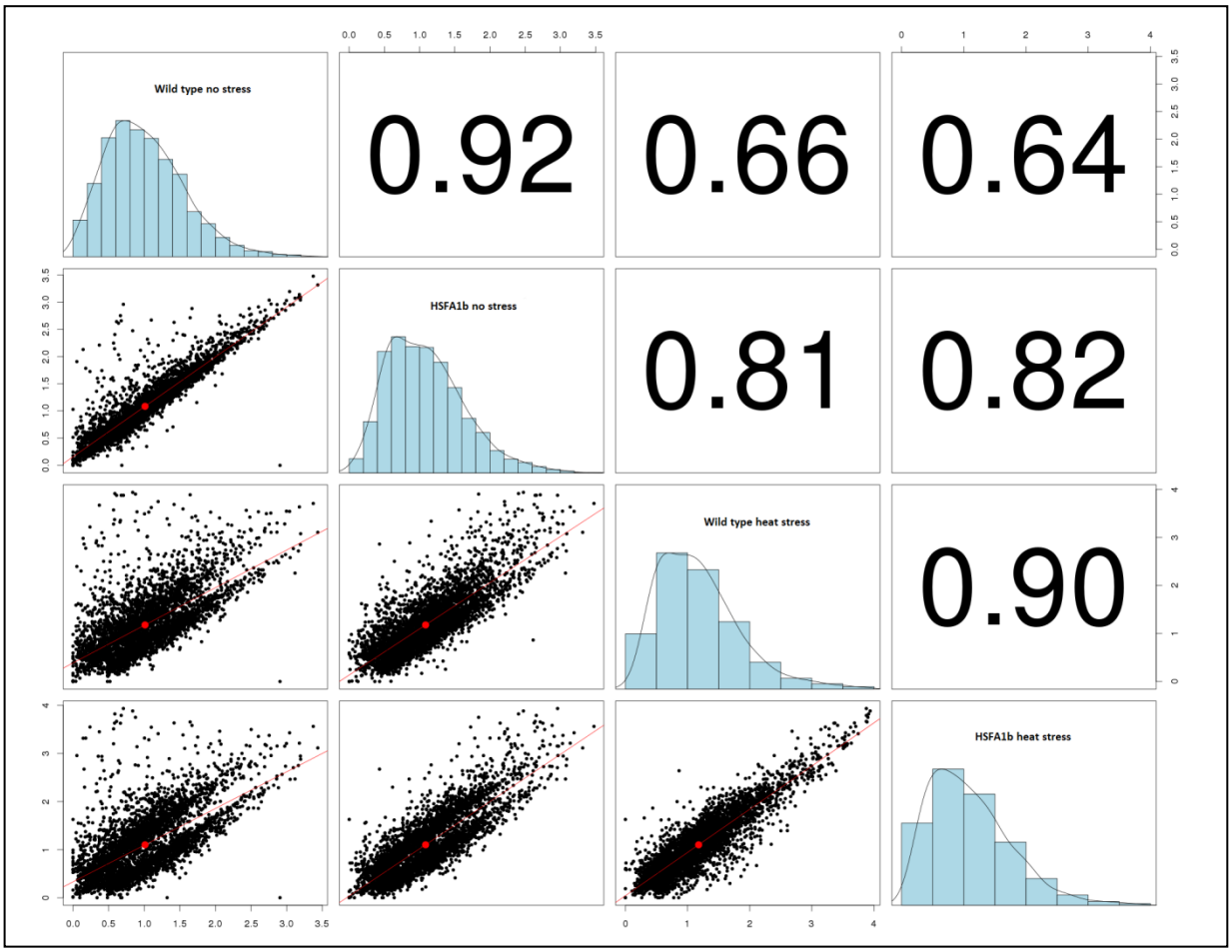


Fig.5.4. AtHSFA1b overexpressing plants exhibit partial heat stress transcriptome under normal growth conditions. Multiple correlation test on the expression profiles of wild type and 35S-AtHSFA1b::mRFP plants under no stress and heat stress conditions. Histograms with smoothing lines show the overall distribution of the data. Numbers represent Pearson's R value of correlation and scatter plots show the degree of dispersion in the data between each dataset in the comparison.

5.2.3 Functional analysis of AtHSFA1b-regulated transcriptome:

Gene ontology analysis on the upregulated genes in the 35S-AtHSFA1b::mRFP plants under no stress condition revealed that the highest enriched groups of genes affected by overexpression of *AtHSFA1b* are involved in abiotic and biotic stress (figure 5.5).

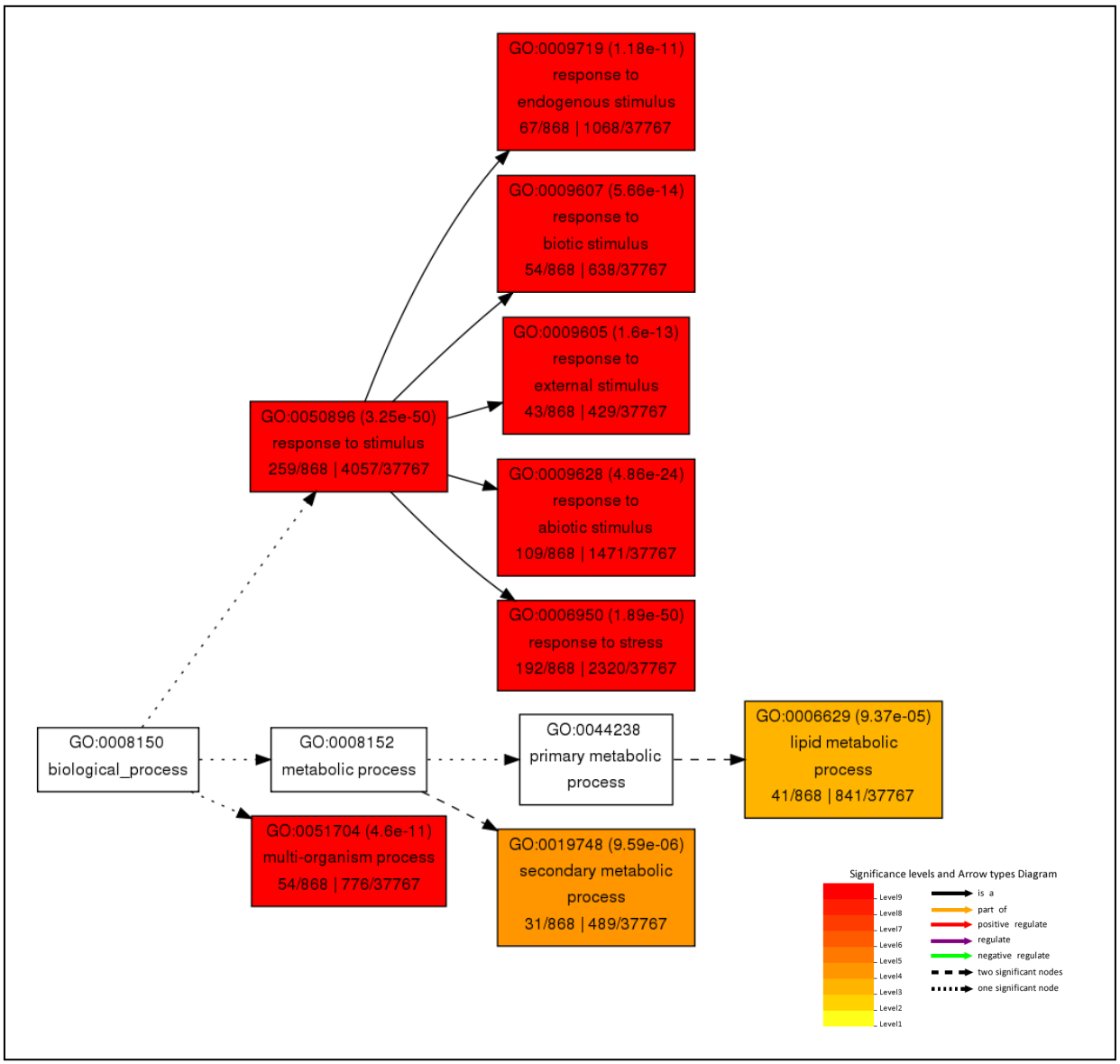


Fig.5.5. Overexpression of AtHSFA1b induces the expression of genes involved stress response under no stress conditions. Gene ontology enrichment map in parent/child relation of the biological processes of groups of genes upregulated as a result of overexpression of *AtHSFA1b::mRFP* under non-stress conditions. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Benjamini-Hochberg *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis was Benjamini-Hochberg adjusted *p-value* ≤ 0.01.

A Similar result was observed on upregulated genes in wild type plants under heat stress. Upregulated genes from heat stressed wild type plants showed high enrichment for groups of genes mainly involved in abiotic and biotic stress responses. However, there were other significantly enriched groups of genes such as, genes that are associated with cell death and transcription factors (figure 5.6).

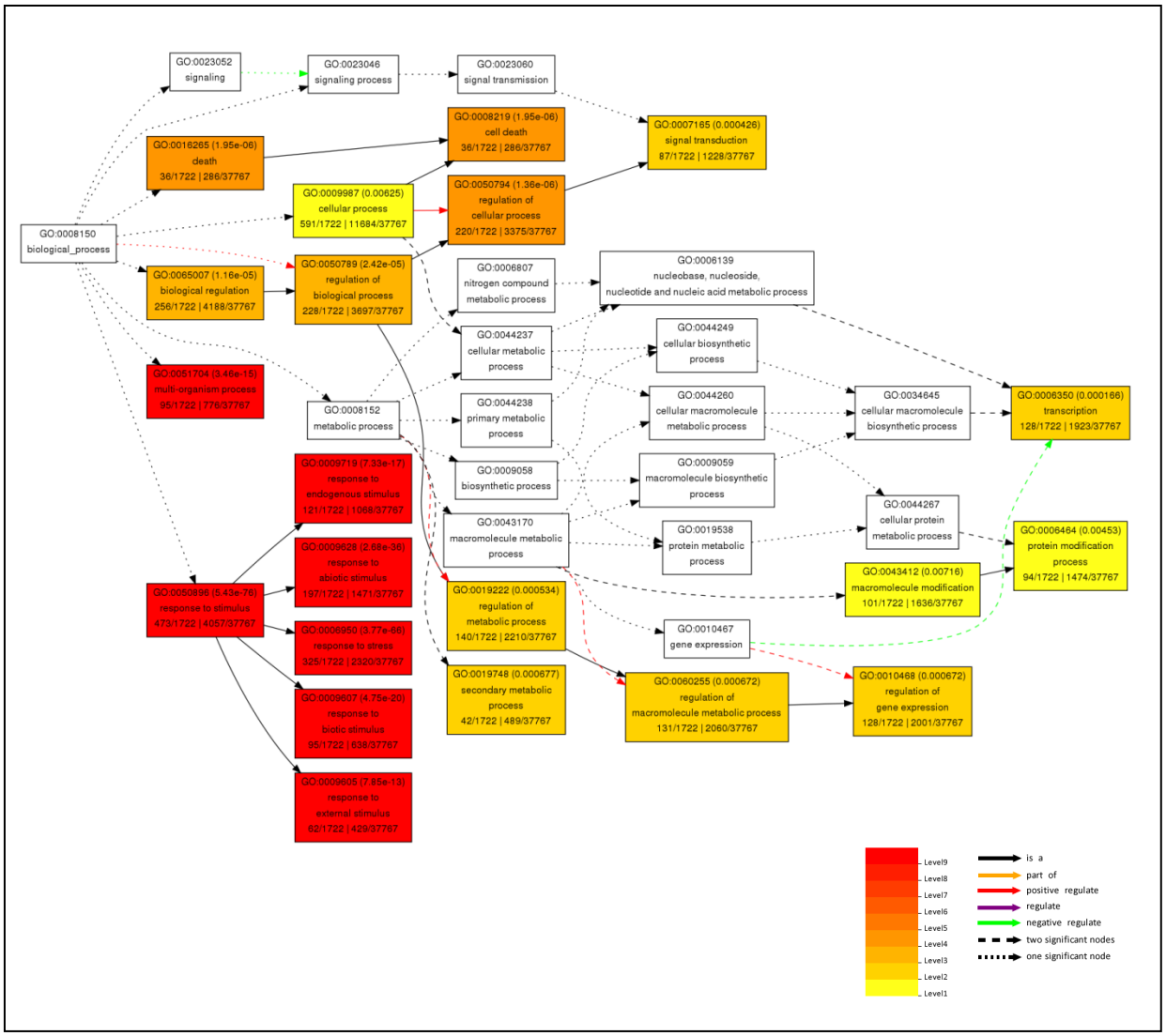


Fig.5.6. Heat stress treatment on wild type plants activates the expression of genes involved in stress response. Gene ontology enrichment map in parent/child relation of the biological processes of groups of genes upregulated as a result of applying heat stress (37°C for 30 minutes) on wild type plants. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Benjamini-Hochberg *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis was Benjamini-Hochberg adjusted *p-value* ≤ 0.01.

Heat stressed 35S-AtHSFA1b::mRFP plants showed very similar enriched groups of genes to those in heat stressed wild type plants. The analysis showed high enrichment for biological processes such as response to abiotic and biotic stress responses and cell death (figure 5.7)

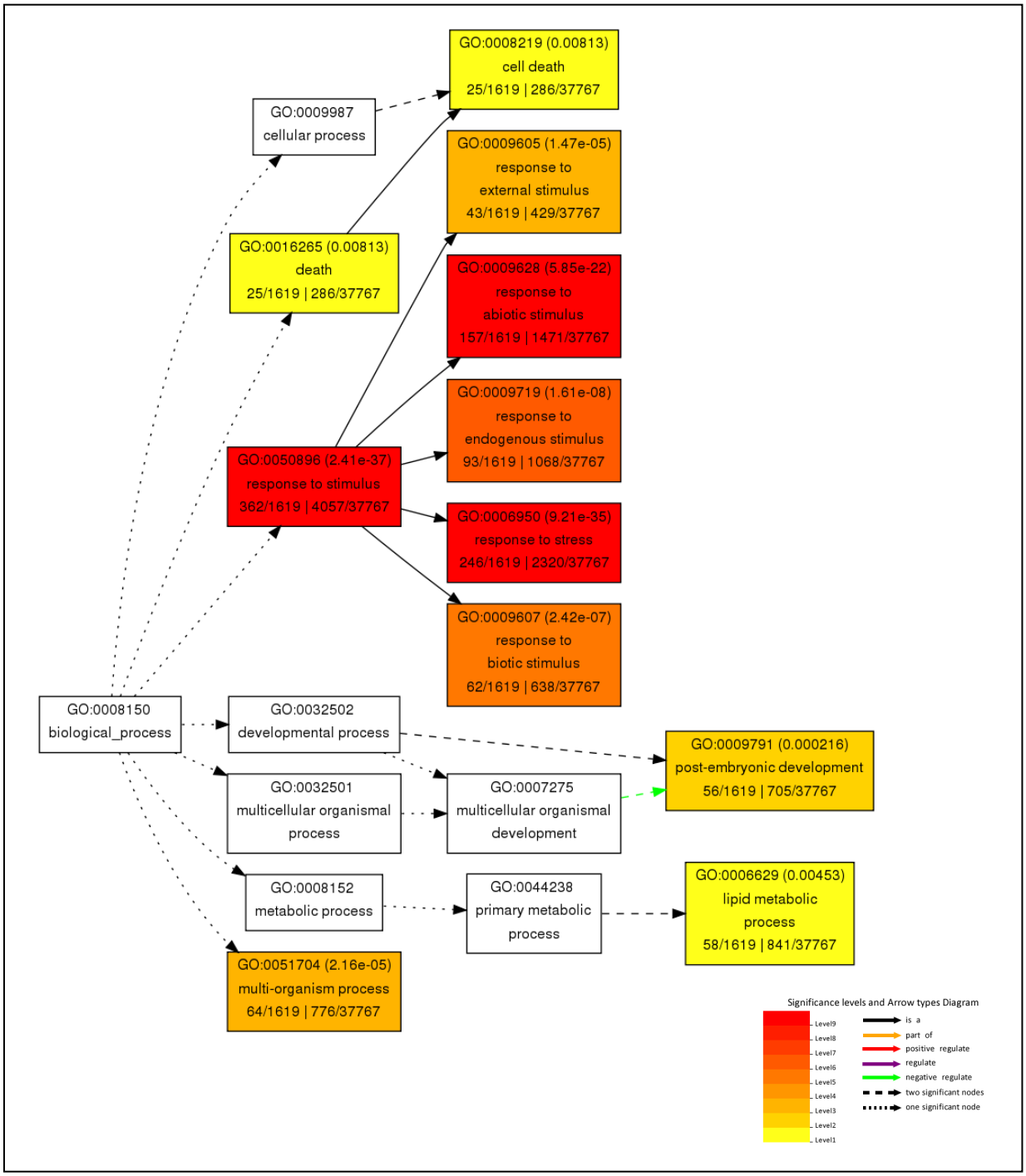


Fig.5.7. Heat stress treatment of 35S-AtHSFA1b::mRFP plants increases the expression of stress response genes. Gene ontology enrichment map in parent/child relation of the biological processes of groups of genes upregulated as a result of applying heat stress (37°C for 30 minutes) on plants overexpression of *AtHSFA1b::mRFP*. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Benjamini-Hochberg *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis was Benjamini-Hochberg adjusted *p-value* ≤ 0.01.

One interesting observation about heat stressed 35S-AtHSFA1b::mRFP and wild type plants DEGs compared to wild type under no stress was the number of downregulated genes. Short heat stress treatment on both 35S-AtHSFA1b::mRFP and wild type plants resulted in the downregulation of 1716 and 1118 genes respectively (table 5.1). Gene ontology analysis on downregulated genes in both 35S-AtHSFA1b::mRFP and wild type plants under heat stress showed a large functional overlap between them. Downregulated genes in both 35S-AtHSFA1b::mRFP and wild type under heat stress showed high enrichment for various biological processes including, stress response and developmental processes. The developmental groups of genes seemed at first to be enriched only in heat stressed 35S-AtHSFA1b::mRFP plants. However, those groups did appear in heat stressed wild type plants when a more relaxed statistical cutoff was applied (p-value ≤ 0.005, Benjamin-Hochberg FDR ≤ 0.05). Strikingly, one of the highest enriched functional groups gene ontology analysis showed in both 35S-AtHSFA1b::mRFP and wild type under heat stress was ‘transcription’ (GO:0006350) (figures 5.8 and 5.9).

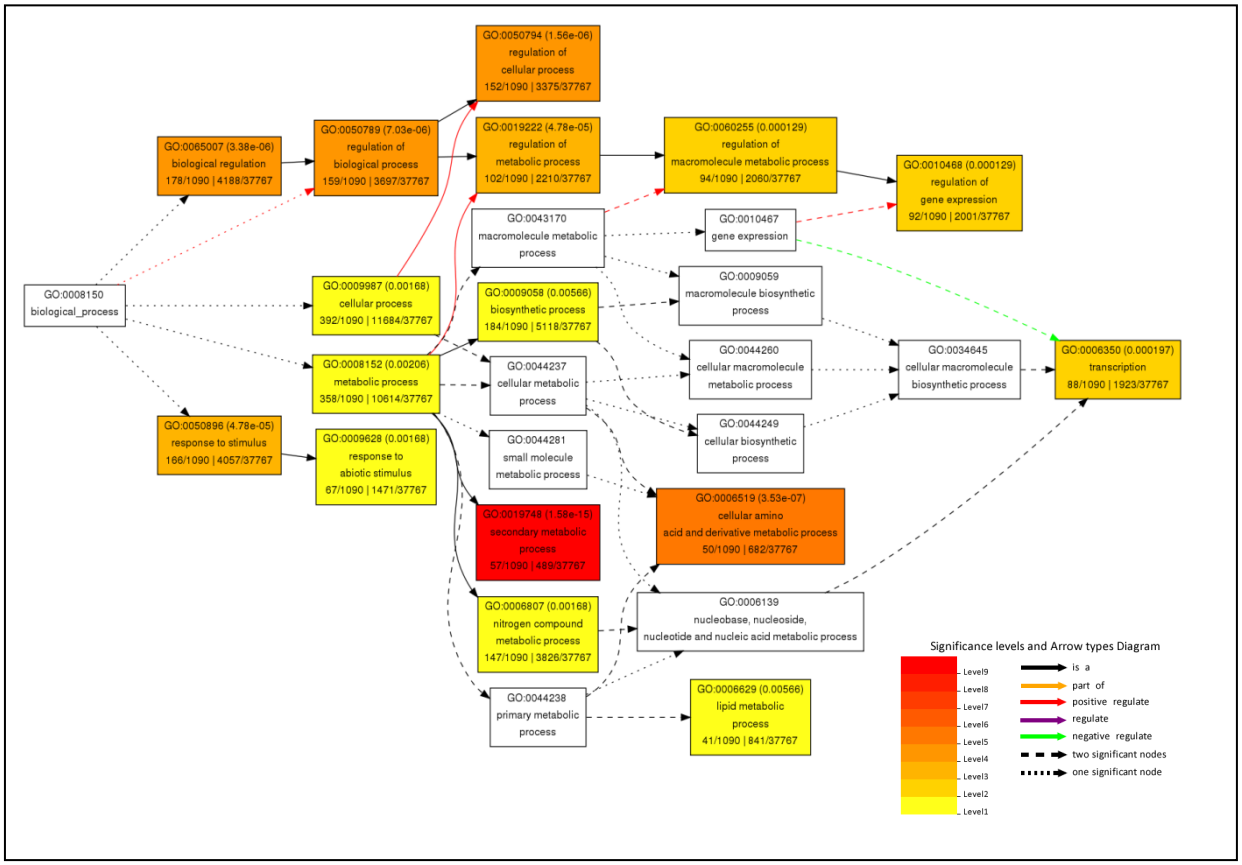


Fig.5.8. Heat stress on wild type plants results in downregulation of genes involved in various functions. Gene ontology enrichment map in parent/child relation of the biological processes of groups of genes downregulated as a result of applying heat stress on wild type plants. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Benjamini-Hochberg *p-value* (red is highest significant and white is least significant). The cutoff used in gene ontology analysis was Benjamini-Hochberg adjusted *p-value* ≤ 0.01.

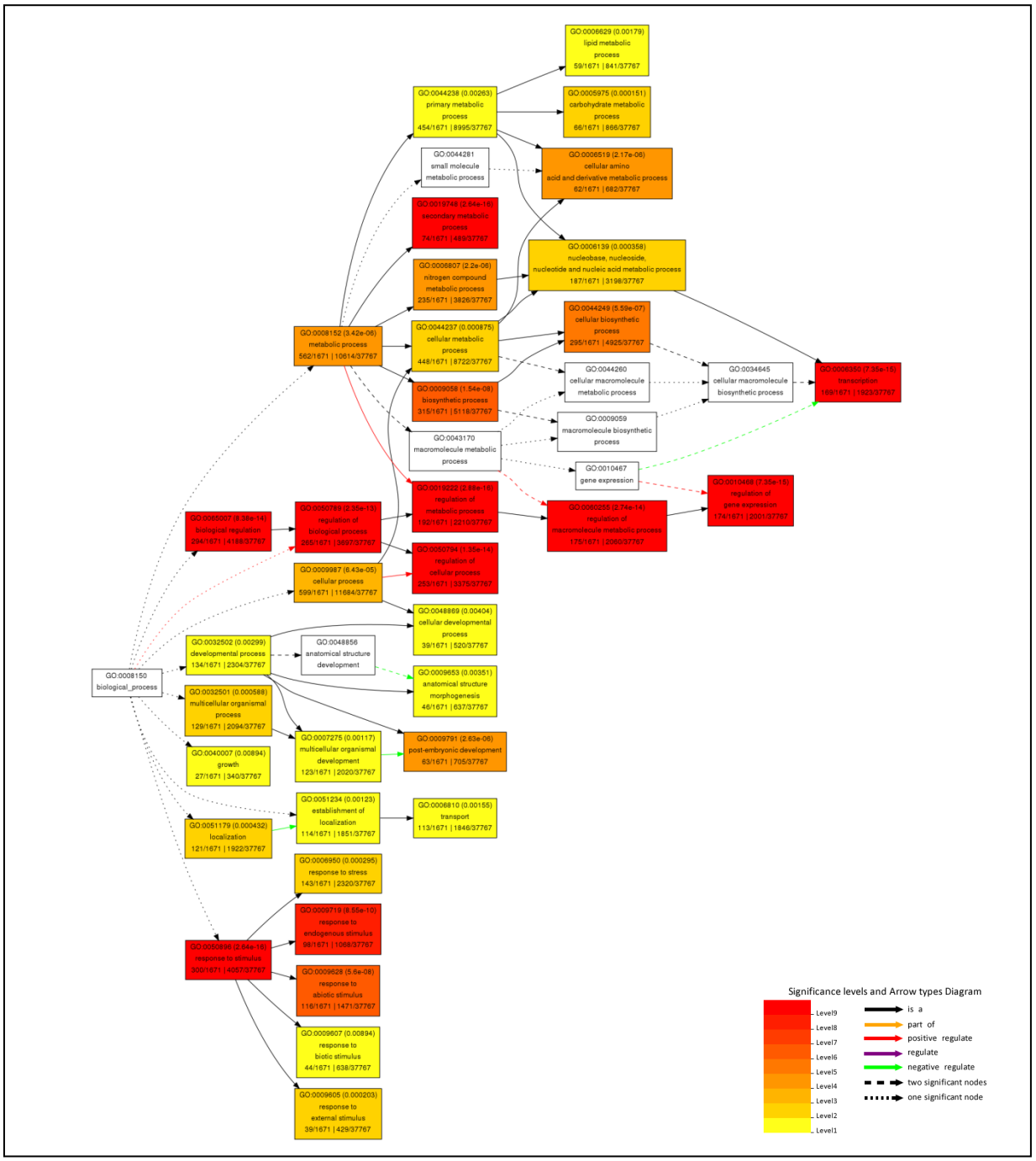


Fig.5.9. Applying heat stress on 35S-AtHSFA1b::mRFP plants results in downregulation of genes involved in various functions. Gene ontology enrichment map in parent/child relation of the biological processes of groups of downregulated genes in heat stressed 35S-AtHSFA1b::mRFP. Each group of genes is assigned a level of significance based on its *p-value* and adjusted Benjamini-Hochberg *p-value* (red is highest significant and white is least significant). The statistical cutoff used was Benjamini-Hochberg adjusted *p-value* ≤ 0.01.

Further gene ontology analysis on the molecular function of downregulated genes in both 35S-AtHSFA1b::mRFP and wild type plants under heat stress revealed two highly significant groups. The first was a group of genes involved in 'glycosyltransferase activity' (GO:0046527) which is directly linked to 'UDP-glycosyltransferase activity' (GO:0035251) and 'cellulose synthase activity' (GO:0016759). The second highest enriched group was 'TF activity' (GO:0003700). The analysis showed that there were 220 downregulated transTFs from the total 1716 downregulated genes in 35S-AtHSFA1b::mRFP plants and 106 downregulated TFs from the 1118 downregulated genes in wild type under heat stress. Surprisingly, examining the biological functions of the downregulated TFs in heat stressed 35S-AtHSFA1b::mRFP and wild type plants revealed that the vast majority of the downregulated TFs are extremely enriched for the biological process growth and developments (p-value = 4.12×10^{-14} in wild type) and (p-value = 5.14×10^{-24} in 35S-AtHSFA1b::mRFP) and other associated processes beside response to chemical stimuli (figure 5.10).

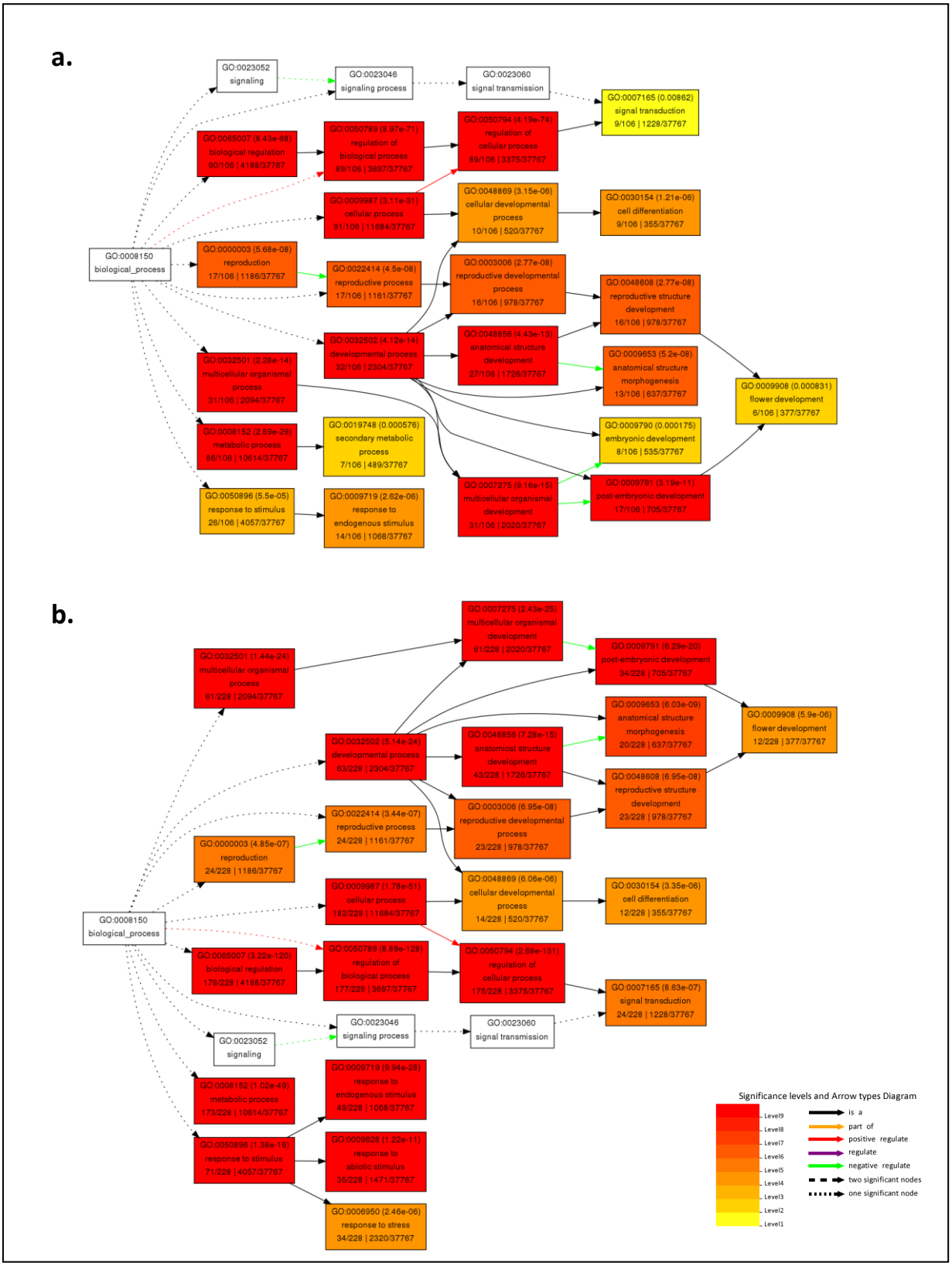


Fig.5.10. Heat stress treatment results in downregulation of TFs involved in growth and development. Gene ontology enrichment map of biological functions of downregulated TFs as a result of applying heat stress (37°C for 30 minutes. **(a.)** In wild type plants and (**b.**) in 35S-AtHSFA1b::mRFP plants. Downregulated TFs in heat stressed plants are highly enriched for developmental processes (Benjamini-Hochberg adjusted *p-value* ≤ 0.01).

5.2.4 Promoter motif analysis:

Motif analysis on the promoter regions of the DEGs in *AtHSFA1b* overexpressing plants under no stress compared to wild type under the same condition resulted in a number of HSE candidates with different structures. However, the HSEs discovered in the analysed promoters were structurally overlapping and it was not possible to distinguish which one is the preferred HSE for AtHSFA1b. The same overlapping HSEs were also discovered in the promoters of upregulated genes in wild type and 35S-AtHSFA1b::mRFP plants under heat stress (figure 5.11)

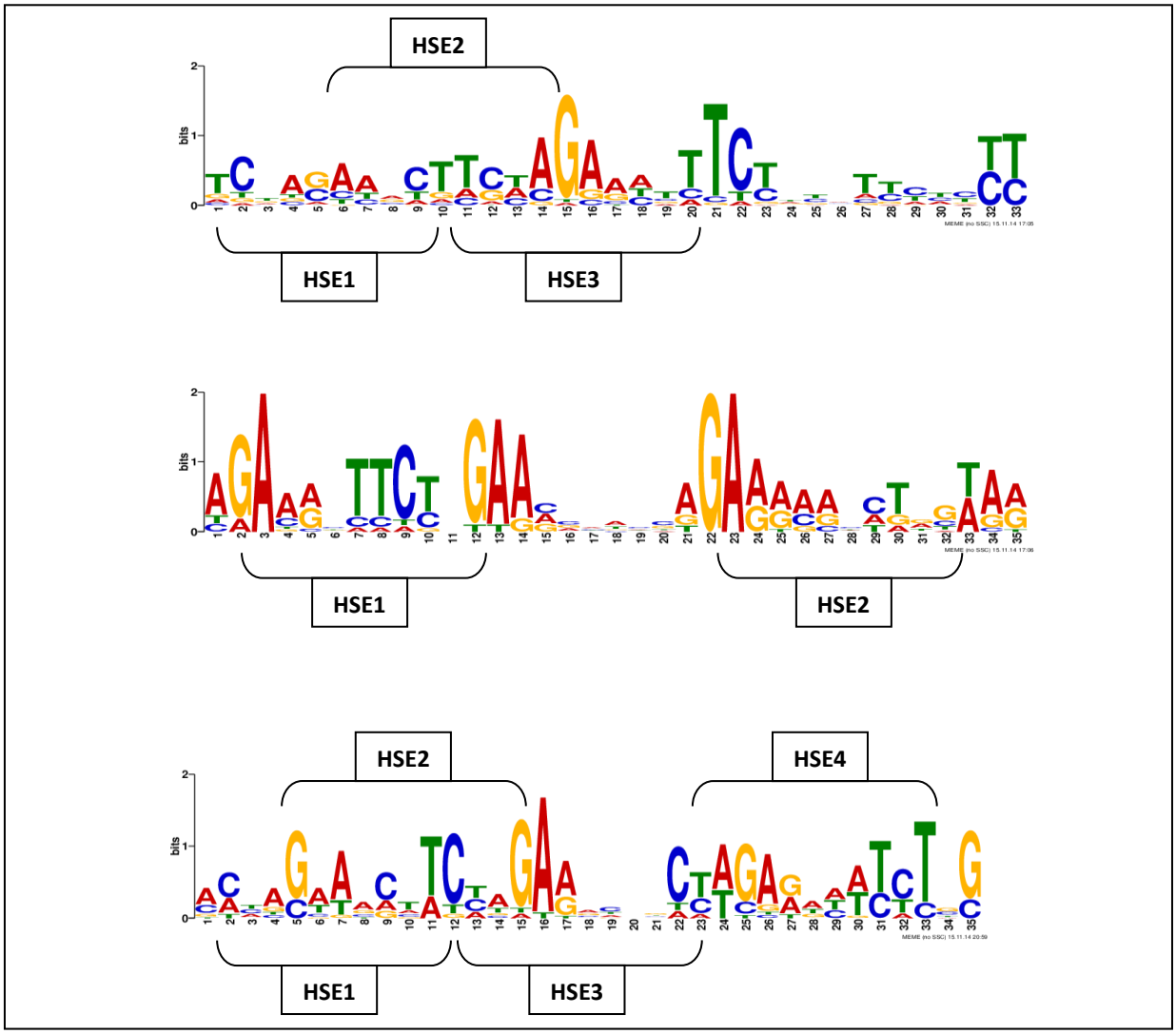


Fig.5.11. **The promoters of upregulated genes in unstressed 35S-AtHSFA1b::mRFP contain various forms of overlapping HSEs.** Upstream promoter sequences (500 bp from the translational start site) of all upregulated genes were used in *de novo* motif analysis to identify enriched *cis*-elements. The analysis showed that a multiple overlapping HSEs in the promoters of upregulated genes as a result of overexpression of *AtHSFA1b::mRFP* (*p*-value < 0.0001).

The same promoters analysed which showed the presence of HSEs also showed a strong presence of Site II binding element (*p*-value < 0.0001). Site II element was present along with HSEs within 500bp from TSS in the promoters of the top 100 DEGs in 35S-AtHSFA1b::mRFP plants under no stress. The analysis also showed high enrichment for G-Box/E-Box element on the promoters of upregulated genes. Unlike Site II, G-Box/E-Box binding elements were not found in promoter regions containing HSEs. However, it was discovered in the promoters of DEGs that do not contain HSEs.

Under heat stress, site II binding element was again highly enriched in the promoters that contain HSEs. Interestingly, G-box/E-Box binding element was not present in the promoters of DEGs in both wild type and 35S-AtHSFA1b::mRFP under heat stress. However, a new *cis*-element, C-Box, was significantly enriched (*p*-value < 0.0001) in the promoters of upregulated gene in heat stressed wild type and 35S-AtHSFA1b::mRFP. Similar to G-Box/E-Box, C-Box binding element was only present in the promoter regions of genes that do not contain HSEs (table 5.2).

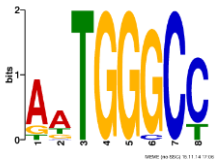
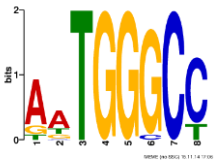
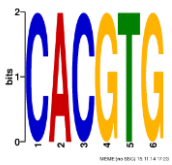
No stress		**Heat stress**	
Motif	**Annotation**	**Motif**	**Annotation**
	Site II element		Site II element
	G-Box/E-box		C-Box (Kaminaka, et al., 2006)

Table 5.2: Other *cis*-elements enriched in the promoters of upregulated genes in 35S-AtHSFA1b::mRFP. Promoters of upregulated genes in all plants under all conditions were analysed for the presence of other *cis*-elements using *de novo* motif discovery. The statistical cutoff for the analysis was *p*-value<0.0001.

5.2.5 Integrating ChIP-SEQ with RNA-SEQ:

Integrating the RNA-SEQ data with ChIP-SEQ revealed that not all of the DEGs in 35S-AtHSFA1b::mRFP plants under both conditions are direct targets of AtHSFA1b. The intersected data showed that AtHSFA1b binds the promoters of 464 of the 798 upregulated genes under no stress and targets only 100 genes out of the 1664 upregulated genes in 35S-AtHSFA1b::mRFP plants under heat stress (figure 5.12).

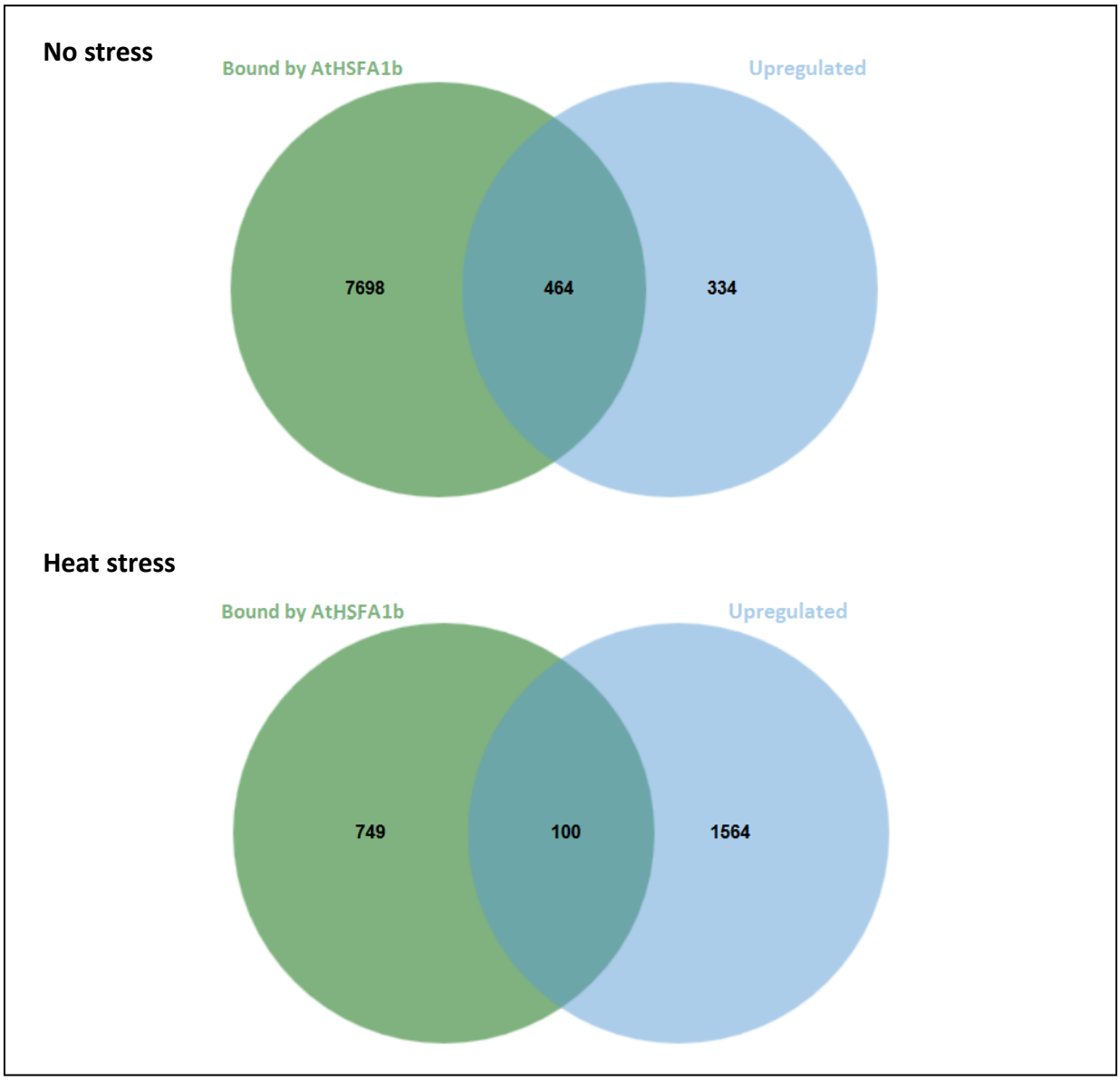


Fig.5.12. Not all upregulated genes are directly controlled by AtHSFA1b under both conditions. Venn diagrams showing a summary of the numbers of genes targeted by AtHSFA1b, the number of genes upregulated under both conditions and the overlap between them.

Further analysis on AtHSFA1b direct targets that were upregulated under no stress condition showed that AtHSFA1b controls the expression of 47 TF genes that belong to different families including, *AtHSF*, *AtAP2* domain, *AtAP3* domain, *AtbHLH*, *AtbZIP*, MADS box, *AtJAZ* and zinc finger TFs (See appendix C for a detailed list of TFs that are bound AtHSFA1b and differentially expressed under no stress). Gene ontology analysis showed that the TF genes targeted by AtHSFA1b and upregulated under no stress are involved primarily in stress defence and signal transduction (figure 5.13)

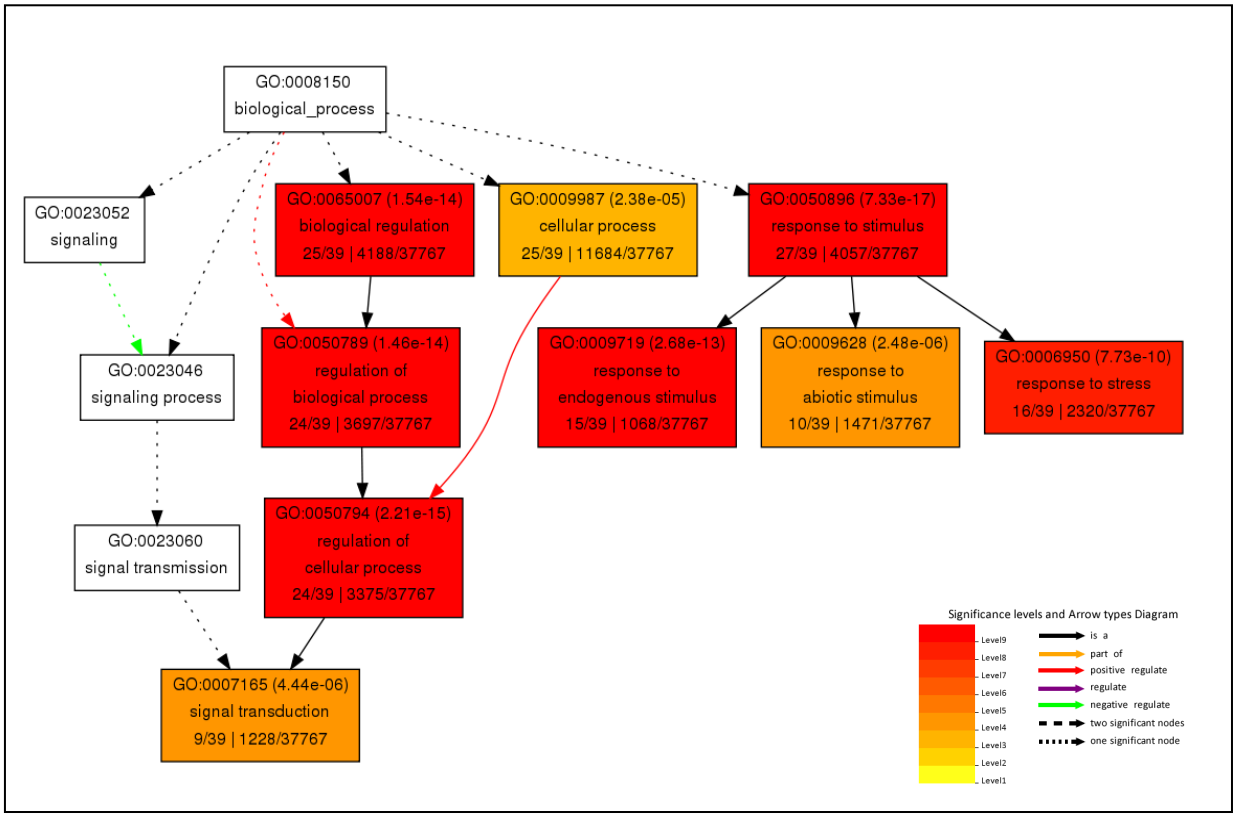


Fig.5.13. TFs bound by AtHSFA1b and upregulated under no stress are enriched for stress response. Gene ontology enrichment map of the biological process of the TFs targeted by AtHSFA1b and upregulated in plants overexpressing *AtHSFA1b::mRFP* under no stress**.** Upregulated TFs targeted by AtHSFA1b under non-stress conditions showed high enrichment for stress response and signal transduction. The statistical cutoff used was Benjamini-Hochberg adjusted *p-value* ≤ 0.01.

The number of target genes of AtHSFA1b that were upregulated under heat stress was significantly less than those targeted and upregulated under no stress condition (figure 5.12). Gene ontology enrichment of the biological processes of AtHSFA1b direct targets that were upregulated under heat stress showed only enrichment for groups of genes involved in 'response to heat' (GO:0009408), 'chemical stress response' (GO:0042221) and 'alcohol metabolic processes' (GO:006066) (figure 5.14).

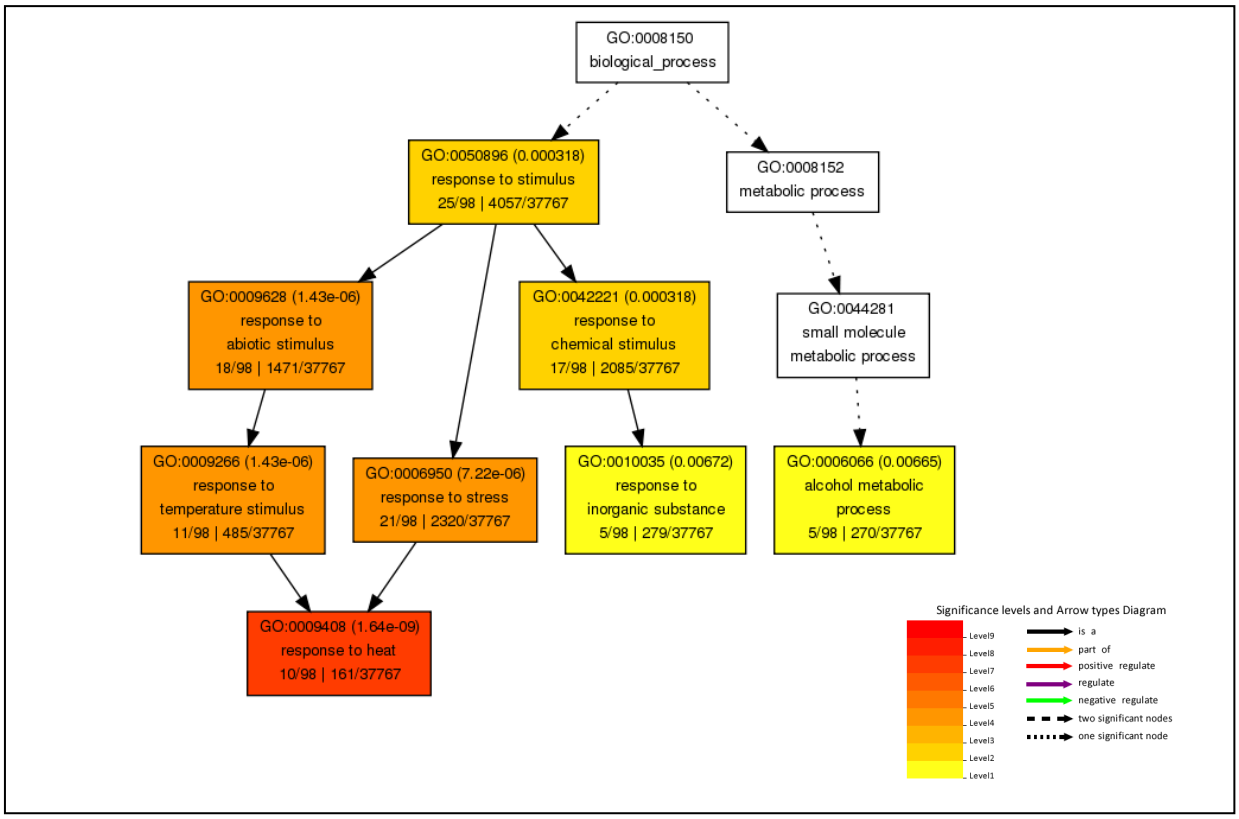


Fig.5.14. Genes targeted by AtHSFA1b and upregulated under heat stress are enriched for stress response. Gene ontology enrichment map of the biological processes of genes directly controlled by AtHSFA1b under heat stress. Only genes that were bound by AtHSFA1b and upregulated under heat stress (37°C for 30 minutes) were analysed for functional enrichment. The map shows that AtHSFA1b targets and induces genes involved in heat stress response.

Despite the loss of AtHSFA1b binding to the promoters of target genes under heat stress, there were groups of genes that maintained strong upregulation in 35S-AtHSFA1b::mRFP plants under heat stress. The groups of genes that were released by AtHSFA1b under heat stress and maintained or showed increased expression under heat stress were mainly genes that code for proteins involved in stress response including members of *AtHSPs* (figure 5.15).

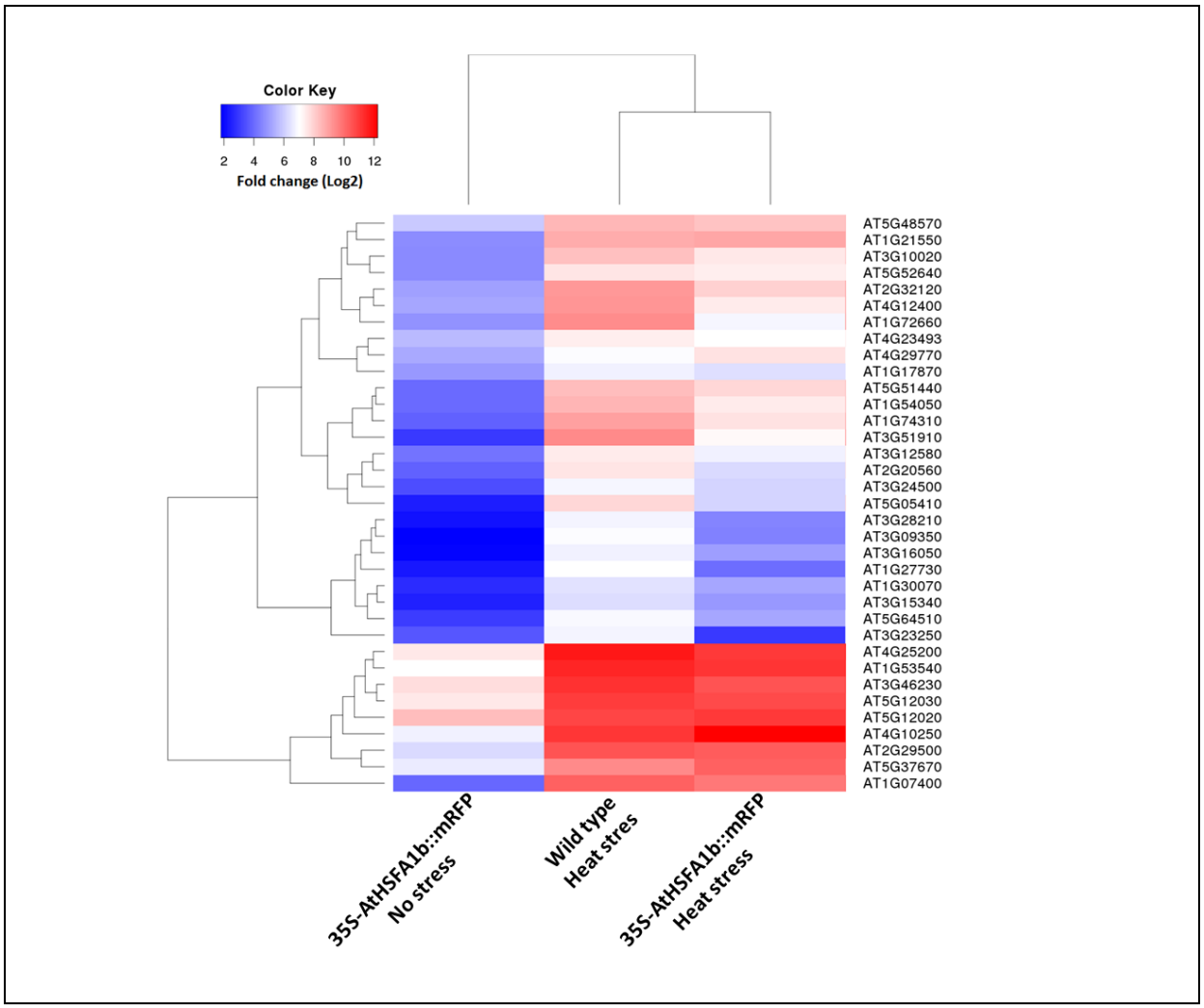


Fig.5.15. Stress response genes released by AtHSFA1b maintain high expression levels under heat stress. Expression pattern of target genes that lost binding of AtHSFA1b under heat stress and maintained high expression levels. Heat map showing the change in expression of selected genes that were targeted by AtHSFA1b under no stress then lost binding under heat stress. The selected genes shown in the heat map are common genes between 35S-AtHSFA1b::mRPF and wild type under heat stress (complete list of the genes and their TAIR annotations are shown in appendix C).

Analysis of the integrated datasets showed that 45.7% and 38% of the downregulated genes under heat stress were targeted by AtHSFA1b under no stress subsequently lost the binding of AtHSFA1b under heat stress in 35S-AtHSFA1b::mRFP and wild type respectively. Gene ontology analysis of the downregulated genes under heat stress condition in both wild type and 35S-AtHSFA1b::mRFP revealed that the vast majority of those genes are involved in crucial plant developmental processes (figure 5.16). Furthermore, analysis of downregulated genes in both wild type and 35S-AtHSFA1b::mRFP plants under heat stress showed that the largest downregulated group of genes ,in terms of molecular functions, was TFs which also showed high enrichment for biological processes such as developmental processes (GO:0032502), signal transduction (GO:0007165) and stress stimulus (GO:0050896) (figure 5.16)

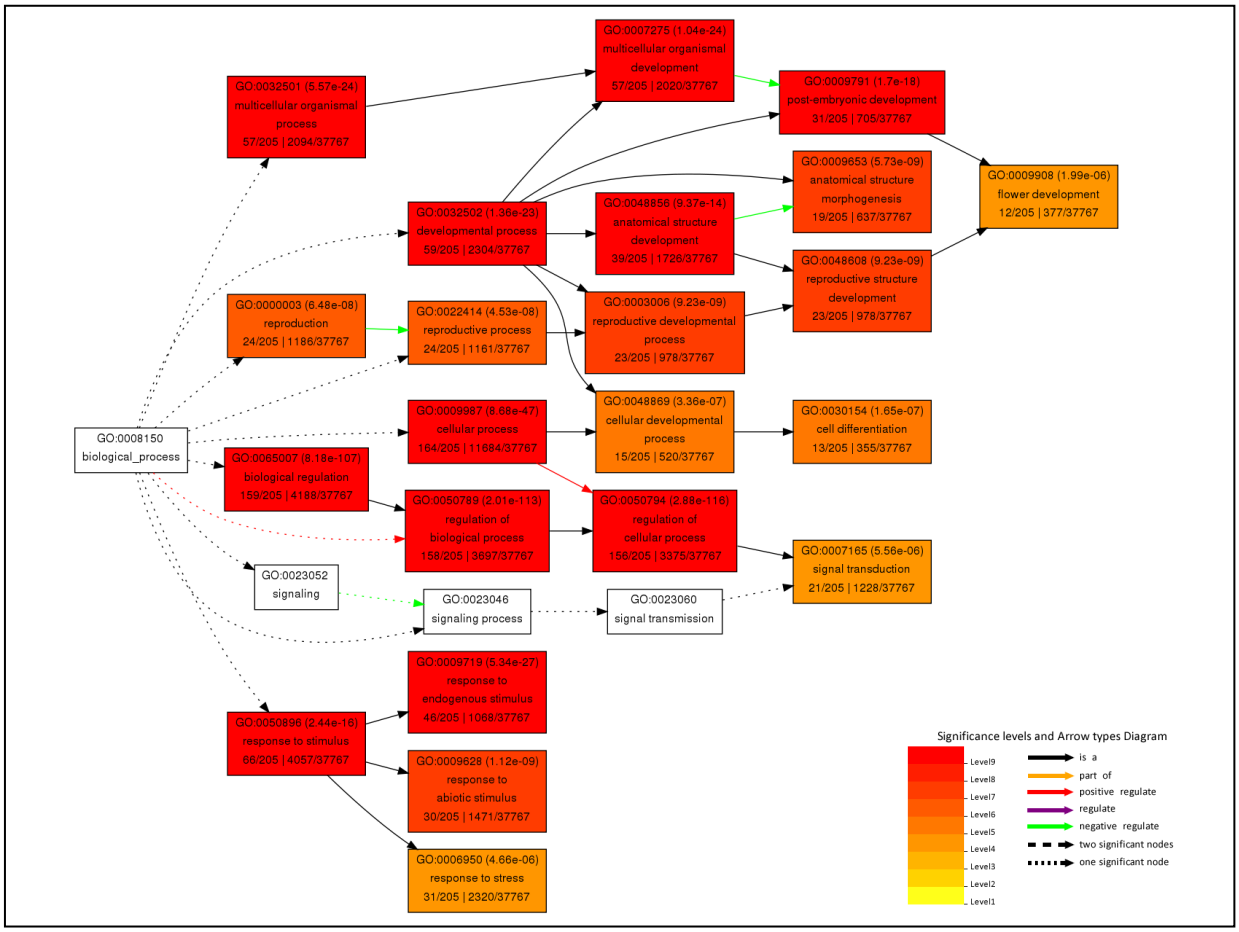


Fig.5.16. The majority of genes released by AtHSFA1b and downregulated under heat stress are TFs involved in plant development. Gene ontology enrichment map of the biological processes of downregulated TFs in both wild type and 35S-AtHSFA1b::mRFP plants that were released from AtHSFA1b under heat stress.

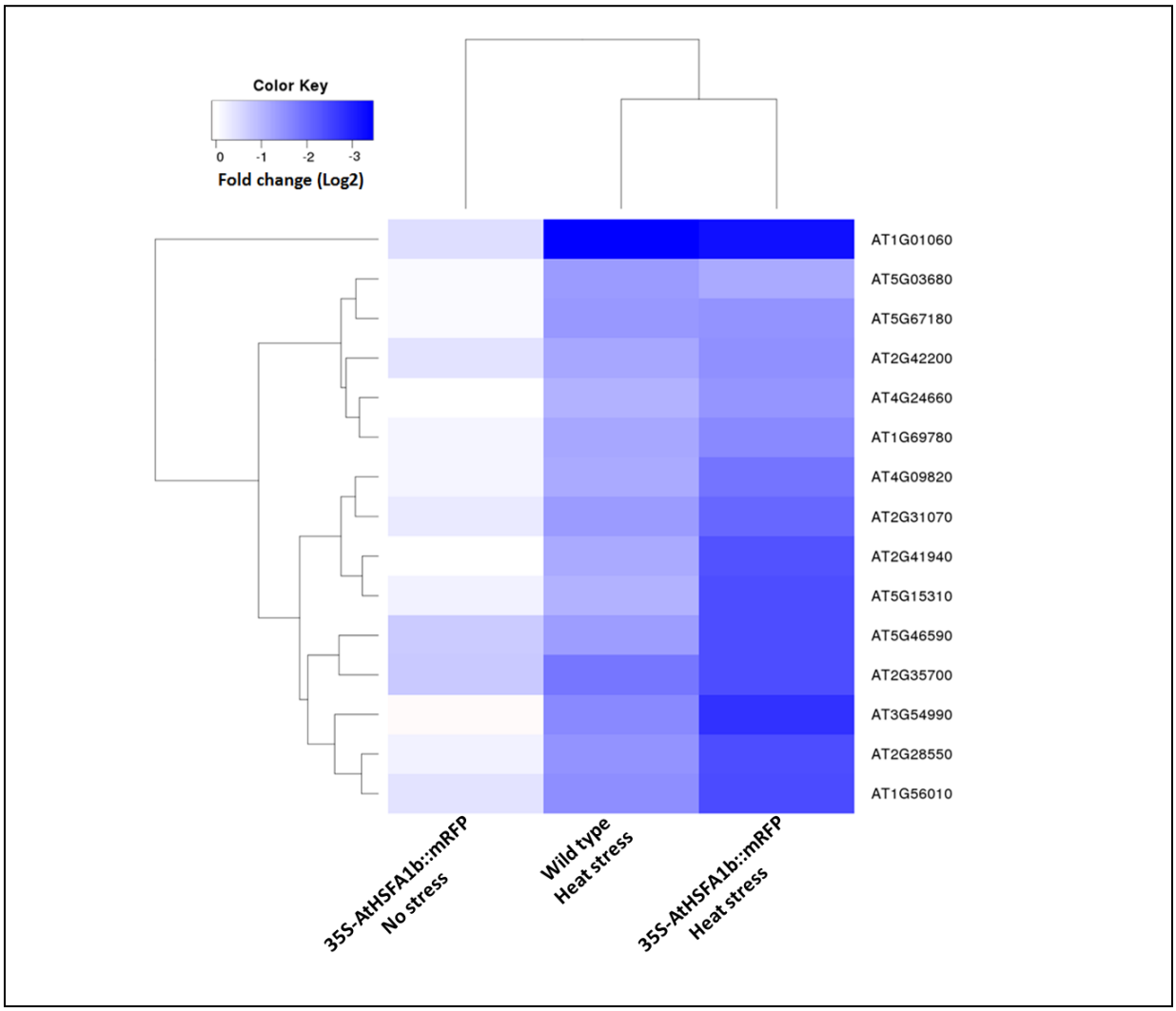


Fig.5.17. Developmental TFs that lost binding of AtHSFA1b were more downregulated in 35S-AtHSFA1b::mRFP plants. Expression pattern of selected developmental TFs targeted by AtHSFA1b under no stress then lost binding of AtHSFA1b under heat stress. Fold change heat map showing the change in expression of 16 selected TFs annotated as involved in developmental processes in 35S-AtHSFA1b::mRFP plants under no stress, wild type and 35S-AtHSFA1b::mRFP plants under heat stress compared to wild type under no stress. The selected TFs are commonly downregulated in both 35S-AtHSFA1b::mRFP and wild type under heat stress.

5.2.5.1 AtHSFA1b binding motif:

The integrated ChIP-SEQ and RNA-SEQ dataset was used in an attempt to generate a specific binding element for AtHSFA1b as described in Section (2.4.2.2.1). The aim of this method was to resolve the specific binding *cis*-element of AtHSFA1b and avoid the overlap between different forms of HSE that result from conventional analysis on the entire promoter sequences of DEGs and could potentially be HSEs occupied by other AtHSFs.

Briefly, the method is based on isolating the genes that are bound by AtHSFA1b and differentially expressed by intersecting the ChIP-SEQ and RNA-SEQ datasets using common AGI codes between genes bound by AtHSFA1b and DEGs. From that intersected list, only the sequences that represent AtHSFA1b binding regions from ChIP-SEQ were used in *de novo* motif analysis (figure 5.18)

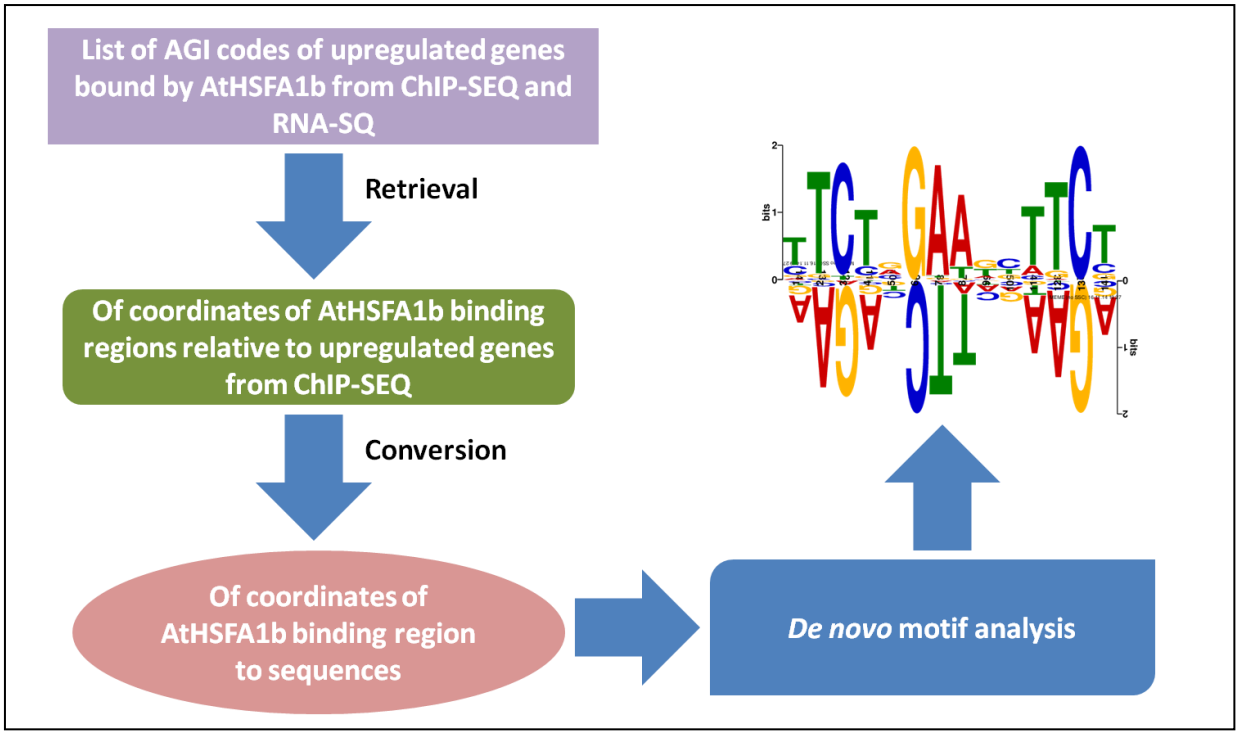


Fig.5.18. Summary of the method used to discover AtHSFA1b binding element. Flow chart showing the general procedure used to discover AtHSFA1b preferred HSE. Both ChIP-SEQ and RNA-SEQ data sets were merged into one dataset which resulted in assigning each gene bound by AtHSFA1b an expression value. Genes that were bound and upregulated compared wild type control were isolated and only the sequences of the regions of those genes that were occupied by AtHSFA1b were used in *de novo* motif analysis using three different programs (rGADEM, MEME and STEME).

This improved motif discovery method solved the problem of multiple overlapping HSEs on the promoters of DEGs and resulted, with total agreement between the programs used, in one unique HSE in high resolution which might be the preferred HSE for AtHSFA1b (figure 5.19).

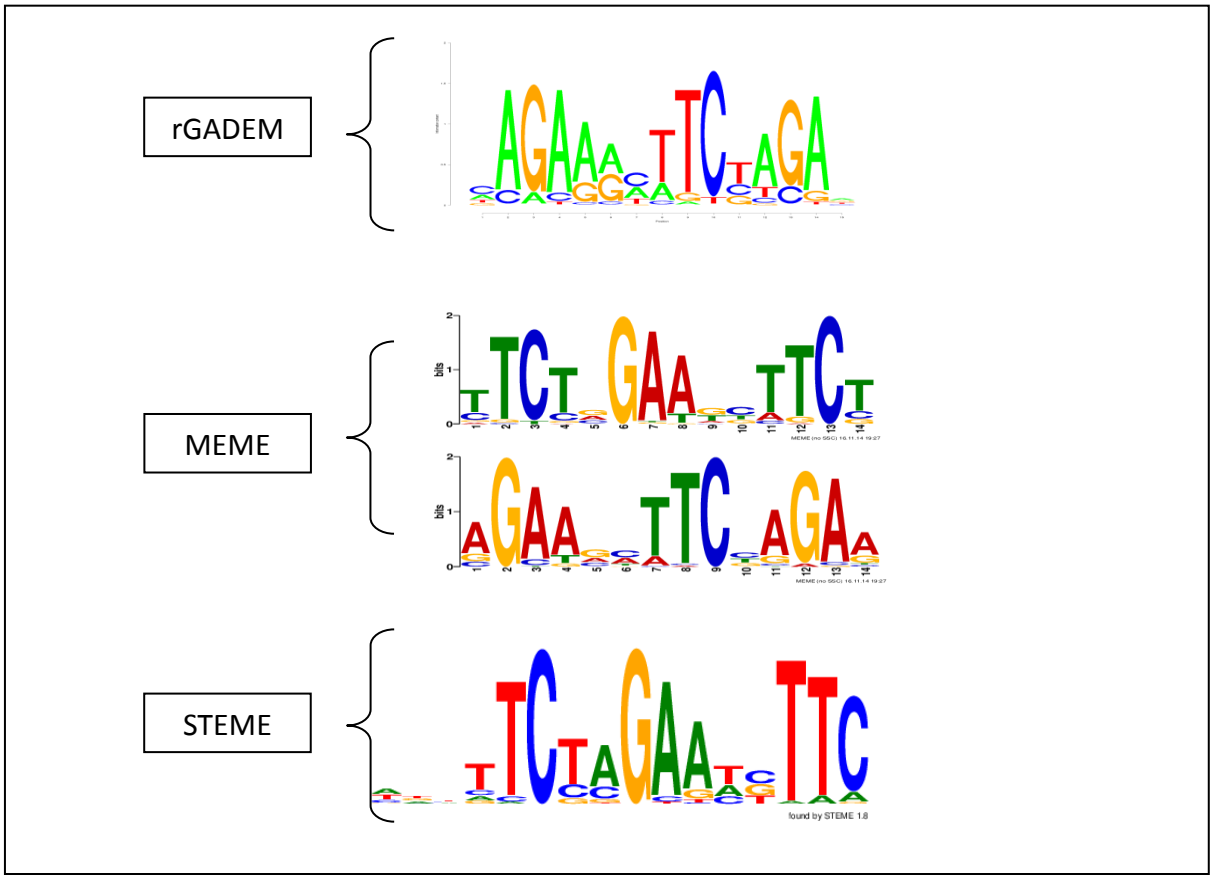


Fig.5.19. Structure of the functional AtHSFA1b binding element (HSE). The AtHSFA1b binding motif was obtained from integration of ChIP-SEQ and RNA-SEQ datasets. The motif obtained from the analysis of the integrated datasets represents one form from the multiple overlapping motifs recovered from analysing the entire promoter sequences of upregulated genes shown in figure 5.11.

5.3 Discussion:

5.3.1 A note about RNA-SEQ expression analysis:

Integrating the ChIP-SEQ data with RNA-SEQ showed that only a small fraction of the genes bound by AtHSFA1b resulted in change in expression in 35S-AtHSA1b::mRFP plants compared to wild type under no stress condition (figure 5.12). However, potential targets of AtHSFA1b including developmental genes, did not show any significant differential expression when the transcript levels of those genes in 35S-AtHSFA1b::mRFP plants were compared to those in wild type under no stress. However, as described in the method Sections (2.1.1 and 2.3.1) RNA was extracted from whole *A. thaliana* rosettes. One drawback of this method is that it is biased towards globally expressed genes and highly abundant transcripts and may not give an accurate measure of expression levels of tissue-specific and cell type-specific genes unless there is an extreme change in expression (Malone and Oliver, 2011). Another drawback is that all plants were tested at a specific fixed developmental stage which was fifth week from germination (Section 2.1.1). This means that the method might not provide an accurate measure of the changes in expression of genes that are expressed at specific developmental stages (Zenoni et al., 2010; Wang et al., 2009). Thus, a further improvement to the technique would be testing the expression of AtHSFA1b target genes at different developmental stages as well as taking into account tissue specificity of target genes of AtHSFA1b.

It is also worth noting that the conventional ratio between expression levels can, on many occasions, be misleading and may not reflect the real levels of upregulation and downregulation of genes. There are examples in this study where the ratio between transcript levels of target genes of AtHSFA1b in 35S-AtHSFA1b::mRFP and wild type plants under no stress showed that the change in expression of some genes is below the 2-fold change cutoff. However, comparing

the absolute expression values (FPKM) of those target genes showed, in many instances, a massive increase/decrease in the transcript levels of those genes in 35S-AtHSFA1b::mRFP plants. For example, the ratio between expression levels showed that the change in expression of *AtAPX1*, which is targeted by AtHSFA1b under no stress, is below the 2-fold change cutoff. However, by looking at absolute FPKM values, the transcript level of *AtAPX1* was significantly increased by overexpression of *AtHSFA1b*. The average absolute expression FPKM value of *AtAPX1* in wild type under no stress was 731.5, whereas in 35S-AtHSFA1b::mRFP plants under the same condition was 1003.7. This shows that the average increase in the transcript level of *AtAPX1* in 35S-AtHSFA1b::mRFP plants under no stress was 272.2 FPKM compared to wild type under the same condition. However, when looking at the ratio between the two FPKM values of *AtAPX1* in wild type and 35S-AtHSFA1b::mRFP under no stress the fold increase is 1.37 which is below the conventional 2-fold change cutoff used in most expression analyses which resulted in marking this change in the expression of *AtAPX1* as 'not significant' (Dalman et al., 2012; Tan et al., 2003; Mutch et al., 2002).

In an opposite scenario, fold change can also result in an overestimation of change in expression levels. For instance, if the starting FPKM value of a target gene in a control sample is low then any small increase or decrease in FPKM value in the test sample will result in an overestimated (or exaggerated) fold change value. For example, one of the genes that was marked as significantly upregulated in 35S-AtHSFA1b::mRFP plants under no stress was germin-like protein 9 (*AtGLP9*). The average FPKM value of *AtGLP9* was 0.16 in wild type under no stress, whereas in 35S-HSFA1b::mRFP plant under the same condition the average FPKM value of *AtGLP9* was 1.65. When the ratio of the transcript level of *AtGLP9* between 35S-AtHSFA1b::mRFP and wild type was calculated the result was 10.3 fold. As a result, the gene was marked as significantly

upregulated. Therefore, it can be concluded that there are instances where it is appropriate to look at and directly compare changes in absolute expression FPKM values beside the conventional ratio of expression. However, care should be taking when using pure FPKM values to report changes in expression as FPKM values can be variable among replicates of the same sample. Therefore, the quality of RNA-SEQ experiment is crucial for the success and reproducibility of this method.

One drawback in this method of analysing RNA-SEQ data, is that it is difficult to decide an appropriate cutoff as the method is affected by the abundance of the genes expressed, similar to the conventional 2-fold change cutoff but in an opposite way. For example, if a highly abundant gene is expressed at an average FPKM value of 1000 in a control sample then that value changes in the test sample to 1010 then it will be automatically marked as 'significantly upregulated' if the cutoff is set at 10 FPKM difference. However, it is not possible to predict whether an increase of 10 FPKM for a gene that is expressed at such high level would have an influence on the overall abundance of that gene and subsequently an impact on the overall biology of the system under study. Therefore, both methods should be considered when analysing RNA-SEQ data in order to get a more in depth view of the real changes in transcript levels.

5.3.2 The stress component of AtHSFA1b:

Analysis of the RNA-SEQ data showed that overexpression of *AtHSFA1b* results in plants having a partial heat stress expression profile under no stress conditions. More detailed statistical analysis of the total transcriptome of plants overexpressing *AtHSFA1b* under no stress revealed that the expression profile of those plants fall in an area in between no stress and heat stress.

Further analysis on the biological processes of upregulated genes in 35S-AtHSFA1b::mRFP plants under no stress showed indeed that the vast majority of genes that were upregulated are involved in abiotic and biotic stress responses (figures 5.3, 5.4, 5.5). Similar result was observed from heat stressed wild type and 35S-AtHSFA1b::mRFP expression data. This result gives strong indication that overexpression of *AtHSFA1b* results in plants in a 'ready-state' to cope with stress by constitutively expressing stress responsive genes under no stress conditions. This result is consistent with previous studies that showed that overexpression of *AtHSFA1b* results in plants with enhanced heat stress response (Bechtold et al., 2013; Yoshida et al., 2011; Lohmann et al., 2002; Prändl et al., 1998; Lee et al., 1995). Furthermore, the mentioned studies showed that overexpressing *AtHSFA1b* does not only lead to changes in the transcript levels of AtHSP genes but it leads to an increase the protein levels of AtHSPs in *A. thaliana* plants under no stress in a similar manner to that which occurs in heat stressed wild type plants (Prändl et al., 1998; Lee et al., 1995)

Similar to HSFs, the HSP family is one of the most structurally and functionally conserved protein families among all eukaryotes. They mainly act as intracellular chaperones by which they aid refolding of misfolded proteins under stress conditions (Section 1.2; Åkerfelt et al., 2010; Kregel 2001; Feder 1999). The accumulation of AtHSPs and other stress-associated proteins in *A. thaliana* plants overexpressing *AtHSFA1b* might be the explanation of the enhanced stress response in those plants. The accumulation of AtHSP chaperones and other stress-related proteins under no stress conditions might allow the plant to cope with stress conditions faster than wild type plants which might need to go through the process of expressing and synthesising those proteins under stress which can slow down the response to stress. Overexpression of *AtHSFA1b* resulted in upregulation of genes that are involved in

various forms of stress not only in HSR including genes that code for proteins involved in anti-oxidant activities such as *AtAPX1* and *AtAPX2*.

There seem to be a large overlap in plants' responses to abiotic stress in particular. One of the main characteristics of plant response to abiotic stress is overproduction of ROS. Nearly all forms of plant abiotic stress result in bursts of ROS in plant cells (Baxter et al., 2013; de Carvalho, 2008). ROS are highly reactive and toxic compounds and overproduction of ROS can damage essential proteins, lipids and carbohydrates in plant cells (Gill and Tutega, 2010; Miller et al., 2008). Plant response to abiotic stress involves induction of chaperones that mainly act in repairing damaged proteins and anti-oxidants that aid detoxifying plant cells from high levels of ROS (Gill and Tuteja, 2010; Vinocur and Altman, 2005; Apel and Hirt, 2004; Sabehat et al., 2002; Parsell and Lindquist, 1993). This seems to be a common mechanism that occurs in plants in response to abiotic stress. Therefore, the enhanced abiotic stress tolerance in *AtHSFA1b* overexpressing plants might be due to not only the accumulation AtHSP chaperones but also other anti-oxidant proteins under no stress condition which might make the process of repairing damaged protein and detoxifying the plant cells from ROS occur in a shorter time avoiding the delay due to the time required for expression and synthesis cycles of those proteins in wild type plants.

The functions of HSPs, in particular, can go beyond just chaperones repairing misfolded proteins and degrading accumulated damaged proteins. It has been shown in mammalian studies that HSPs can play a key role in the regulation of the function of HSFs under different conditions. Studies have shown that hHSP90 can have an inhibitory effect on hHSF1 under no stress condition by binding to their oligomerisation domain preventing it from forming an active

homotrimer and maintaining it in an inactive monomeric state (Wang, et al., 2005; Zou, et al., 1998). Other studies showed that hHSP70 have a similar effect on hHSF1 to that of hHSP90 where it blocks hHSF1 from binding to the DNA (Abravaya, et al., 1992). The inhibitory role of HSP90/70 on HSFs is well studied in biological systems other than plants and models of the regulatory mechanisms of HSFs by HSPs were proposed and in many instances were proven by appropriate experimental procedures (Wang, et al., 2005; Zou, et al., 1998; Abravaya, et al., 1992). However, it is not clear if that is the case in *A. thaliana* or not especially with AtHSFA1b.

Both ChIP-SEQ and RNA-SEQ experiments showed clues that a similar mechanism might be occurring in *A. thaliana*. The loss of binding of AtHSFA1b from its target regions in the *A. thaliana* genome under heat stress shown by ChIP-SEQ and the extreme increase of transcript levels of many isoforms of *AtHSP90* and *AtHSP70* indicate that AtHSFA1b might be subject to a similar regulatory mechanism to those suggested in mammalian systems (table 5.4). The high conservation of the HSR process among eukaryotes also strengthens this argument (Section 1.1). However, without appropriate protein-protein interaction experiments to validate those possible regulatory mechanism, this theory remains a speculation.

AGI code	Gene name	FPKM in wild type no stress	FPKM in wild type heat stress	FPKM in 35S-AtHSFA1b::mRFP heat stress	Fold Change in wild type heat stree (Log2)	Fold Change in 35S-AtHSFA1b::mRFP heat stress (Log2)
AT1G16030	*AtHSP70b*	2.83	2353.60	4288.43	9.69	10.56
AT2G32120	*AtHSP70T-2*	1.71	965.89	437.15	9.13	7.99
AT5G52640	*AtHSP90.1*	25.21	4859.59	4316.89	7.59	7.42
AT3G12580	*AtHSP70*	43.0751	7721.60	4525.47	7.49	6.72
AT5G09590	*AtHSC70-5*	44.55	1353.08	337.49	4.93	2.92
AT5G02490	*AtHSP70*	17.99	402.58	23.80	4.48	0.40
AT5G56030	*AtHSP90.2*	269.66	3667.45	1478.46	3.77	2.46
AT1G79920	*AtHSP70*	99.26	992.44	544.56	3.32	2.46
AT3G07770	*AtHSP90.6*	29.73	289.05	82.45	3.28	1.47
AT2G04030	*AtHSP90.5*	111.80	808.1	461.30	2.85	2.10
AT3G09440	*AtHSP70*	517.13	3122.57	1462.62	2.59	1.50
AT4G24190	*AtHSP90.7*	136.69	716.83	366.20	2.39	1.42
AT5G56010	*AtHSP90.3*	316.00	1417.19	320	2.17	0.018
AT4G16660	*AtHSP70*	37.53	161.03	167.691	2.10	2.16
AT1G11660	*AtHSP70*	11.80	41.63	146.163	1.82	3.63
AT5G02500	*AtHSP70-1*	1504.93	4315.89	1636.01	1.52	0.12

Table 5.4. Upregulated Isoforms of *AtHSP70* and *AtHSP90* genes in wild type and 35S-AtHSFA1b::mRFP plants under heat stress. The table shows TAIR codes, names of the genes, the FPKM values in wild type under no stress, wild type and 35S-AtHSFA1b::mRFP plants under heat stress and the fold change.

The heat stress experiment showed that the expression of stress responsive genes that were targeted by AtHSFA1b under the no stress condition and lost binding of AtHSFA1b under heat stress was maintained or further increased even when AtHSFA1b was no longer binding to those genes. This result suggests that the role of AtHSFA1b might be taken by another AtHSF under prolonged heat stress. Heat stress treatment led to increased transcript levels of a number of AtHSFs (table 5.5). Microarray and RNA-SEQ analyses on *A. thaliana* plants overexpressing heat inducible *AtHSFs* such as, *AtHSFA2* and *AtHSFA3*, showed a degree of overlap in the genes they induce with the genes induced by overexpression of *AtHSFA1b* (figure 4.17). Stress inducible AtHSFs seem to be controlling much smaller groups of genes than those controlled by AtHSFA1b. It has been shown, for instance, that overexpression of *AtHSFA2* leads to the upregulation of only 57 genes. All the genes controlled by AtHSFA2 code only for AtHSPs and other stress-associated proteins (Bechtold et al., 2013; Ogawa, et al., 2007).

AGI code	Gene name	FPKM in wild type no stress	FPKM in wild type heat stress	FPKM in 35S-AtHSFA1b::mRFP heat stress	Fold change in wild type heat stress (Log2)	Fold change in 35S-AtHSFA1b::mRFP heat stress (Log2)
AT3G51910	*AtHSFA7a*	1.11	768.17	159.01	9.44	7.17
AT3G63350	*AtHSFA7b*	0.12	105.00	31.22	10.03	8.28
AT2G26150	*AtHSFA2*	6.48	2175.9	577.605	8.40	6.48
AT1G32330	*AtHSFA1d*	3.95	9.07	24.80	1.20	2.65
AT3G02990	*AtHSFA1e*	1.46	3.73	14.94	1.35	3.35
AT5G03720	*AtHSFA3*	0.58	10.50	1.70	4.17	1.54
AT4G18880	*AtHSFA4a*	7.13	24.40	31.17	1.77	2.13
AT1G67970	*AtHSFA8*	3.17624	13.0414	4.98	2.04	0.65
AT4G36990	*AtHSFB1*	10.60	518.70	546.59	5.61	5.68
AT5G62020	*AtHSFB2a*	5.62	73.67	139.30	3.71	4.63
AT4G11660	*AtHSFB2b*	4.78	91.48	65.62	4.26	3.78
AT1G46264	*AtHSFB4*	5.20	10.10	13.63	1.0	1.39
AT1G58130	*AtHSF30*	0.18	0.92	1.12952	2.33	2.63

Table 5.5. AtHSFs that showed increase in expression in response to heat stress in both wild type and 35S-AtHSFAb::mRFP plants. The table shows TAIR codes, names of the genes, the FPKM values in wild type under no stress, wild type and 35S-AtHSFA1b::mRFP plants under heat stress and the fold change.

When the group of genes controlled by *AtHSFA2* was intersected with the group of genes controlled by *AtHSFA1b*, it showed that *AtHSFA2* possibly controls a subgroup of the genes controlled by *AtHSFA1b* also similar output was observed when the genes upregulated by overexpression of *AtHSFA3* was overlapped with the genes upregulated as a result of overexpression of *AtHSFA1b* (figure 4.17). This supports the evidence that AtHSFA1b involvement in heat stress response is somehow limited compared to those stress inducible AtHSFs. Moreover, these results further support the study that showed that AtHSFA1b is only involved in the very early phases of HSR (Lohmann et al., 2004). Therefore, it can be speculated that during stress AtHSFA1b induces the expression of *AtHSP* genes and other stress responsive genes for a very short period of time then its role is taken by other stress inducible and highly specialised AtHSFs in prolonged heat stress conditions. Therefore, the output of this study shows glimpses on the possible hierarchical organisation in the AtHSF regulatory network where involvement of AtHSFA1b in the regulation of at least HSR could be limited to rapid initiation of

expression of *AtHSP* genes and a group of *AtHSFs* for a short duration. The regulation of HSR could then be taken by other stress inducible AtHSFs that might be responsible for maintaining the expression of *AtHSPs* under prolonged heat stress conditions.

So far, this part of the research has shown that overexpression of *AtHSFA1b* leads to accumulation of AtHSPs and other stress responsive protein in *A. thaliana* plants under no stress conditions which could potentially be the explanation of the improved stress tolerance in those plants. This leads to faster and improved stress response and high survival rate under stressful conditions in those plants compared to wild type. However, such response may have an impact on other plant processes. The accumulation of AtHSPs and other stress responsive genes under no stress conditions can also be looked at as a biological system being 'stressed' under no stress conditions. Generally, plants' response to stress comes at the expense of other non-stress associated processes (Pajerowska-Mukhtar et al., 2012; Tian et al., 2003). However, one of the interesting observations is that overexpressing of *AtHSFA1b* in *A. thaliana* plants leads to subtle developmental effect in the form of increased seed yield and harvest index under normal and stressful growth conditions (Bechtold et al., 2013). Therefore, looking at the possible developmental component of AtHSFA1b is crucial to further understand the impact of AtHSFA1b on other biological processes out of the stress context.

5.3.3 The developmental component of AtHSFA1b:

It was suggested that overexpression of *AtHSFA1b* results in a developmental effect in *A. thaliana* and *B. napus* plants (Bechtold, et al., 2013). Due to the domination of *AtHSPs* and other stress-associated genes in 35S-AtHSFA1b::mRFP plants among those altered in expression under no stress conditions, gene ontology analysis did not show any other groups of genes associated with other processes (figure 5.5). However, examining genes individually showed that there were upregulated genes in 35S-AtHSFA1b::mRFP plants under no stress that might be involved in developmental processes. Further analysis on individual genes showed only a small number of upregulated genes that could potentially be developmental genes in 35S-AtHSFA1b::mRFP plants under no stress. Those genes include, *AtCDKD1*, *AtAGL19*, *AtPSK2* and *AtPSK4* (Table 5.6). The fold changes in the expression of those candidate genes were low compared to that of stress-related genes. However, looking at absolute expression FPKM value showed increases in the transcript levels of those genes in 35S-AtHSFA1b::mRFP plants under no stress compared to their FPKM values in wild type plants under the same condition (figure 5.22 and table 5.6).

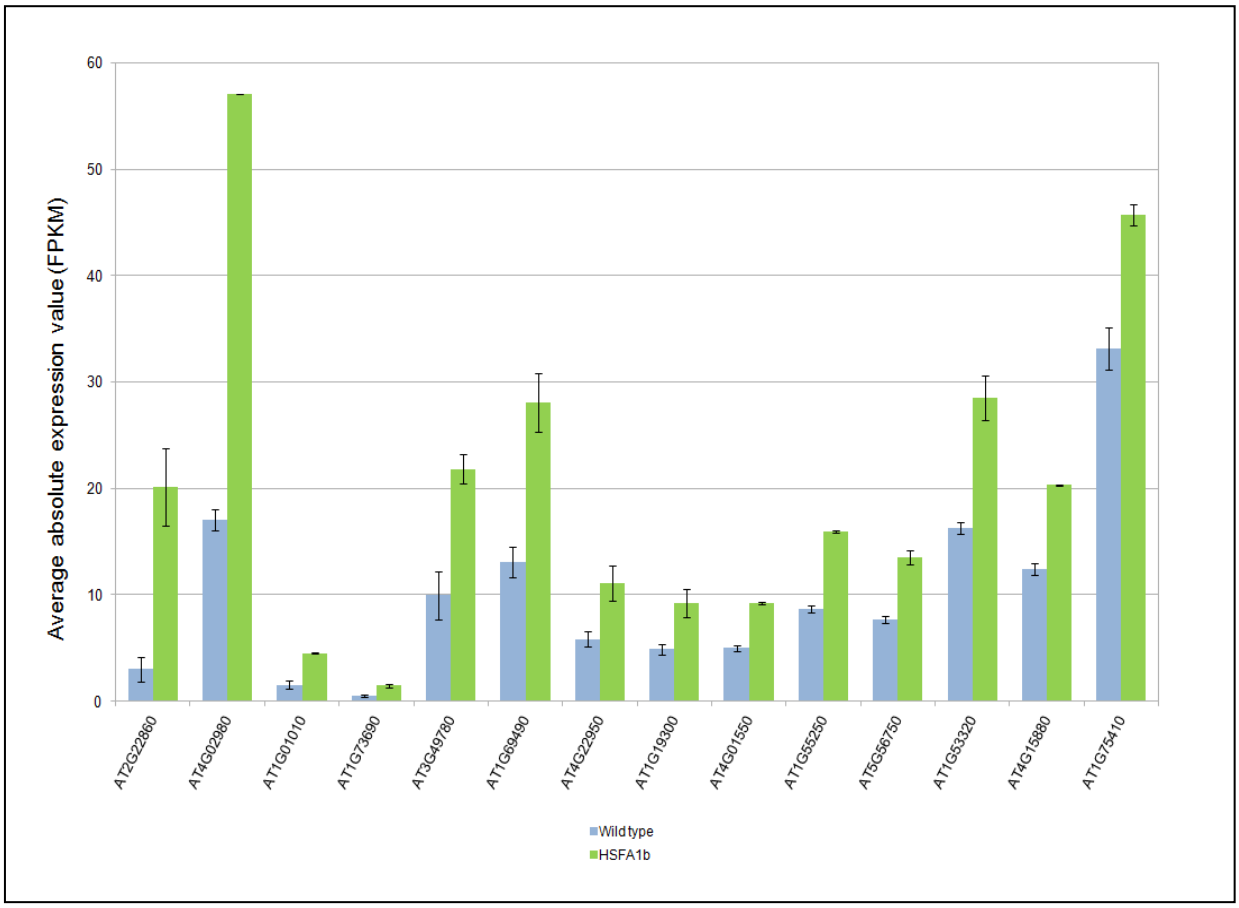


Fig.5.22. Overexpression of AtHSFA1b induces the expression of genes annotated as developmental genes. Changes in expression levels of possible 14 developmental genes in 35S-AtHSFA1b::mRFP plants under no stress. Bar plot showing the increase in expression values of 14 developmental genes in 35S-AtHSFA1b::mRFP plants under no stress condition compared to wild type under the same condition.

AGI code	Gene name	Description
AT2G22860	*AtPSK2*	Phytosulfokine 2 precursor, coding for a unique plant peptide growth factor.
AT4G02980	*AtABP1*	Auxin binding protein involved in cell elongation and cell division. ABP1 is ubiquitinated in vitro and in planta by AtRma2.
AT1G01010	*AtNAC001*	NAC domain containing protein 1 (NAC001); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription
AT1G73690	*AtCDKD1*	cyclin dependent kinase activator CDKD;1. Nuclear localization. Involved in cell cycle regulation and cell differentiation.
AT3G49780	*AtPSK4*	Phytosulfokine 3 precursor, coding for a unique plant peptide growth factor. Plants overexpressing this gene (under a 35S promoter), develop normal cotyledons and hypocotyls but their growth, in particular that of their roots, was faster than that of wild type.
AT1G69490	*AtNAP*	Encodes a member of the NAC transcription factor gene family. It is expressed in floral primordia and upregulated by AP3 and PI. Its expression is associated with leaf senescence.
AT4G22950	*AtAGL19*	MADS-box protein AGL19
AT1G19300	*AtPARVUS*	The PARVUS/GLZ1 gene encodes a putative family 8 glycosyl transferase that contributes to xylan biosynthesis. Its gene expression shows good co-variance with the IRX3 gene involved in secondary cell wall synthesis. PARVUS/GLZ1 is predicted to have galacturonosyltransferase activity and may be involved in the formation of the complex oligosaccharide sequence present at the reducing end of xylan. PARVUS is expressed in cells undergoing secondary wall thickening, and parvus mutants have thinner cell walls.
AT4G01550	*AtNAC069*	Encodes a plasma-membrane bound NAC transcription factor, whose controlled proteolytic activation allows it to enter the nucleus.
AT1G55250	*AtHUB2*	Encodes one of two orthologous E3 ubiquitin ligases in Arabidopsis that are involved in monoubiquitination of histone H2B.
AT5G56750	*AtNDL1*	N-MYC downregulated-like 1 (NDL1)
AT1G53320	*AtTLP7*	Member of TLP family
AT4G15880	*AtESD4*	EARLY IN SHORT DAYS 4 Arabidopsis mutant shows extreme early flowering and alterations in shoot development. It encodes a SUMO protease, located predominantly at the periphery of the nucleus. Accelerates the transition from vegetative growth to flowering. Probably acts in the same pathway as NUA in affecting flowering time, vegetative and inflorescence development.
AT1G75410	*AtBLH3*	BEL1-like homeodomain 3 (BLH3)

Table 5.6. Upregualted developmental genes in 35S-AtHSFA1b::mRFP plants under no stress. The table shows the TAIR accession codes, names and descriptions of upregulated genes in non-stressed 35S-AtHSFA1b::mRFP plants that are annotated as developmental. Annotations were obtained from the *A. thaliana* database (TAIR; http://www.arabidopsis.org).

Further analysis on the expression levels of the 14 developmental genes under both conditions showed that their expression is generally affected by heat stress. The expression level of the 14 developmental genes is generally increased by heat stress as well as overexpression of *AtHSFA1b* (figure 5.23).

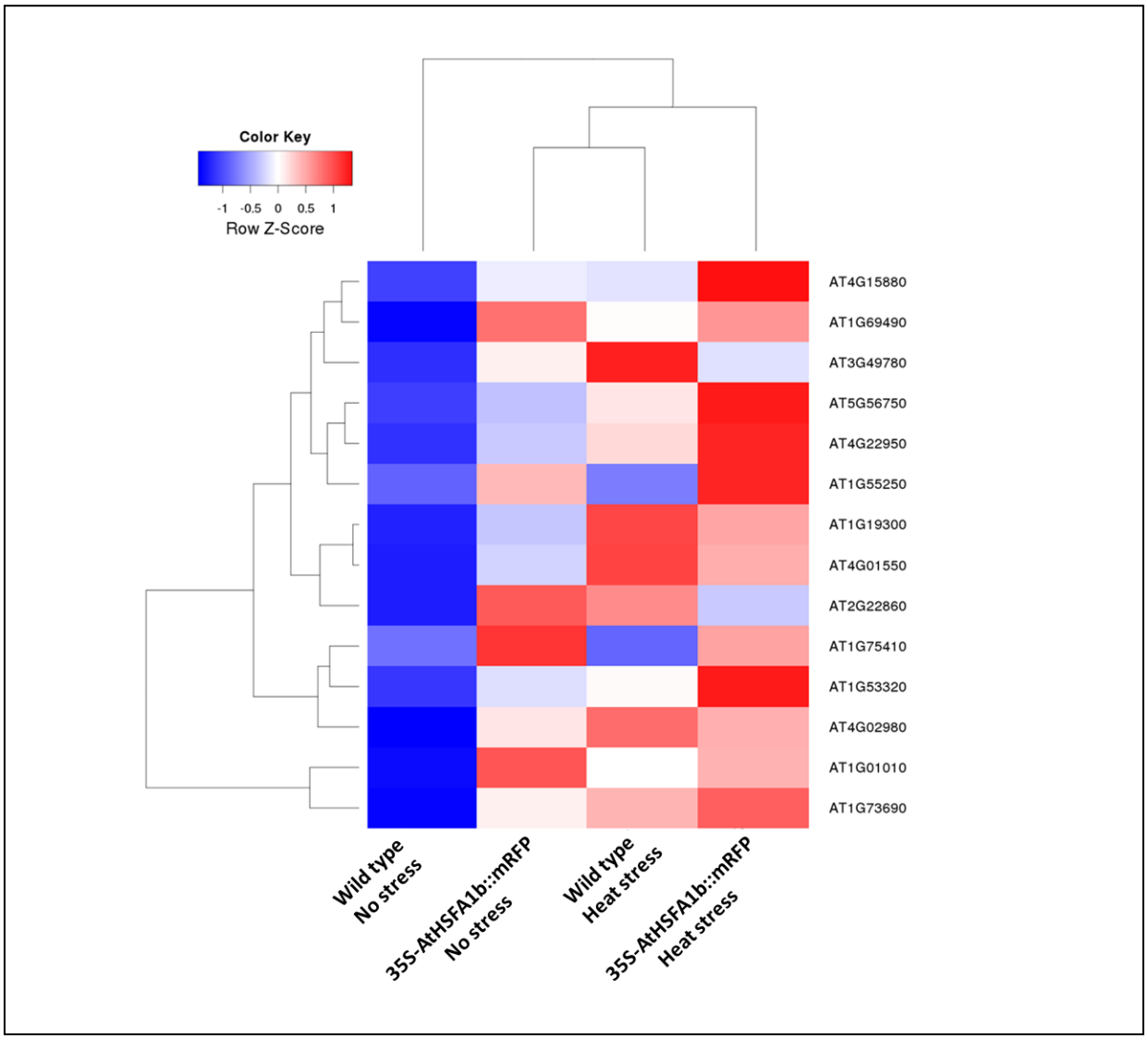


Fig.5.23. The developmental genes induced by overexpression of AtHSFA1b are also induced by heat stress. The expression pattern of the 14 possible developmental genes is responsive to heat. Heat map showing the change in expression of the 14 developmental genes in all treatments from the RNA-SEQ experiment. The heat map shows that the expression of the 14 developmental genes in increased in response to heat stress in both wild type and 35S-AtHSFA1b::mRFP plants.

It is still not clear whether the increase in the transcript levels of these genes would have an influence on plant developmental process or not. However, further analysis of these genes should be taken into consideration as they might be expressed in tissue-specific manners and/or at specific stages in the plant life cycle. For example, the gene *AtCDKD1* which, according to TAIR annotation is involved in the regulation of plant cell cycle and differentiation has a specific expression pattern. According to TAIR annotation of *AtCDKD1*, the expression of this gene is

mainly occurring during the bilateral stage, expanded cotyledon stage, mature embryo stage, petal differentiation and expansion stage. In terms of tissue specificity, *AtCDKD1* is expressed in flower, guard cell, plant embryo, root, seed, and shoot apex (http://www.arabidopsis.org/servlets/TairObject?id=29548&type=locus). Therefore, further detailed expression analysis of this genes and the other 13 should be considered taken into account their tissue specificity and developmental stages where they are expressed.

The phytosulfokine precursor 2 and 4 genes (*AtPSK2* and *AtPSK4*) are also potential developmental candidates in plants overexpressing *AtHSFA1b*. Both genes code for small peptides that functions as growth factors prohormones (Brenner et al., 2005). Both AtPSK genes code for short peptides with identified structures as sulfated peptides. There are two types of PSK peptides in plants, PSK-α and PSK-β. Both peptides were structurally characterised with PSK-α being determined as sulfated pentapeptide and PSK- β as sulfated tetrapeptide (Matsubayashi and Sakagami, 1996). It was later shown that PSK-α is more biologically active than PSK- β and the latter being a product of enzymatic cleavage of the c-terminal of PSK-α (Yang et al., 1999). AtPSKs are involved in various functions including accelerating root development, cell growth and differentiation, cell-to-cell communication and pollen germination (Chen et al., 2000; Matsubayashi et al., 1999b; Kobayashi et al., 1999; Yamakawa et al., 1998). Overexpression of *AtPSK2* was shown to increase the growth of *A. thaliana* twice as much as wild type (Yang et al., 1999). Overexpression of *AtPSK4*, on the other hand, was shown to enhance root growth in *A. thaliana* plants (Kutschmar et al., 2009).

This study and others have shown that expression levels of AtPSK2 and AtPSK4 are affected by heat stress (figure 5.23), wound and microbial interaction in *A. thaliana* (Loivamäki et al., 2010;

Benner et al., 2005). Interestingly, overexpression of *AtPSK2* and *AtPSK4* resulted in increased susceptibility to *Pseudomonas syringae pv. tomato* DC3000 (*Pto DC3000*) and increased resistance to the necrotrophic fungus *A. brassicicola* (Mosher et al., 2012). These results suggest that both AtPSK2 and AtPSK4 might be involved in broad functions in *A. thaliana* beside their main functions as promoters of plant cell growth. However, it is not clear how such peptides that were characterised as growth factors would have an influence on plant resistance to stress.

Both AtPSK2 and AtPSK4 were upregulated in 35S-AtHSFA1b::mRFP plants under no stress (figure 5.22 and table 5.6). However, the ChIP-SEQ experiment under no stress showed that AtHSFA1b directly targets only *AtPSK4* suggesting that *AtPSK2* is indirectly regulated by AtHSFA1b (Appendix B). Interestingly the upregulation of *AtPSK2* was much higher than *AtPSK4* in 35S-AtHSFA1b::mRFP plants under no stress (figure 5.19). This might be because AtPSK4 is expressed predominantly in *A. thaliana* roots (Kutschmar et al., 2008).

Possibly the most striking event in the AtHSFA1b regulatory network is the collapse of the network under heat stress and the association of the collapse with the general downregulation of developmental genes under heat stress. Looking at the ChIP-SEQ and RNA-SEQ data from the heat stress experiments, they clearly demonstrate that the loss of binding of AtHSFA1b to the promoters of its target genes was associated with a general downregulation of genes involved in growth development and other non-stress related processes (figures 5.16 and 5.17). It is not possible to relate that downregulation just to the loss of AtHSFA1b binding to the promoters of those genes. However, it is widely known that plant and eukaryotic responses to stress in general results in downregulation of other processes including growth and development to allocate resources and energy for the protection of the system (Pajerowska-Mukhtar et al.,

2012; Tian et al., 2003). Furthermore, the short and limited involvement of AtHSFA1b in the regulation of HSR for a very short duration, the constitutive expression pattern, and the large number of genes controlled by AtHSFA1b suggest that this particular AtHSF might have evolved to become more than just HSF.

5.3.4 AtHSFA1b and HSEs:

De novo motif analysis on the promoters of upregulated genes from RNA-SEQ data showed high enrichment for multiple structurally overlapping HSEs (figure 5.11). From previous studies on the structure and the form active HSFs *in vivo* which indicated that each HSF trimer binds to inverted repeats of the core consensus sequence nGAAn where each HSF monomeric unit binds to one nGAAn sequence (Åkerfelt et al., 2010; Mishra et al., 2002; Ahn et al., 2001; Liu and Thiele, 1999). It can be concluded that it is not possible for an AtHSFA1b trimers to bind to all of those element s simultaneously. The possible scenarios are that either an active AtHSFA1b trimer would bind to those elements interchangeably or it binds to one preferred element in higher affinity than other HSEs. Integrating ChIP-SEQ with RNA-SEQ provided what seems to be the answer to this problem. Only the genes that were upregulated from RNA-SEQ were intersected with AtHSFA1b target genes from ChIP-SEQ. Then instead of analysing the entire promoters of those genes, short sequences that represent AtHSFA1b binding sites from ChIP-SEQ were analysed using *de novo* motif discovery. The result showed one unique HSE. This unique form could be the preferred binding *cis*-element for AtHSFA1b *in vivo*. Furthermore, the fact that this element was retrieved from target genes that showed upregulation as a result of *AtHSFA1b* overexpressing indicates that this could be the functional element for AtHSFA1b. It is also worth mentioning that the structure of the HSE discovered using this method is in an agreement to a large extent with the structure of the HSF1b element suggested by Bechtold et

al., (2013)

The other part of the argument is that those multiple HSEs discovered on the promoters of the upregulated genes in 35S-AtHSFA1b::mRFP plants could be part of the homotypic clustering phenomenon. However, homotypic clustering involves appropriate spacing between *cis*-elements on the same promoter which allows simultaneous multiple TF molecules to bind on the same promoter (Whitefield et al., 2012; Gotea et al., 2010). As mentioned above, the multiple HSEs discovered on the promoters of the upregulated genes were structurally overlapping that it may not be possible for multiple AtHSFA1b trimers to bind to all of those elements at the same time. However, this might be an important indication for the possible events that occur under heat stress. Those other HSEs might be preferred binding elements for other AtHSFs. For instance, stress inducible AtHSFs might have higher affinity to those elements than AtHSFA1b which by turn might result in displacing AtHSFA1b from the promoters of its target genes in favour of the binding of other AtHSFs that may be involved in the regulation of *AtHSPs* and other stress associated genes at later stages of stress conditions.

CHAPTER 6

The *A. thaliana* HSFA1b gene in Yeast

6.1 Introduction:

Analysis of AtHSFA1b binding sites in the *A. thaliana* genome and the AtHSFA1b-regulated transcriptome revealed some surprising insights about the AtHSFA1b transcriptional regulatory network. So far, it has been shown that AtHSFA1b has the potential to be more than just a regulator of HSR in *A. thaliana* (Chapters 4 and 5). Briefly, this research demonstrated evidence that AtHSFA1b might be involved in broad functions including developmental processes in *A. thaliana* and showed that the involvement of AtHSFA1b in the regulation of HSR is limited to possibly a very short duration of time (Chapters 4 and 5). However, it is important to prove that what was observed in previous parts of this study is translated into real biological events. Therefore, confirming those possible scenarios of AtHSFA1b network in a biological system and observing those involvements of AtHSFA1b in development and its limited involvement in the regulation of HSR in a biological system is a crucial part to verify pervious findings in this research.

Previous parts of this study strongly suggest that AtHSFA1b directly controls the expression of a set of genes which in turn control the expression of other sets of genes leading to possibly terminal genes in the transcriptional cascade. However, it is not possible to establish the link between AtHSFA1b and its indirect targets from the experiment carried out so far. This aspect of the study is important to establish the depth of the AtHSFA1b network and the possible cascade of events within the network that might lead to the activation of the indirectly regulated target genes of AtHSFA1b which might be of great importance.

Another important finding that begs to be proven in this study is the limited involvement of AtHSFA1b in the regulation of HSR. The result obtained from ChIP-SEQ and ChIP-PCR from heat

stressed 35S-AtHSFA1b::mRFP plants which showed reduced genome-wide occupancy of AtHSFA1b (Chapter 4) combined with the maintained or, in some instance, increased expression levels of heat and stress-associated genes in 35S-AtHSFA1b::mRFP plants under heat stress may create doubts about the outcome of the ChIP-SEQ experiment despite showing evidence that the expression of those genes might be controlled by other AtHSFs under heat stress (Chapter 5). Therefore, translating the outcomes of those experiments into real biological events is crucial to confirm and add credibility to those finding from Chapters 4 and 5.

Baker's yeast (*S. cerevisiae*) is a true eukaryotic unicellular model organism that has been used extensively to study the functions of numerous proteins from different organisms including plants (Uetz et al., 1999; Zinser et al., 1991; Fields and Song, 1989). For instance, yeast has been used in functional studies of human cell cycle proteins (Serrano et al., 1993; Harper et al., 1993; Pines and Hunter; 1989). Furthermore, the yeast system is widely used to determine the ability of TFs to activate gene expression *in vivo* through their interactions with their specific *cis*-elements on promoter regions fused to yeast reporter genes in what is known as yeast one-hybrid (Y1H) system (Deplancke et al., 2004; Kim et al., 2002; Li and Herskowitz,1993). Moreover, yeast is also used to study protein-protein interactions *in vivo* with a technique named yeast two-hybrid (Y2H) (Walhout et al., 2001; Ito et al., 2000; Sato et al., 1994). However, one of the most informative experiments is using *S. cerevisiae* to study the functional conservations of proteins from other biological systems in yeast in rescue or complementation experiments (Gécz et al., 2003; Piao et al., 1999; Schirmer et al., 1994; Liang et al., 1994). This experiment relies on the deletion of genes that code for proteins that are essential for survival or tolerance of yeast cells to various forms of stresses and test if genes that code for homologues of those proteins from other organisms can rescue the deletion mutations in yeast

(Gécz et al., 2003; Piao et al., 1999; Schirmer et al., 1994; Liu et al., 1997; Liang et al., 1997). For instance, it was shown the expression of hHSF2 but not hHSF1 can rescue the *yhsf* deletion mutation which was shown to be lethal to yeast cells (Section 1.2.1.1; Liu et al., 1997; Jakobsen and Pelham, 1988).

This part of the research was mainly designed to examine three aspects. The first is to use Y1H to identify potential TF candidates from the AtHSFA1b network that are responsible for the regulation of genes that are indirectly regulated by AtHSFA1b as well as testing the interaction between AtHSFA1b and the proposed HSE1b element from (Bechtold et al., 2013). The second is to examine the ability of AtHSFA1b to functionally complement the *yhsf* deletion mutation. The third test is to examine the possible involvement of AtHSFA1b in the regulation of HSR in yeast.

6.2 Results

6.2.1 Yeast one-hybrid and AtHSFA1b indirect target genes:

Three 1000bp promoter regions of three genes indirectly regulated by AtHSFA1b were used, in addition to an HSE1b-containing promoter fragment to test the interaction between AtHSFA1b and the HSE1b element *in vivo*. In order to remove the distance effect on the ability of TFs to activate the reporter gene in yeast, each promoter was amplified using PCR into three overlapping fragments. Each fragment was then cloned upstream of the histidine reporter gene (*HIS3)* in the yeast bait vector pHIS3LEU2.

The yeast one-hybrid (Y1H) system is based on the ability of TFs to activate the expression of the *HIS3* gene via bindings to their cognate *cis*-elements on promoter fragments of target genes fused upstream of the histidine (*HIS3*) reporter gene. The activation of the *HIS3* gene results in survival and growth of yeast colonies in SD media lacking the amino acid histidine (SD-H). The Y1H consist of three major components, Bait vector containing the reporter *HIS3* reporter genes, Prey vector expressing TF fused to GAL4 activation domain (AD-TF), and yeast strain that is unable to grow on synthetic dropout media lacking the amino acid histidine (SD-H). The bait vector containing the promoter fragment-reporter gene fusion is transformed into an appropriate yeast strain compatible with the auxotrophic selectable markers in the bait vector. Then the bait strain is mated with another yeast stain containing the prey vector on minimal SD media lacking the three amino acids synthesised by the auxotrophic selectable markers present in both the bait and the prey vectors in addition to the amino acid synthesised by the reporter gene (figure 6.1)

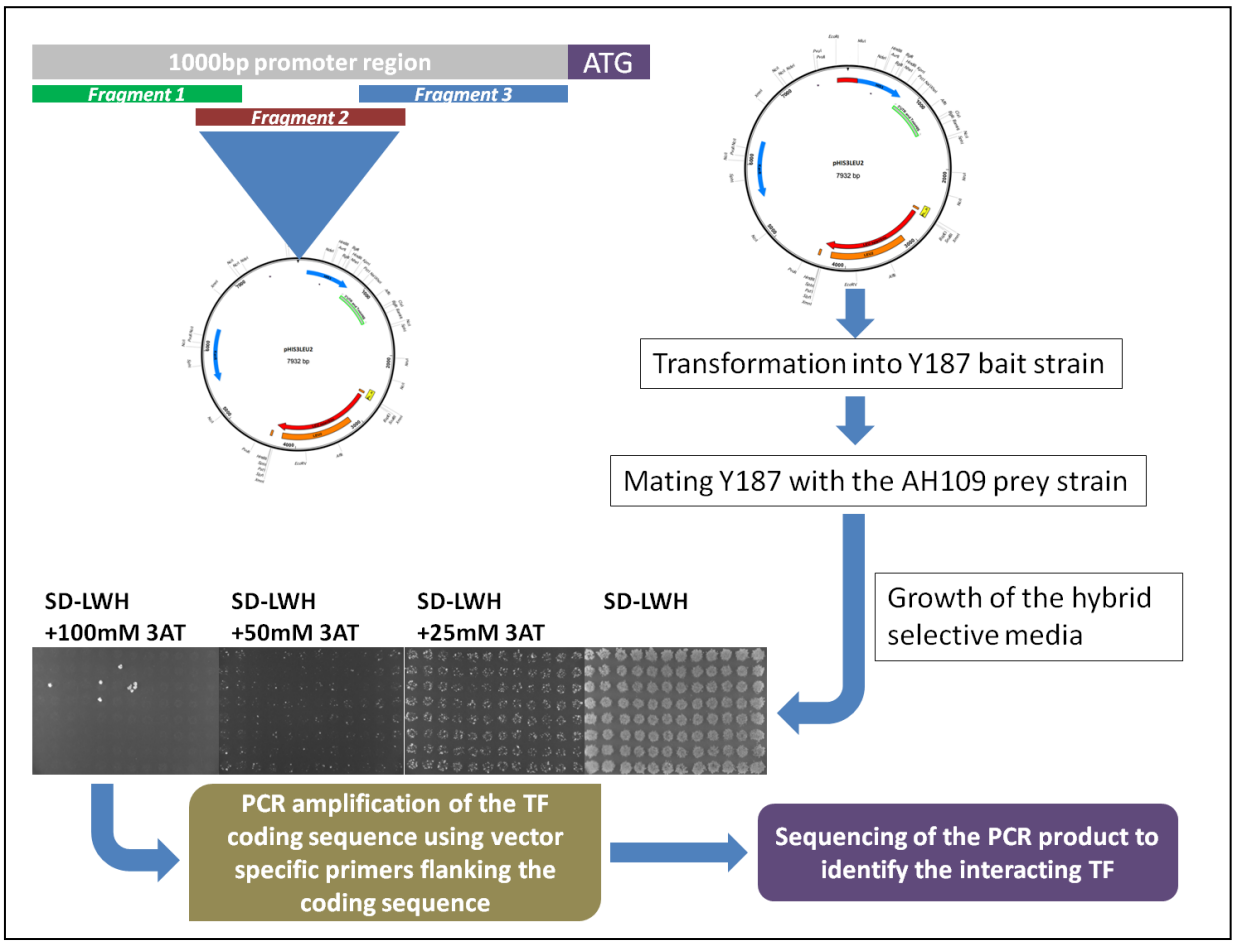


Fig.6.1. Summary of the Y1H experimental design. The diagram shows the general PCR strategy used to divide each promoter into three overlapping fragments then fusing each fragment upstream of the *HIS3* reporter gene in the yeast plasmid pHIS3LEU2 and the subsequent steps for the Y1H assay.

Only 3 promoter fragments showed clear positive results in the Y1H assay. The rest of the promoter fragments showed either no interaction or showed very high levels of auto-activation that could not be avoided with 3AT concentration up to 100 mM (figure 6.2).

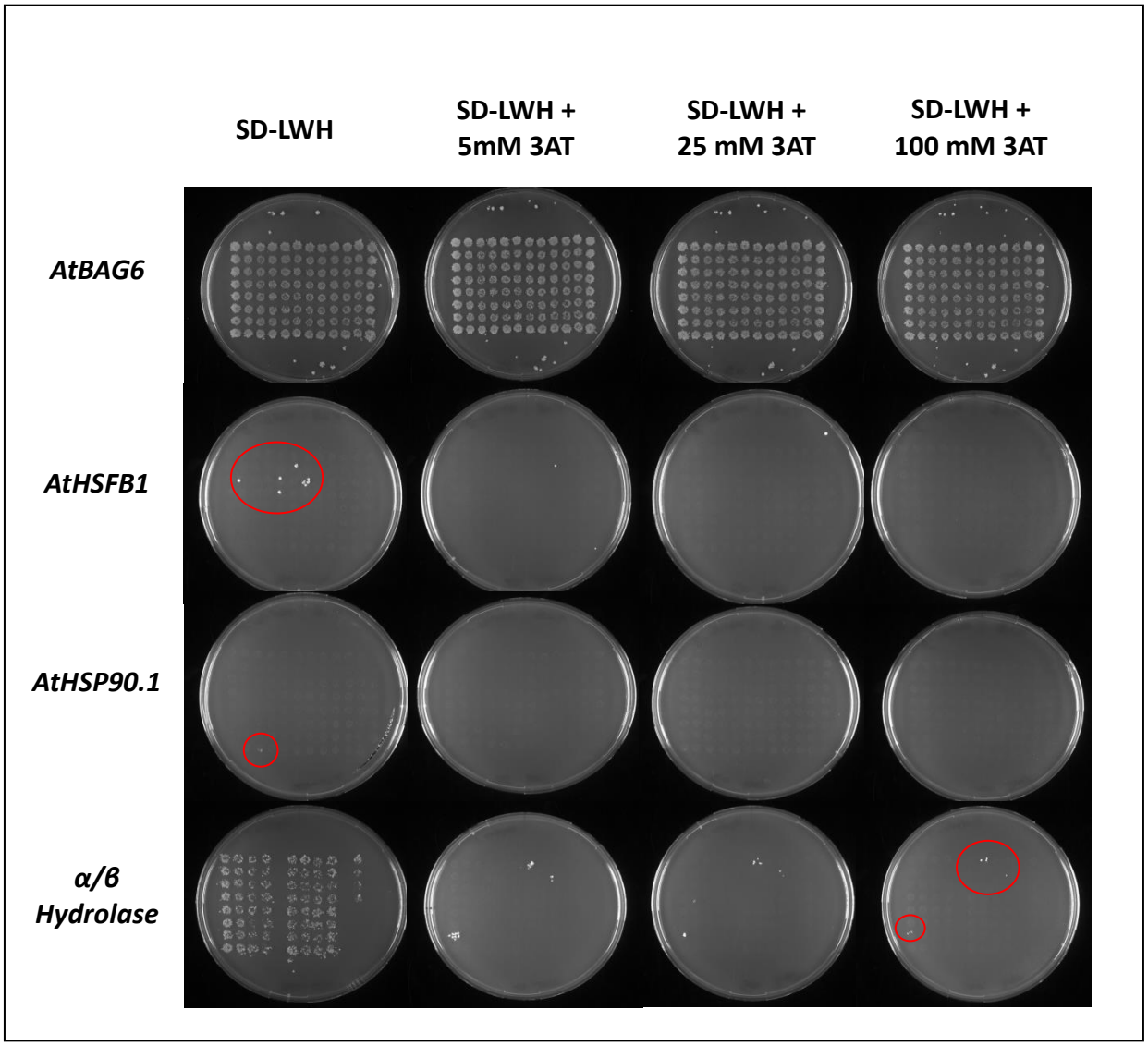


Fig.6.2. The output of Y1H screen on selective plates SD-LWH (-/+ 3AT). Positive Y1H colonies are highlighted in red circles. The first row (*AtBAG6*) shows an example of a strong auto-activation of the *HIS3* reporter gene that could not be avoided even with a concentration as high as 100 mM of 3AT. Second to the fourth are positive results that showed yeast growth with and without 3AT indicating potential positive interaction between AtTFs and the promoter fragments used in the assay. Only single yeast colonies that represent real TF promoter interaction were selected for PCR and sequencing analysis. Each part of the Y1H was repeated three times with no difference in the output of each repeat.

The yeast one-hybrid assay showed that 7 TFs can interact with the 3 promoter fragments tested (table 6.1). The interacting TFs comprised members of 4 TF families in *A. thaliana* including CCCH-type zinc finger, high mobility group (AtHMG) box protein, AtTCP, and AtANAC TFs.

Promoter fragments	**Interacting Transcription Factors**
AT4G36990 (*AtHSFB1*) [-1bp-375bp]	**AT3G19360**: Zinc finger (CCCH-type) family protein. **AT1G76110**: HMG (high mobility group) box protein with ARID/BRIGHT DNA-binding domain.
AT5G52640 (*AtHSP90.1*) [665bp-1000bp]	**AT1G58100**: Encodes TCP8, belongs to the TCP transcription factor family known to bind site II elements in promoter regions. **AT3G12910**: NAC (No Apical Meristem) domain transcriptional regulator superfamily protein.
AT1G68620 (*α/β Hydrolase*) [1bp-150bp]	**AT3G12910**: NAC (No Apical Meristem) domain transcriptional regulator superfamily protein. **AT4G27410**: Encodes a NAC transcription factor induced in response to dessication. It is localized to the nucleus and acts as a transcriptional activator in ABA-mediated dehydration response. **AT1G61110**: NAC domain containing protein 25 (NAC025); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription

Table 6.1. *A. thaliana* TFs that interacted with the promoter fragments tested. Promoter fragments that showed positive interactions with *A. thaliana* TFs in Y1H screens. Left column shows the AGI codes and the promoter fragment that passed the Y1H assay. Right column shows the interacting *A. thaliana* TFs with the tested promoter fragments. All TF annotations were retrieved from the *A. thaliana* database (TAIR; http://www.arabidopsis.org)

Interestingly, members of AtANAC TF family showed interactions with more than one promoter fragment. The RNA-SEQ data from 35S-AtHSFA1b::mRFP plants under no stress showed upregulation of a number of TF that belong to that group including ANAC075 (RD26) (table2)

AGI code	Gene name	Bound by AtHSFA1b	FPKM in wild type No stress	FPKM in 35S-AtHSFA1b::mRFP No stress	Fold Change (Log2)
AT1G01010	*AtANAC001*	No	1.53	4.49	1.55
AT1G34180	*AtANAC016*	No	0.39	1.52	1.97
AT1G52890	*AtANAC019*	Yes	0.90	2.86	1.66
AT3G04060	*AtANAC046*	No	0.30	1.15	1.94
AT4G01550	*AtANAC069*	Yes	4.97	9.22	0.89
AT3G04070	*AtANAC047*	Yes	0.21	0.75	1.84
AT1G69490	*AtANAC029* (*AtNAP*)	Yes	13.02	28.10	1.11
AT4G27410	*AtANAC072* (*AtRD26*)	Yes	9.14	26.0	1.51

Table 6.2. Members of *AtANAC* TFs controlled by AtHSFA1b. The table shows a list of upregulated *AtANAC* TFs in 35S-AtHSFA1b::mRFP plants under no stress with their FPKM values in wild type and 35S-AtHSFA1b::mRFP and the log2 fold expression change. The gene highlighted in red colour is a TF gene that was present in the Y1H screen.

The yeast one-hybrid screen did not show any interaction between AtHSFA1b and the α/β Hydrolase promoter fragment that contains the HSE1b element. Moreover, there was no presence of interaction between any AtHSF present in the Y1H library and the promoter fragments tested although two of them (*AtHSFB1* and *AtHSP90.1*) are known to contain HSEs. To test that, promoter motif scan was carried out on the promoter fragments tested in Y1H. The motif analysis test showed that each promoter fragment tested contains at least one potential HSE. Yet, no interaction between AtHSFs present in the Y1H library and those HSEs was observed (figure 6.3).

Fig.6.3. AtHSFs did not interact with promoter fragments that contain HSEs. Possible HSEs present in the promoter fragments tested in yeast one-hybrid assay. The figure shows the location of possible HSE candidates on the promoter fragments tested in Y1H assay.

Interestingly, all the TFs that showed positive interaction with the promoter fragments in the yeast one-hybrid screen were plant exclusive TFs, such as AtANAC, AtTCP, AtHMG and AtCCC-H TFs (Yamasaki, et al., 2013). None of the conserved TFs between plants and yeast, including HSFs, showed any interaction with the promoter fragments tested.

6.2.2 Functional analysis of AtHSFA1b in yeast:

The result obtained from Y1H was valuable in the sense that it gave a deeper view of the AtHSFA1b transcriptional network and it showed a potential link between AtHSFA1b and AtANAC072 (AtRD26). It showed that other TFs candidates have possible involvement in the regulation of AtHSFA1b direct and indirect target genes. However, the absence of interaction between AtHSFA1b and other AtHSFs with HSEs including the suggested HSE1b element (Bechtold et al., 2013) raised a number of concerns. The first concern was about the result of the ChIP-PCR if it was a real result or an experimental artefact as one of the fragments tested from the promoter of the genes (*α/β Hydrolase*) showed positive interaction between AtHSFA1b and HSE1b in ChIP-PCR (Section 3.2.2). The second concern was about AtHSFs whether they are functional in yeast or not especially when they are tagged with GAL4 activation domain. The first concern was quickly ruled out as ChIP-SEQ added an extra confirmation to the ChIP-PCR result (Section 4.2.4). The effect of the tag on the function of AtHSFA1b was also ruled out as it was shown in this study and other studies that fluorescent and other protein tags do not affect the function of AtHSFs including AtHSFA1b (Bechtold et al., 2013; Prändl et al., 1998). The only possibility left was whether AtHSFA1b is functional in yeast cells.

Previous studies have shown that yHSF is required for the survival of yeast cells under normal growth conditions and deletion of the yHSF is lethal to yeast cells (Section 1.2.1.1; Liu et al., 1997; Jakobsen and Pelham, 1988). Researchers have taken advantage of this condition and used it as a tool to study functional conservation between HSFs in yeast by replacing the endogenous yHSF with other HSFs from other organisms (Batista-Nascimento et al., 2011; Liu and Thiele, 1999; Liu, et al., 1997; Boscheinen et al., 1997). The majority of the complementation experiments showed that HSFs can rescue the *yhsf* deletion mutation by

showing the ability of yeast cells to grow normally. However, there were instances where other HSFs failed to complement loss of function of yHSF. For instance, it has been reported that hHSF2 but not hHSF1 can functionally substitute the yHSF in *S. cerevisiae* (Liu et al., 1997).

The questions asked in this part of the study are, can AtHSFA1b functionally complement the loss of the endogenous yHSF in yeast under no stress? The second question is to what extent AtHSFA1b can cover the various functions of yHSF as a major regulator of stress response in yeast? The answers to these question might provide insights about the possible involvement of AtHSFA1b in the regulation of HSR in yeast in an attempt to show whether the result obtained from ChIP-SEQ on heat stressed 35S-AtHSFA1b::mRFP plants is a real biological event.

6.2.2.1 AtHSFA1b can functionally complement the loss of endogenous yHSF in yeast:

This part of the research has taken advantage of the conditionally *yhsf* knockout mutant PS145 strain (Liu et al., 1997) to test the ability of AtHSFA1b to functionally complement the loss of the endogenous yHSF. PS145 was a kind gift from Prof. Dennis Thiele, Duke University, NC, USA. PS145 (*ade2-1 trp1, can1-100, leu2-3,-112, his3-11,-15, ura3, Δhsf::LEU2, YCpGAL1-yHSF*) is a *yhsf* deletion mutant; however, as described in Section (1.1.2.1), deletion of yHSF is lethal to yeast cells. To get around this problem, the strain was transformed with a yeast centromere plasmid containing yHSF under control of inducible GAL promoter (*YCpGAL1-yHSF::URA3*) (Jakobsen and Pelham, 1988). The GAL1 promoter controlling yHSF in PS145 contains the GAL1 TF binding site. The expression of *GAL1* is only induced when the yeast cells are grown on media supplemented with galactose as a carbon source. However, the expression of *GAL1* is repressed when galactose is replaced with glucose as a carbon source. Therefore, if it is not expressing other HSFs, the strain PS145 grows only on media supplemented with galactose (Liu et al.,

1997).

To test the ability of other HSFs to functionally complement the loss of *yhsf* in PS145, the coding sequences of HSFs are cloned downstream of constitutive promoters such as ADH (weak) or GPD (strong) (Mumberg et al., 1995). If the yeast strain PS145 expressing HSF from another organism under control of a constitutive promoter grows on media supplemented with glucose, then the HSF expressed in PS145 can functionally complement the loss of *yhsf* (Liu et al., 1997). In a similar experiment, the *AtHSFA1b* genes was cloned downstream of the GPD constitutive promoter in the plasmid pAG424 (Alberti, et al., 2007). The plasmid pAG424 also contains an eYFP coding sequence downstream of the coding sequence of AtHSFA1b. Figure 6.4 shows a summary of the HSF functional complementation experimental strategy.

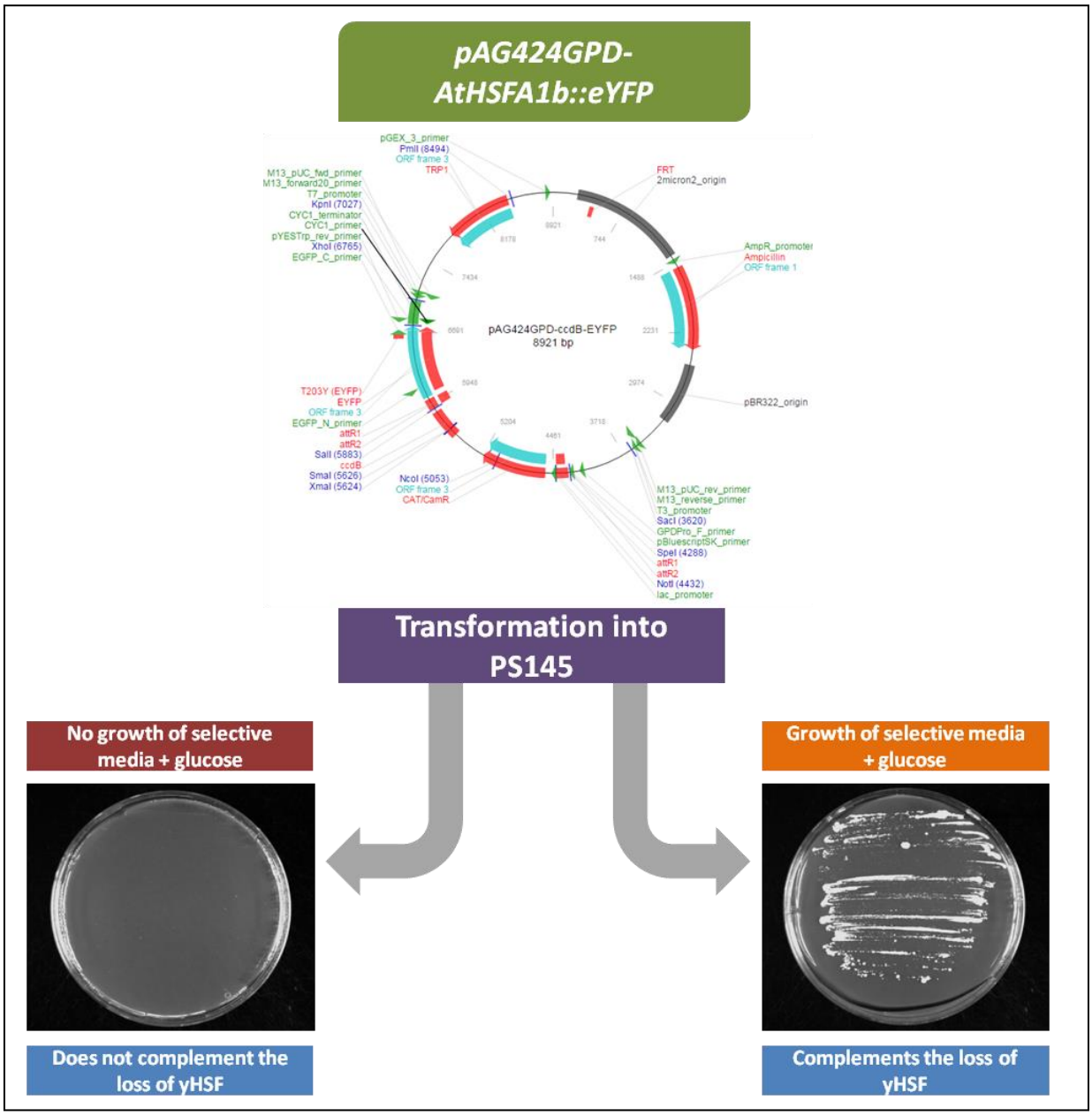


Fig.6.4. Summary of the procedure of the HSF functional complementation experiment in yeast. The protein coding sequence of AtHSFA1b was cloned downstream of the constitutive promoter GPD in the yeast expression vector pAG424-ccdB-eYFP. Subsequently the vector was transformed into the *yhsf* knockout strain PS145. Functional complementation was tested by growing the yeast strain PS145 expressing AtHSFA1b (GPD-AtHSFA1b) in selective media supplemented with glucose side by side with the same knockout strain expressing yHSF under the same constitutive promoter (GPD-yHSF) as a control. The pAG424GPD-ccdB-eYFP map was obtained from Addgene website (https://www.addgene.org/14224/).

Expressing the *AtHSFA1b* gene in the yeast strain PS145 complemented the viability defect due to the loss of the single endogenous *yhsf* in that strain. The PS145 expressing *AtHSFA1b* GPD-AtHSFA1b grew well on SD-media supplemented with glucose and lacking Tryptophan (SDDEX-W). The knockout strain PS145 without the AtHSFA1b gene, here named GAL1-yHSF, failed to

grow on SDDEX-W (figure 6.5).

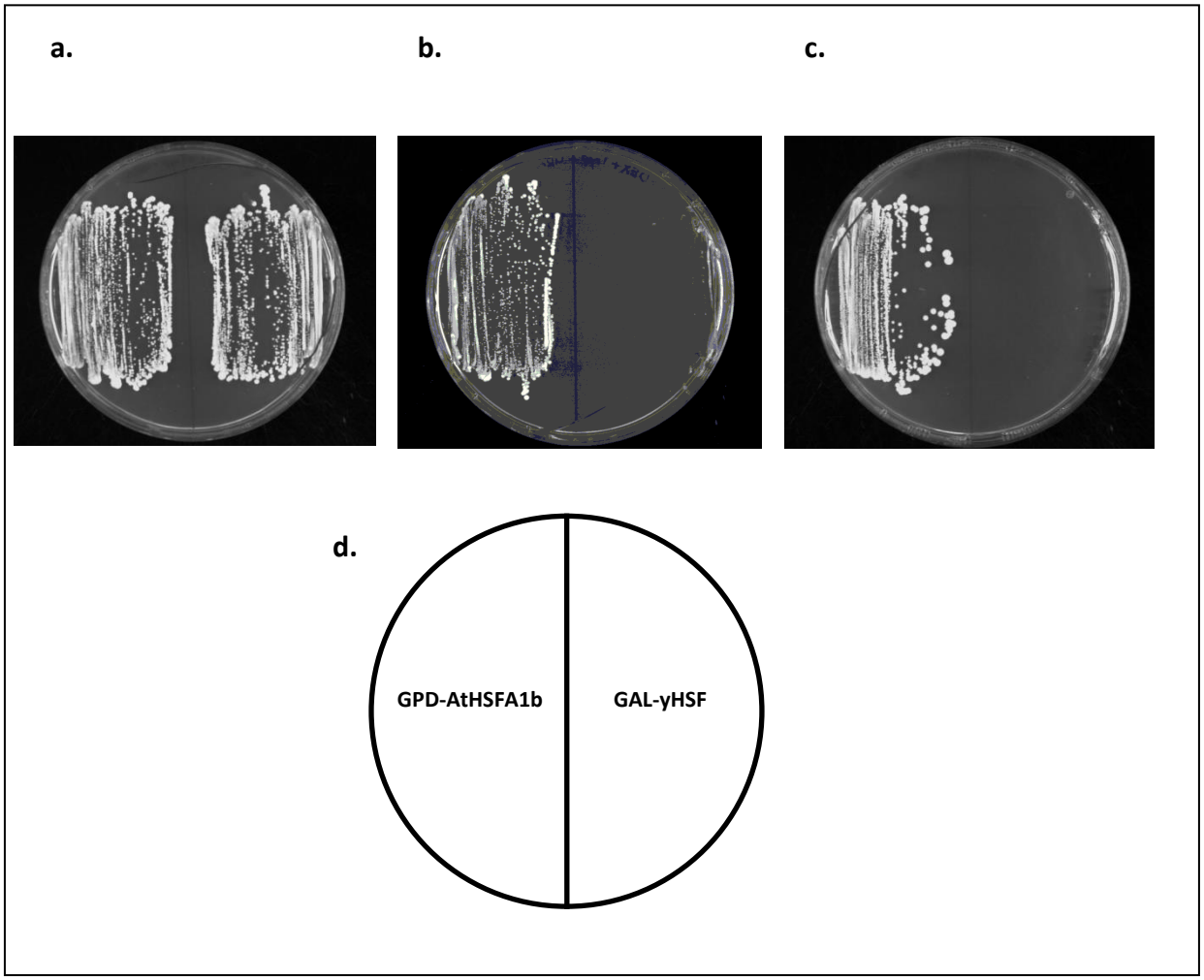


Fig.6.5. AtHSFA1b functionally complements the *yhsf* deletion in PS145. Growth of yeast cells on different types of media at 28°C. (**a.**) Both GPD-AtHSFA1b and GAL1-yHSF grown on complete media supplemented with galactose (YPGA). (**b.**) grown on complete media supplemented with glucose (YPDA). (**c.**) grown on synthetic minimal dropout media supplemented with glucose and lacking tryptophan (SDDEX-W). (**d.**) plate template showing the positions of GPD-AtHSFA1b and GAL1-yHSF cells on each agar plate used.

The growth of the yeast strain GPD-AtHSFA1b was quantified and compared to the growth of the same strain constitutively expressing the endogenous yHSF (GPD-yHSF). The result showed no difference between the growth rates the strains GPD-AtHSFA1b and GPD-yHSF expressing (figure 6.6)

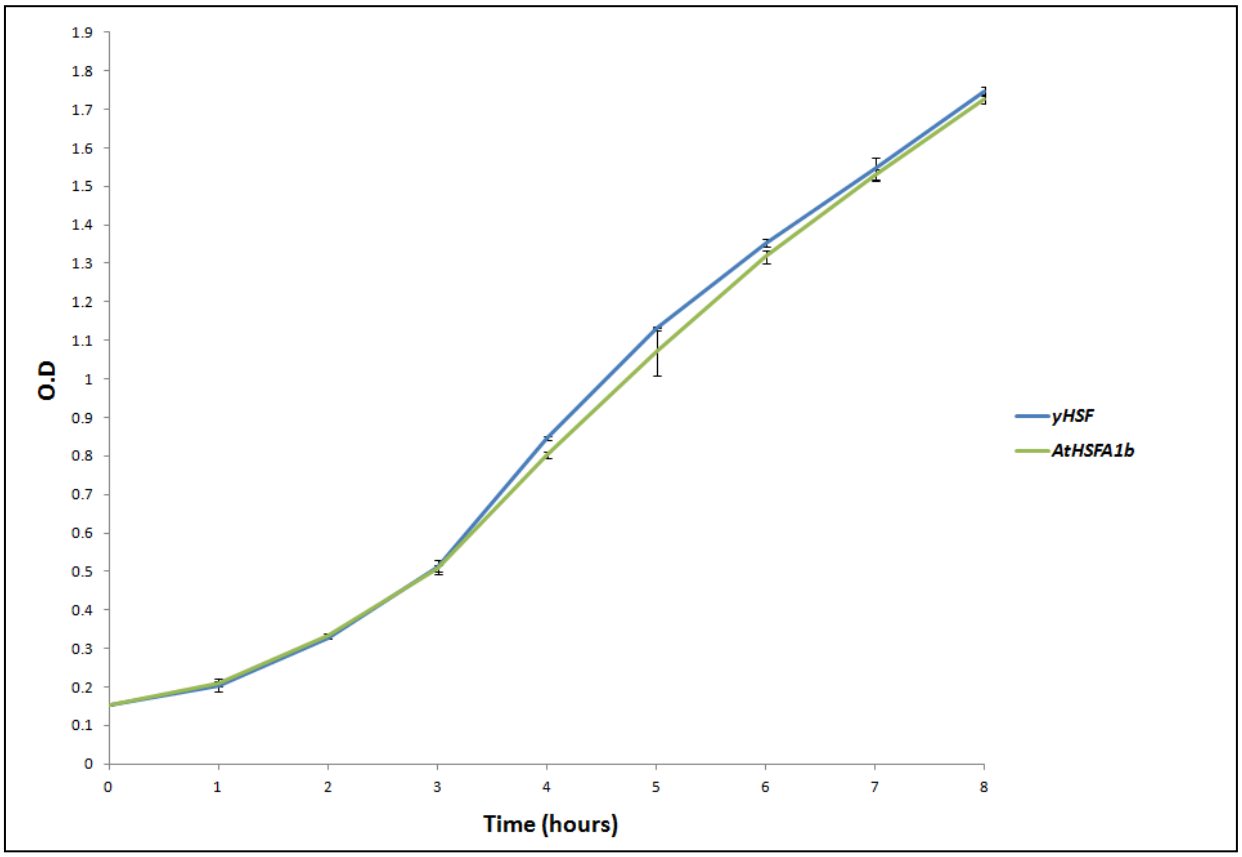


Fig.6.6. The growth rate of GPD-AtHSFA1b is identical to GPD-yHSF under normal growth conditions. Growth curve showing comparison between the rates of growth of the GPD-AtHSFA1b and GPD-yHSF at 28°C in liquid SDDEX-W. Y-axis is the optical density of the yeast cells (O.D) at 600 nm, X-axis show the time points where the O.D measurement was taken, values are means ± SD (n = 3). (This experiment was carried out by D. Kumar and S. Prasad).

6.2.2.2 The yeast strain GPD-AtHSFA1b is intolerant to heat stress:

In order to examine the ability of AtHSFA1b to functionally complement the function of yHSF as a regulator of yeast HSR, the same growth experiment was carried out at an elevated temperature. Both GPD-AtHSFA1b and GPD-yHSF were grown in SDDEX-W media at 37°C. The yeast strain GPD-yHSF grew at 37°C but GPD-AtHSFA1b failed to grow under the same condition (figure 6.7).

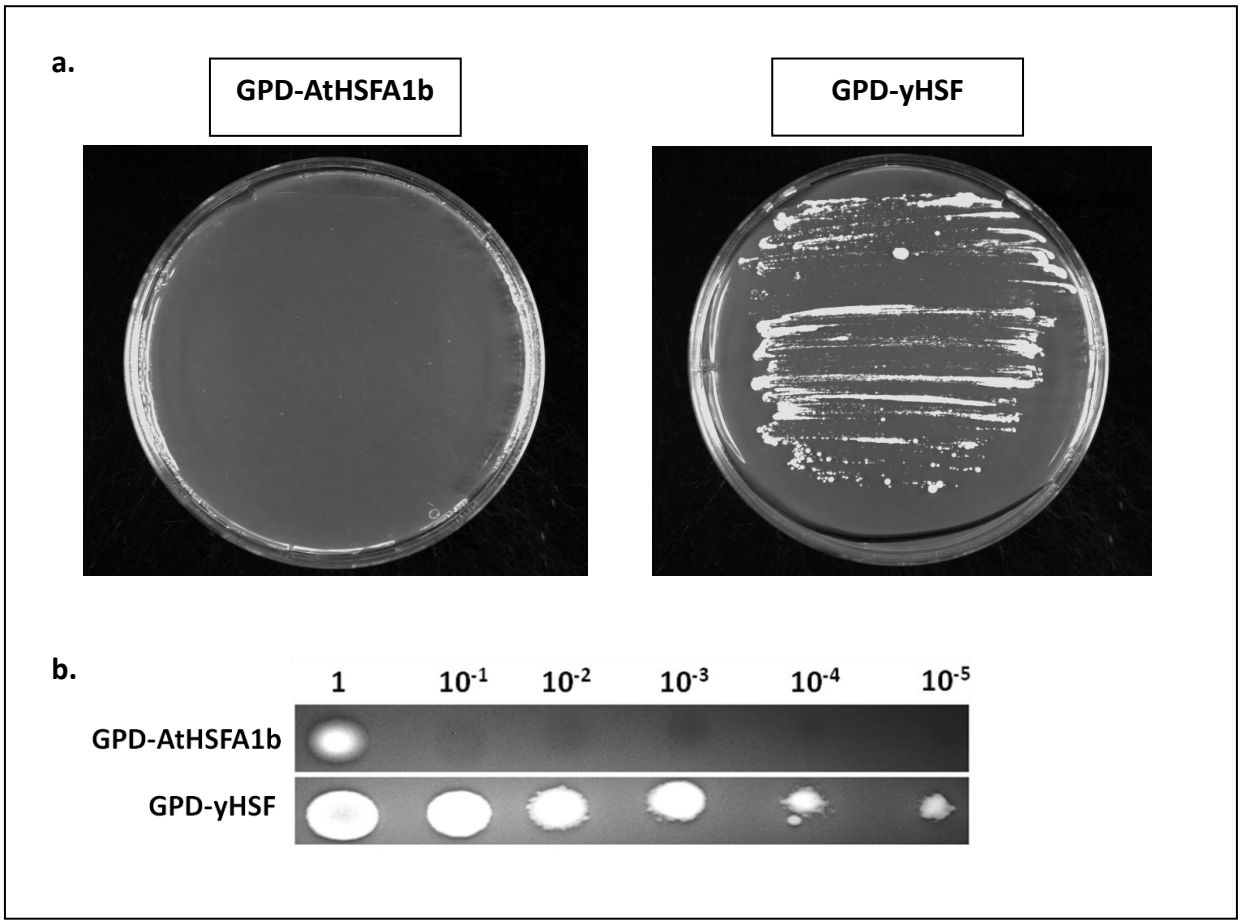


Fig.6.7. AtHSFA1b does not regulate HSR in yeast (a.) GPD-AtHSFA1b (left photo) and GPD-yHSF (right photo) grown on SDDEX-W media plates at 37°C. (**b.**) comparison between the growth of serially diluted GDP-AtHSFA1b and GPD-yHSF liquid cultures at 37°C. (experiment was carried out by D. Kumar and S. Prasad).

6.2.2.3 Elevated temperature inhibits the function of AtHSFA1b in yeast:

The growth of the yeast strain GPD-AtHSFA1b was measured and compared to the growth of the strain GPD-yHSF under shifted temperature condition (grown at 28°C then temperature shifted to 37°C). The growth of the strain GPD-AtHSFA1b was stunted when the growth condition was shifted from 28°C to 37° whereas GPD-yHSF maintained normal growth rate under the same conditions (figure 6.7).

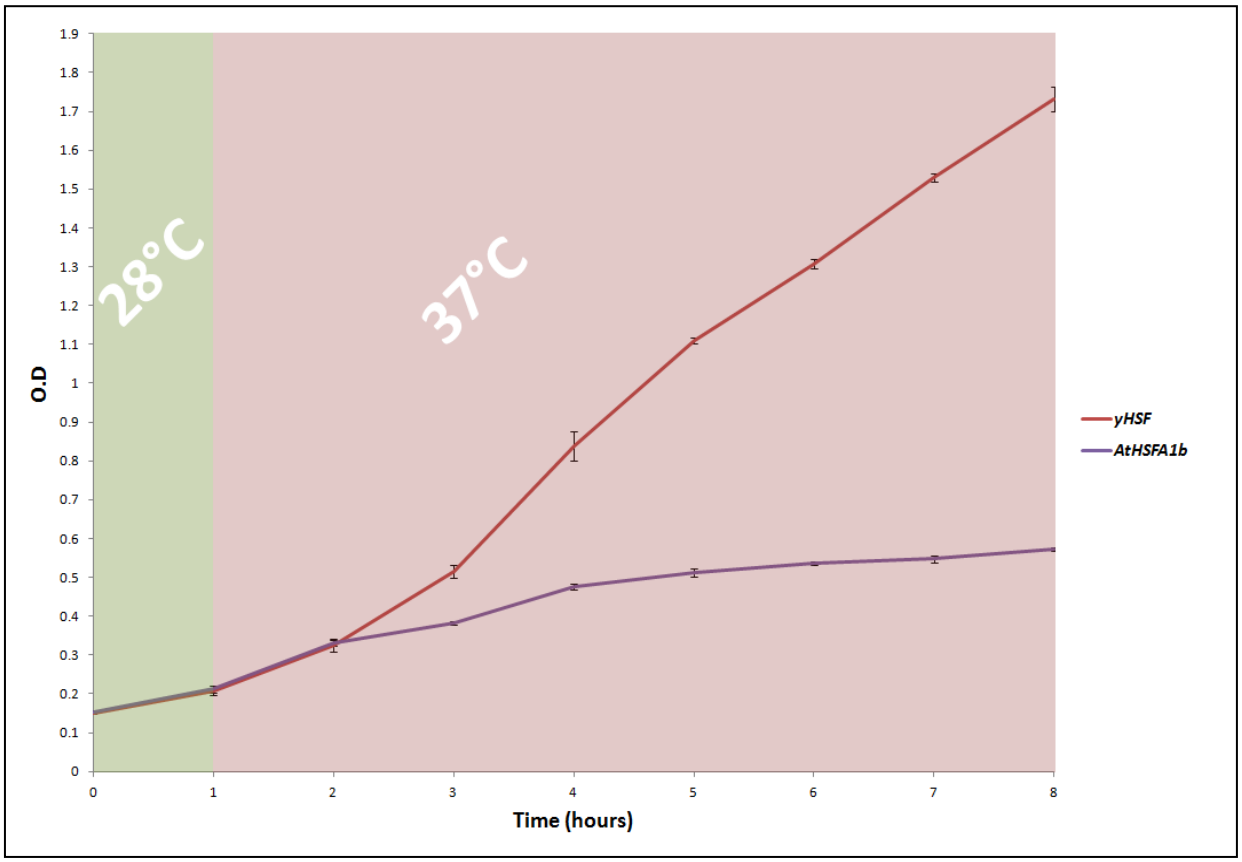


Fig.6.8. Effect of elevated temperature on the function of AtHSFA1b in yeast. The growth of the yeast strain GPD-AtHSFA1b was measured and compared to the growth of GPD-yHSF on liquid SDDEX-W media under change in temperature conditions. Y-axis is the optical density of the yeast cells (O.D) at 600 nm, X-axis shows the time where the O.D measurements were taken. (This experiment was carried out by D. Kumar and S. Prasad). Values are means ± SD (n = 3).

6.3 Discussion:

6.3.1 Possible involvement of other TFs in the AtHSFA1b network:

6.3.1.1 AtTCPs might not be involved in the AtHSFA1b network

The Y1H experiment showed one positive interaction between the promoter fragment of the gene *AtHSP90.1*, which is also a target gene of AtHSFA1b, with the TF AtTCP8 (table 6.1). This result is consistent with what was shown in Sections (4.2.6 and 5.2.4) about the possible roles of AtTCPs in regulation of AtHSFA1b direct and indirect genes. In both Sections (4.2.6 and 5.2.4) the analysis showed a high enrichment for the site II element, which is a DNA binding site for AtTCPs (Giraud, et al., 2010; Welchen and Gonzalez, 2006; Trémousaygue, et al., 2003), in some of the promoters that contain HSEs and some others that are not targeted by AtHSFA1b.

Despite the growing evidence from this study of the involvement of AtTCPs in the AtHSFA1b transcriptional regulatory network, none of the *AtTCP* genes showed upregulation in 35S-AtHSFA1b::mRFP plants under no stress and under heat stress. In fact, the general expression pattern of *AtTCP* genes was towards downregulation in 35S-AtHSFA1b::mRFP plants under no stress and was further repressed in wild type and 35S-AtHSFA1b::mRFP plants under heat stress (figure 6.8). There were some exceptions to the general expression theme of *AtTCPs* in plants overexpressing *AtHSFA1b* and heat stressed wild type and 35S-AtHSFA1b::mRFP plants. However, this is due to the algorithm used to generate the heat map where it somehow shows an exaggeration in the expression change (figure 6.9). The changes in the FPKM values of those members in unstressed 35S-AtHSFA1b::mRFP and heat stressed wild type and 35S-AtHSFA1b::mRFP plants were marginal compared to the FPKM values in wild type under no stress condition.

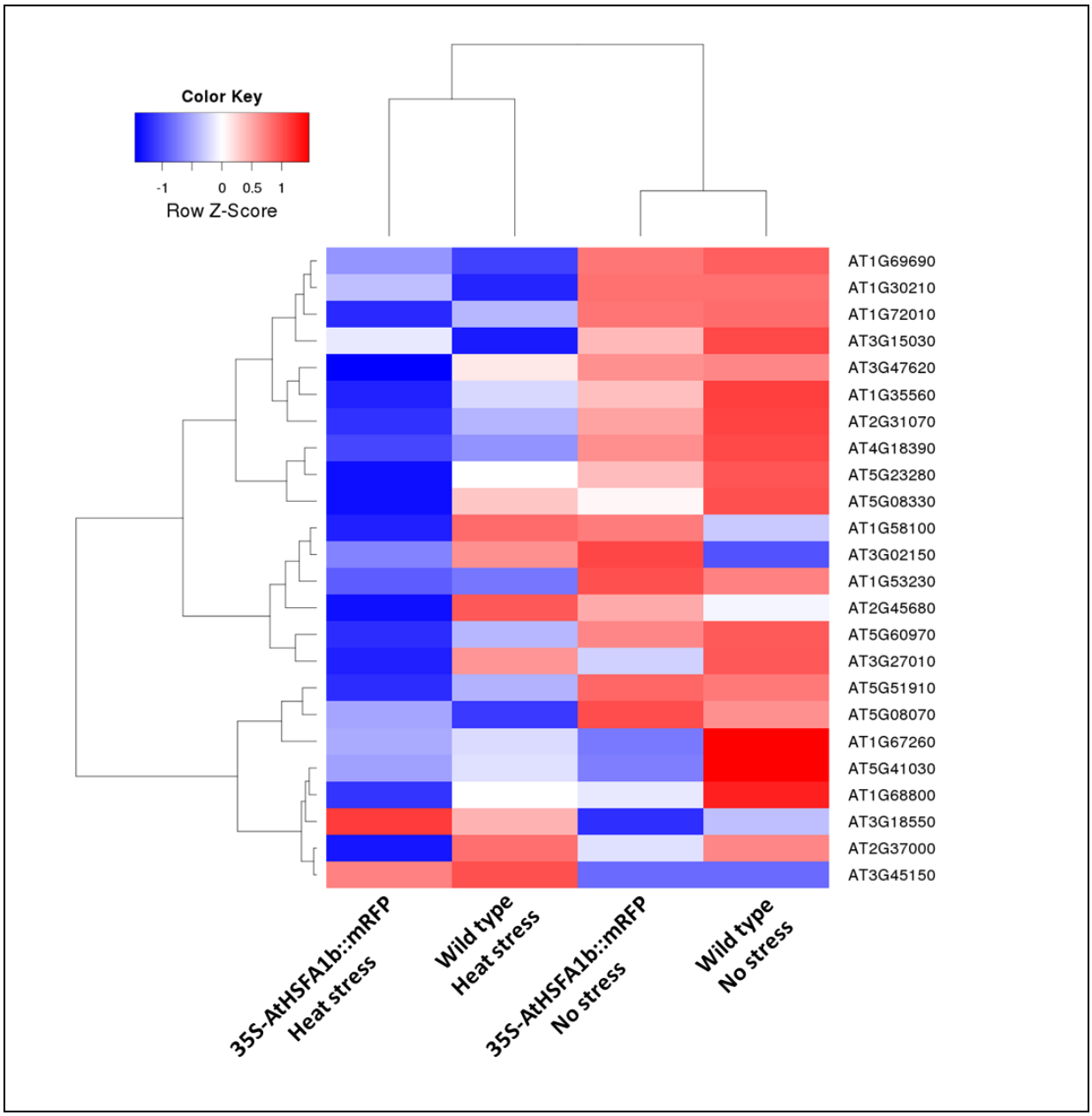


Fig.6.9. The expression of *AtTCP*s is repressed in plants overexpressing *AtHSFA1b* and by heat stress. Heat map showing comparison between the averages of the absolute FPKM expression values obtained from RNA-SEQ experiment of *AtTCP*s in all plants under all treatments.

6.3.1.2 Possible involvement of AtANACs in the AtHSFA1b network

The Y1H assay showed what might be an indication of involvement of AtANAC TF in the AtHSFA1b transcriptional network (table 6.1). The results obtained from Y1H experiment when integrated with transcriptomics data from unstressed 35S-AtHSFA1b::mRFP plants strongly suggests a direct involvement of AtANAC TFs in the AtHSFA1b regulatory network (table 6.2). The expression of *AtANAC* genes in plants overexpressing *AtHSFA1b* under no stress condition

was variable compared to that of *AtTCPs*. There were members of *AtANAC* gene family upregulated as a result of overexpression of *AtHSFA1b* under no stress condition some of which were bound by AtHSFA1b. Furthermore, the expression pattern of members of *AtANAC* genes changes with in response to heat stress suggesting a possible reconfiguration of that sub-network. Only one member of AtANAC TFs, AtRD26, identified in the Y1H assay was upregulated in 35S-AtHSFA1b::mRFP plants under no stress conditions. Furthermore, AtRD26 was bound by AtHSFA1b in 35S-AtHSFA1b::mRFP plants under no stress (appendix B). This suggested that AtRD26 might be linking AtHSFA1b to some of its indirect target genes.

6.3.2 AtHSFA1b can functionally complement the loss of yHSF in yeast:

Expressing *AtHSFA1b* in the *yhsf* knockout yeast strain PS145 has given a deeper view of the diverse roles of AtHSFA1b. It was shown that *AtHSFA1b* when expressed in the yeast knockout strain complements the viability defect due to the loss of the endogenous yHSF (Section 6.2.3). The ability of AtHSFA1b to functionally substitute the loss of the endogenous yHSF for cell viability and maintenance in yeast indicates that AtHSFA1b is possibly able to recognise, bind and control the expression of the same targets of yHSF required for those functions. The ability of the strain GPD-AtHSFA1b to survive under normal growth conditions suggests that AtHSFA1b is able to control the expression of genes that are required for survival and cell cycle in yeast. This result also gives an indication that some of the cell cycle and developmental genes in yeast contain HSE elements. This is consistent with the result obtained from mapping AtHSFA1b binding sites in the *A. thaliana* genome (chapter 4) where it clearly showed that AtHSFA1b does not target only stress response genes but also targets genes that are involved in plant growth and development, cell differentiation and signal transduction (Section 4.2.7).

6.3.3 AtHSFA1b is not involved in the regulation of HSR in yeast:

The heat stress experiment on the yeast stain GPD-AtHSFA1b showed that AtHSFA1b is not involved in the regulation of HSR in yeast. AtHSFA1b failed to rescue the *yhsf* deletion mutation under mild heat stress conditions (37°C). The strain GPD-AtHSFA1b showed no growth under that mild heat stress condition while the same strain expressing the endogenous *yHSF* (GPD-yHSF) showed positive growth under the same condition. Furthermore, analysis of the growth of the yeast strain GPD-AtHSFA1b clearly showed that heat stress resulted in a complete shutdown of growth when the condition was changed from 28°C to 37°C (figure 6.7). This suggests that elevated temperature might have and inhibitory effect of AtHSFA1b. More importantly, this result shows that AtHSFA1b was unable to control the expression *HSP* genes in yeast under heat stress conditions.

From the experiments carried out, it is not possible to know what molecular changes occur on AtHSFA1b in yeast. However, the results show a strong indication that there is no involvement of AtHSFA1b in the regulation of HSR in yeast. This perhaps can be interpreted as AtHSFA1b releasing its target genes under heat stress in a similar manner to the situation in *A. thaliana* under heat stress (chapter 4). This assumption is consistent with what the result obtained from mapping AtHSFA1b binding profile under heat stress in the *A. thaliana* genome where it was shown that the total number of binding sites of AtHSFA1b was extremely reduced under heat stress compared to that under no stress (Section 4.2.2). Therefore, a reasonable assumption would be that AtHSFA1b in yeast might be subject to a similar regulatory mechanism to that in plants that leads it to unbind its target genes under heat stress.

The overall output of studying the AtHSFA1b in yeast gave more information about the molecular roles of AtHSFA1b and about the functional conservation of HSFs in general between plants and yeast. The results suggest that AtHSFA1b transcriptional regulation goes beyond just the control of heat shock and stress response to cell maintenance and viability. The results also suggest that the involvement of AtHSFA1b in the regulation of HSR is very limited in both *A. thaliana* and yeast.

CHAPTER 7

Final discussion and future directions

7.1 General overview of the outcomes of this research:

So far, it has been shown that AtHSFA1b has the potential to be more than just a regulator of HSR in *A. thaliana*. For instance, the ChIP-SEQ experiment showed that AtHSFA1b is not limited to targeting genes that are involved in plant stress responses but goes beyond that to targeting genes associated with development and other important processes in *A. thaliana* (Chapter 4). The analysis of AtHSFA1b-regulated transcriptome also showed that overexpression of AtHSFA1b results in an increase in the transcript levels of a number of developmental genes (Section 5.3.3). Furthermore, the integration of ChIP-SEQ and RNA-SEQ data also showed evidence that the release of AtHSFA1b from its target regions in the *A. thaliana* genome under heat stress is associated with a general downregulation of the expression of genes involved in various non-stress related functions (Section 5.2.5). These results suggest that the role of AtHSFA1b is not limited to the activation of HSR and stress response in *A. thaliana* plants but extends to non-stress associated processes. This is consistent with what was previously suggested about the possible developmental component of AtHSFA1b (Bechtold et al., 2013).

Perhaps the most striking aspect shown in this study is the limited involvement of AtHSFA1b in the regulation of HSR in *A. thaliana.* A duration of 30 minutes of heat stress at 37°C was enough to reduce the number of binding sites of AtHSFA1b in the *A. thaliana* genome by 90.6% (Section 4.2.2). This surprising result showed that AtHSFA1b has probably very little involvement in the regulation of prolonged HSR in *A. thaliana*. This output is consistent with the study that suggested that the involvement of AtHSFA1b in the regulation of HSR is only limited to the immediate and very early phases (Lohmann et al., 2004). That study showed evidence that knockout of both *AthsfA1a/AthsfA1b* impairs plant immediate HSR by delaying the accumulation of mRNA levels of some AtHSPs that are involved in that stage of HSR (Busch et

al., 2005; Lohmann et al., 2004). The outputs of aforementioned and this study show important clues to why plants possess more HSFs than other organisms and also shows indications that HSR and stress response involving HSFs in plants is a highly complex and sophisticated process compared to other species.

The output of the ChIP-SEQ and RNA-SEQ experiments suggest that there might be a rearrangement in the AtHSF regulatory network which takes place in response to changes in growth conditions where stress-inducible AtHSFs become the dominant regulators of HSR in *A. thaliana* under prolonged heat stress conditions. This assumption is supported by the maintained or, in some instances, the increased transcript levels of *AtHSPs* and other stress-associated genes under heat stress from the RNA-SEQ experiment despite the disengagemet of AtHSFA1b from the promoters of stress-response genes (Section 5.2.5).

7.2 Possible regulatory mechanism(s) acting on AtHSFA1b

7.2.1 Possible intrinsic response of AtHSFA1b to elevated temperature:

The heat stress experiment showed evidence that AtHSFA1b is unable to regulate HSR in yeast and also showed that elevated temperatures inhibit the activity of AtHSFA1b (Sections 6.2.4 and 6.2.5). This result is consistent and adds extra validation to the output of ChIP-SEQ experiment under heat stress (Chapter 4). Therefore, it can be predicted that AtHSFA1b also disengages from its target genes in yeast genome under heat stress which results in a collapse in the HSR in yeast cells expressing AtHSFA1b.

It is not clear from a mechanistic point of view what exactly happens that leads the disassociation of AtHSFA1b from its target genes under heat stress in both *A. thaliana* and yeast. A number of studies showed that HSFs from other organisms such as dmHSF and hHSF1 have the ability to intrinsically convert into active trimers in response to elevated temperatures and hydrogen peroxide *in vitro* (Ahn and Thiele, 2003; Zhong et al., 1998). However, that intrinsic sensing of temperature and hydrogen peroxide has a positive effect in the case of dmHSF and hHSF1. There is a possibility that AtHSFA1b might have the same ability to intrinsically respond to elevated temperature and reducing chemicals such as hydrogen peroxide. However, in the case of AtHSFA1b it could be in an opposite manner to those in dmHSF and hHSF1. If AtHSFA1b is able to intrinsically sense temperature then, based on what has been observed from the experiments carried out throughout this study, it would be part of a negative regulatory mechanism that leads to dissociation of the active trimers of AtHSFA1b rather than conversion to an active form (figure 7.1)

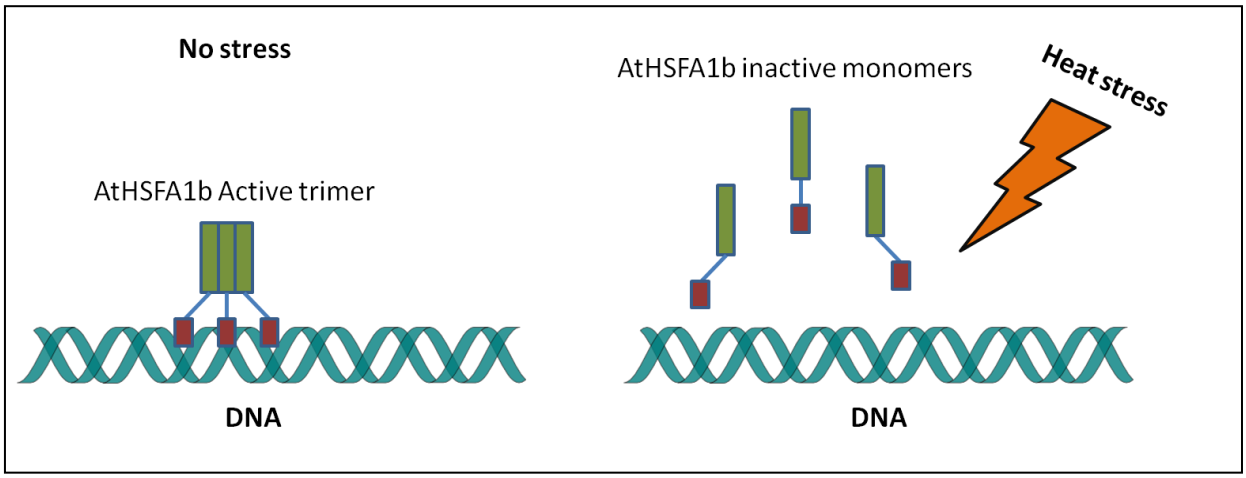


Fig.7.1. AtHSFA1b might intrinsically sense high temperatures. Model of the possible intrinsic temperature sensing of AtHSFA1b. The model shows the possible intrinsic dissociation and conversion of AtHSFA1b to inactive monomers in response to high temperature. The model suggests a negative regulation of AtHSFA1b by increased temperatures.

7.2.2 Possible posttranslational modification events on AtHSFA1b:

7.2.2.1 Redox regulation:

It was briefly mentioned in sections (1.1.2.1, 1.1.2.2 and 1.1.2.3) that *in vitro* experiments showed that yHSF, dmHSF, and hHSF1 respond to reducing agent and can form active trimers in response to superoxide (in yeast) and hydrogen peroxide (in dmHSF and hHSF1) treatments. Studies have shown that the redox state of HSFs play a central role in the multimerisation and subsequently activation of HSFs in a number of species. It has been shown by Ahn and Thiele, (2003) that hHSF1 is subject to redox regulation which leads to the activation of hHSF1 by allowing it to trimerise and, subsequently, translocate to the nucleus. It was shown that two cysteine residues located in the DBD of hHSF1 are responsible for stress sensing (temperature and hydrogen peroxide) and engage in redox regulation via the formation of disulfide bonds. Mutation of those two Cys^{35} or Cys^{105} to serine (C35S and C105S) in hHSF1 impaired its ability to trimerise, translocate in the nucleus and to transactivate the expression of hHSP genes *in vivo*. The study provided very convincing evidence of the importance of both Cys^{35} and Cys^{105} residues in the redox regulation of hHSF1. The reason is that when they did the mutation analyses they mutated Cys^{35} or Cys^{105} but not both of them at the same time and both mutations showed identical effect of the function of hHSF1.

A similar study has shown that also plant HSFs are subject to redox regulation. Jung et al., (2013), have shown that AtHSFA1d is also subject to redox regulation in response to excess light stress. The study showed that the amino acid residues Cys^{153} and Cys^{357} are involved in the regulation of redox state and subsequent activation of AtHSFA1d via formation of disulfide bridges under excess light stress. However, there are few points that need commenting on from that study. First, the study that showed the involvement of Cys^{35} and Cys^{105} in the regulation of

redox state of hHSF1 showed that both residues are located at the same functional domain in hHSF1 which is the DBD. In contrast, the two cysteine residues identified by Jung, et al., (2013) are located on two different domains, Cys^{153} is located on the HR-A/B domain and Cys^{357} is located on a non-functional domain. Therefore, the functional relation between those two cysteine residues is not clear. Second, Ahn and Thiele, (2003) showed a clear evidence of the importance of both cysteine residues in the regulation of the redox state of hHSF1 by mutating a single cysteine residue at a time. Whereas, in Jung, et al., (2013) they mutated both cysteine residues at once and reported the effect. Therefore, it is not clear from Jung, et al., (2013) whether one or both Cys^{153} and Cys^{357} are involved in the regulation of the redox state of AtHSFA1d under light stress. However, it is interesting to note that cysteine residue in the HR-A/B domain is highly conserved among all group-A1 AtHSFs and it is also present in AtHSFA2 (figure 7.2).

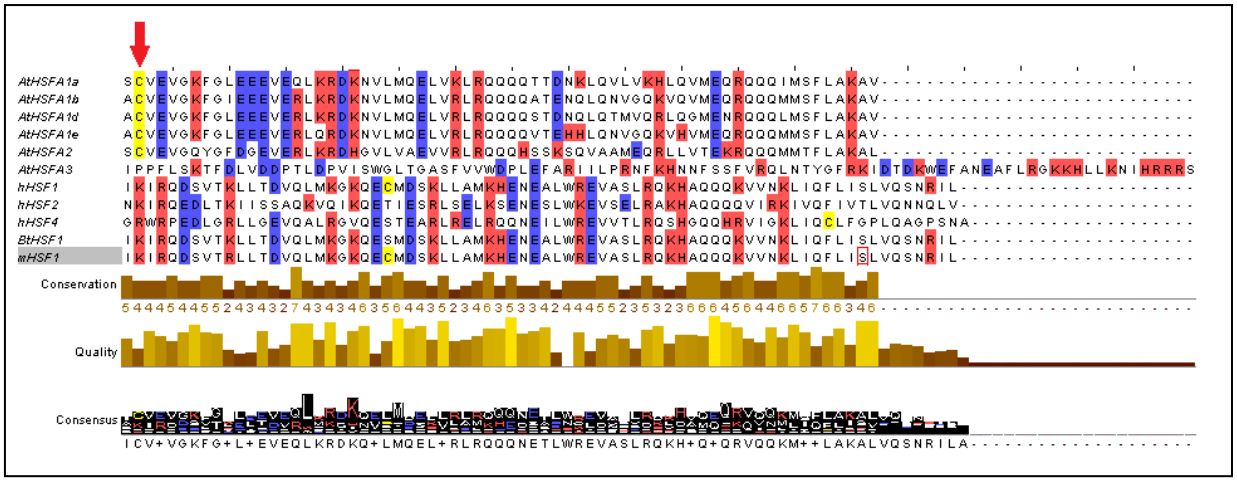


Fig.7.2. Conservation of the cysteine residue located with the HR-A/B domain among all group-A1 AtHSFs and AtHSFA2. Multiple sequence alignment of the HR-A/B domain from *A. thaliana* and mammalian HSFs. The conserved cysteine residue in the HR-A/B domains of group-A1 AtHSFs and AtHSFA2 is marked with the red arrow.

The availability and the constitutive expression pattern of *AtHSFA1b* might make it exposed to the ROS bursts that occur during abiotic stress. Furthermore, the evidence shown from previous studies that AtHSFA1b can induce the expression of AtHSPs and other stress-related genes (Busch et al., 2005; Lohmann et al., 2004) gives a further indication that AtHSFA1b might be

subject to redox regulation during stress that would allow it to bind and transiently activate the expression of AtHSPs and other stress-associated genes in the early phases of stress. Therefore, redox regulation of AtHSFA1b is an area that is worth investigating to reveal some of the regulatory mechanisms acting on AtHSFA1b.

7.2.2.2 Phosphorylation:

In vivo studies of HSFs revealed that they are subject to numerous posttranslational modifications in response to changes in the overall physiological state of the cell. There are many examples of well-studied posttranslational event on HSFs from other organisms. For example, it has been show that phosphorylation is a major posttranslational modification event on dmHSF under heat stress. Groups of serine residues on dmHSF are phosphorylated and other groups are dephosphorylated under heat stress. However, it has been shown that disruption of those phosphorylation events does not impair dmHSF from binding the DNA. As a result it was concluded that those phosphorylation events have no significant role in regulation of the DNA binding activity of dmHSF (Fritsch and Wu, 1999). Similar result was observed on hHSF1 in heat stressed cells. It was shown that hHSF1 is phosphorylated on multiple serine residues upon activation under heat stress including Ser^{121}, Ser^{230}, Ser^{292}, Ser^{303}, Ser^{307}, Ser^{314}, Ser^{319}, Ser^{326}, Ser^{344}, Ser^{363}, Ser^{419}, and Ser^{444}. However, none of those phosphorylation events was shown to play a significant role in the activation of hHSF1. Furthermore, none of those events showed any effect on the ability of hHSF to bind the DNA and translocate in the nucleus in *hhsf1* cells except that the replacement of Ser^{326} with Ala (S326A) stimulated the expression of hHSP70 several folds less than wild type hHSF (Guettouche et al., 2005).

Despite showing limited effect of phosphorylation of HSFs on their molecular functions, some other studies showed that phosphorylation of specific serine residues can play a key regulatory role on HSFs. For instance, it was shown that phophorylation of Ser^{230} enables hHSF1 to become active and bind the DNA. The study showed that Ser^{230} is phosphorylated via the enzyme calcium/calmodulin dependant protein kinase (CaMKII). Overexpression of that enzyme resulted in a significant increase in the amount of active hHSF1 *in vivo*. Moreover, the same study

showed that mutation of Ser^{230} to alanine (S230A) resulted in markedly reduced activity of hHSF1 when expressed in *hhsf1* cells compared to wild type hHSF1 (Holmberg et al., 2001).

Looking at the RNA-SEQ data from heat stressed wild type and 35S-AtHSFA1b::mRFP plants, there was a number of upregulated genes that code for protein kinases and mitogen activated protein kinases (MAPK). This indicates that phosphorylation is perhaps a major event during heat stress. However, it is not clear if AtHSFA1b is part of that phosphorylation event during heat stress or not. Furthermore, even if there is a change in the phosphorylation state of AtHSFA1b, it is difficult to determine, which of the upregulated genes coding protein/MAP kinases is the one that is involved in phosphorylation of AtHSFA1b under heat stress conditions without carrying out an appropriate experiment (table 7.1)

AGI code	Gene name	Description
AT1G67000	--	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation
AT1G72540	--	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation
AT3G57730	--	Protein kinase superfamily protein; FUNCTIONS IN: protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation
AT5G37350	--	Serine/threonine-protein kinase Rio1; FUNCTIONS IN: protein serine/threonine kinase activity, catalytic activity, ATP binding
AT3G46930	--	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine/tyrosine kinase activity, kinase activity; INVOLVED IN: protein amino acid phosphorylation
AT5G25440	--	Protein kinase superfamily protein; FUNCTIONS IN: protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, N-terminal protein myristoylation
AT1G56145	--	Leucine-rich repeat transmembrane protein kinase; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, ATP binding
AT1G48210	--	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation
AT5G45780	--	Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding
AT1G61640	--	Protein kinase superfamily protein
AT5G50780	--	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase family protein; FUNCTIONS IN: ATP binding
AT5G25930	--	Protein kinase family protein with leucine-rich repeat domain; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation
AT1G33770	--	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding
AT5G66210	*AtCPK28*	member of Calcium Dependent Protein Kinase
AT3G57530	*AtCPK32*	Calcium-dependent Protein Kinase. ABA signaling component that regulates the ABA-responsive gene expression via ABF4. AtCPK32 has autophosphorylation activity and can phosphorylate ABF4 in vitro
AT4G18700	*AtCIPK12*	Encodes CBL-interacting protein kinase 12 (CIPK12).
AT2G25090	*AtCIPK16*	Encodes a member of the SNF1-related kinase (SnRK) gene family (SnRK3.18), which has also been reported as a member of the CBL-interacting protein kinases (CIPK16).
AT3G55950	*AtCCR3*	CRINKLY4 related 3 (CCR3); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation
AT4G23190	*AtCRK11*	Encodes putative receptor-like protein kinase that is induced by the soil-borne vascular bacteria, Ralstonia solanacearum. Naming convention from Chen et al 2003 (PMID 14756307)
AT4G04540	*AtCRK39*	Encodes a cysteine-rich receptor-like protein kinase.
AT2G18170	*AtMPK7*	MAP kinase 7 (MPK7)
AT3G50500	*AtSNRK2.2*	encodes a member of SNF1-related protein kinases (SnRK2) whose activity is activated by ionic (salt) and non-ionic (mannitol) osmotic stress. Enzyme involved in the ABA signaling during seed germination, dormancy and seedling growth.
AT3G48530	*AtKING1*	SNF1-related protein kinase regulatory subunit gamma 1 (KING1)

Table 7.1. Upregulated genes that code for proteins involved in kinase activity in heat stressed wild type and 35S-AtHSFA1b::mRFP. Annotations were obtained from the *A. thaliana* database (TAIR; http://www.arabidopsis.org).

7.2.2.3 Acetylation:

Acetylation is also a major posttranslational modification event that has been shown to have a dramatic effect on the function of hHSF1. At least 9 lysine residues have been shown to be acetylated under stress. However, acetylation of the Lys^{80} has the greatest impact on the function of hHSF1. It was shown that acetylation of Lys^{80} inhibits hHSF1 from binding to HSE *in vitro*; despite showing that acetylation of that specific residue does not affect trimerisation of hHSF1 upon heat stress. Replacement of the Lys^{80} with Glutamine (K80Q) to mimic constitutive acetylation of Lys^{80} led to complete inhibition of hHSF1 binding activity to the DNA and activation of expression of hHSP70 *in vivo* resulting in a non-functional hHSF1 (Westerheide et al., 2009). The study also identified the protein SIRT1 responsible for deacetylation of hHSF1 upon activation under stress.

Scanning the DBD sequence on AtHSFA1b showed that this lysine residue at that particular position is among at least three conserved between AtHSFA1b and many other HSFs from other species suggesting that acetylation of that lysine residue on the DBD of HSFs could be a highly conserved regulatory process (figure 7.3)

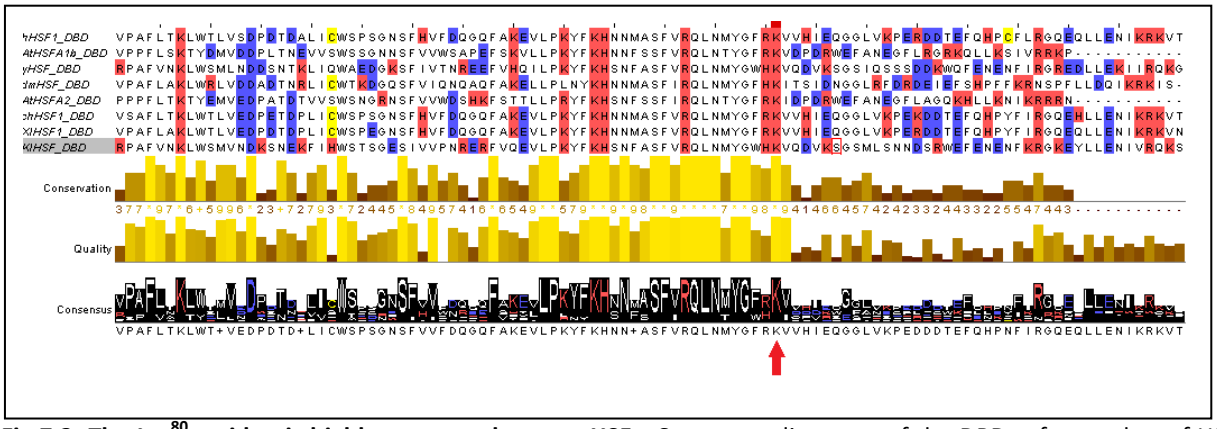


Fig.7.3. The Lys^{80} residue is highly conserved among HSFs. Sequence alignment of the DBDs of a number of HSFs from different species. The conserved Lys^{80} is marked with the red arrow.

Looking back again at the RNA-SEQ data from heat stressed wild type and 35S-AtHSFA1b::mRFP plants, there was not a single gene that codes for a protein involved in acetylation activity among the upregulated genes under heat stress. This gives an indication that acetylation of that lysine residue in AtHSFA1b might not be occurring under heat stress conditions. However, this assumption needs to be validated with appropriate experiments. From what has been discussed about the possible posttranslational events that might be occurring on AtHSFA1b during heat stress, it is clear that this is a point that needs addressing in future work.

7.2.4 Possible protein-protein interactions:

It was reported that interaction between hHSF1 and hHSP90 occurs under no stress conditions which is responsible for maintaining hHSF1 in an inactive monomeric state. Under stress conditions, the hHSP90-hHSF1 complex dissociates allowing hHSF1 monomers to interact with each other and subsequently forming an active trimer that can translocate in the nucleus and activate the expression of stress genes (Wang et al., 2005; Zou et al., 1998). Similarly, hHSP70 was shown to interact with hHSF1 which results in repressing hHSF1 from binding the DNA (Abravaya et al., 1992). In *D. melanogaster*, however, a study reported interaction between dmHSF and DROJ1 protein which is the counterpart of hHSP40. The study again showed that the interaction is of a repressive nature. The study showed that overexpression of DROJ1 in *D. melanogaster* SL resulted in delaying HSR. In contrast, depletion of DROJ1 resulted in constitutive activation of endogenous heat shock genes (Marchler and Wu, 2001).

Unlike HSFs from other species, the area of protein-protein interaction between AtHSFs and other proteins is poorly studied. There are only few examples from the literature of interacting proteins with AtHSFs. For instance, an interaction between AtHSFA1a and AtHSP70 has been reported *in vitro* and *in vivo*. The study showed that AtHSP70 interacts with the TAD of AtHSFA1a which results in repressing the transcriptional activity of AtHSFA1a. However, the interaction was shown using EMSA and Y2H assay and has never been shown to occur in *planta* (Kim and Schöffl, 2001). Interaction between group-A1 AtHSFs was also shown to occur. It was shown that AtHSFA1a and AtHSFA1b can interact with each other and form active heterotrimers in *A. thaliana* protoplasts (Kim and Schöffl, 2001).

The interaction between AtHSFA1a and AtHSFA1b was also validated using immunoprecipitation. Study also showed that both AtHSFs interact with each other via their HR-A/B domains as deletion of those domains from both AtHSFs abolished their ability to interact with each other. Furthermore, the same study showed that this is not the case with class-B AtHSFs as AtHSFB1 and AtHSFB2b did not show any signs of interaction or formation of heterotrimers in the BiFC assay (Li et al., 2010). *In vitro* and Y2H but not in *planta* binding assays also showed positive interaction between AtHSFA1a and TATA-binding protein (TBP) (Reindl and Schöffl, 1998). A more recent study have reported the interaction between AtHSFA4a and two MAPKs (MAPK3 and MAPK6) which leads to change of the phosphorylation state of AtHSFA4a in response to stress (Pérez-Salamó et al., 2014). However, this is related to posttranslational modifications more than formation of inhibitory complexes.

The HSP families in plants are more complex than those in mammals and *D. melanogaster*. There is an overwhelming number of different isoforms of both HSP70 and HSP90. For instance, there are 7 members in the AtHSP90 family and at least 18 members in the AtHSP70 family (Krishna and Gloor, 2001; Lin et al., 2001). This multiplicity of the AtHSP70/90 proteins in *A. thaliana* makes it difficult to predict which one from each family is the interacting partner of AtHSFA1b. However, the regulation of AtHSFA1b by HSPs, particularly AtHSP70/90, under heat stress remains a strong possibility that is worth investigating in order to understand the regulatory mechanism occurring on AtHSFA1b during heat stress that leads to its dissociation for its target region in the *A. thaliana* under heat stress.

7.3 Analysis of existing models of transcriptional regulation of HSR by AtHSFs

Looking back at previous published work, only a few models were suggested for the regulation of HSR by group-A1 AtHSFs including AtHSFA1b. However, the models generated seem to be either flawed or vague. The first model generated was by Nishizawa-Yokoi et al., (2011) where it was suggested that AtHSFA1a and AtHSFA1b control the expression of a few *AtHSFs* in response to heat stress. Moreover, the study suggested that AtHSFA1d and AtHSFA1e control the expression of two main heat inducible AtHSFs (*AtHSFA2* and *AtHSFA3*). The second suggestion seems to be valid based on their analysis of an *AthsfA1d/AthsfA1e* double knockout mutant. However, they suggested that *AtHSFA2* controls the expression of *AtHSFA3* which by turn controls the expression of *AtHSFA1e* in a feedback loop manner (figure 7.4)

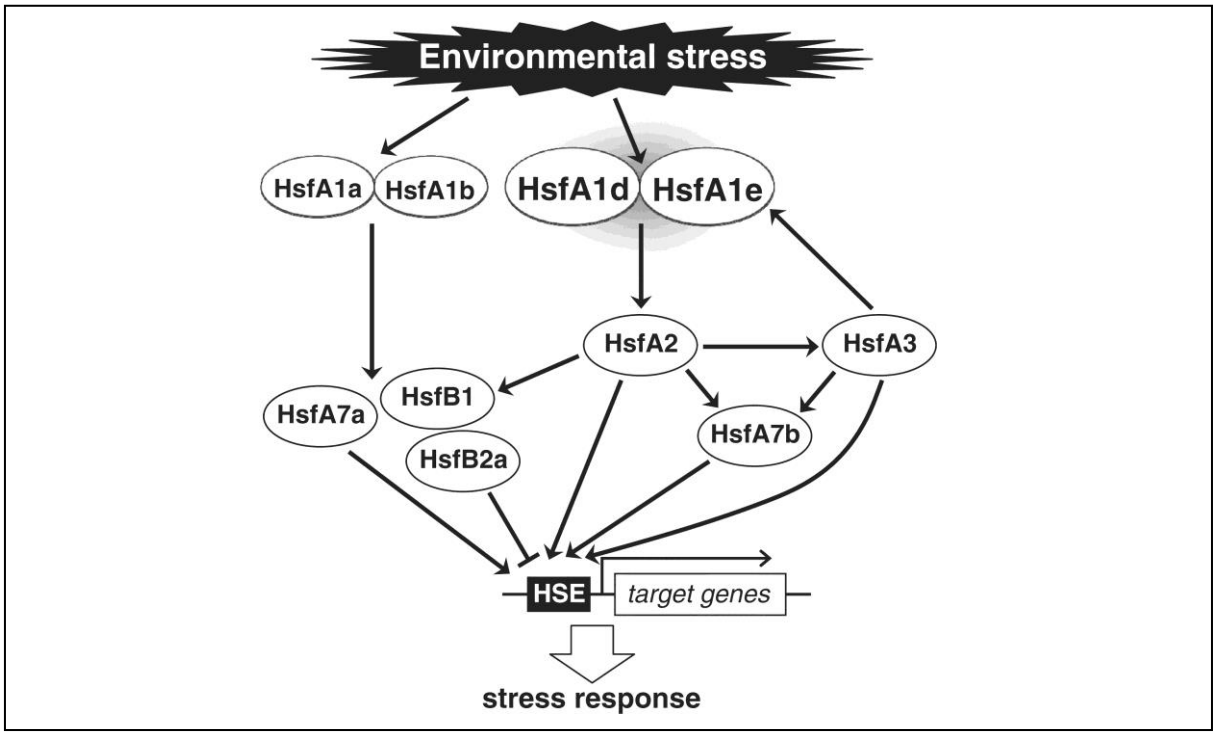


Fig.7.4. Suggested model for the AtHSFs signalling pathways via other AtHSFs in response to environmental stress. The model was generated by Nishizawa-Yokoi et al., (2011)

However, essential parts in this model seem to be flawed especially the AtHSFA1e part of the model. The authors used microarray analysis from plants overexpressing *AtHSFA3* which showed that *AtHSFA1e* was upregulated by overexpression of *AtHSFA3* (Yoshida et al., 2008).

The same microarray data showed also upregulation of other *AtHSFs* including *AtHSFA7b*. Clearly the authors did not consider, for some reason, that AtHSFA7b might be the one that affects the expression of *AtHSFA1e* which is not shown to be linked to AtHSFA1e in their model. Furthermore, the model generated by Nishizawa-Yokoi et al., (2011) suggests that *AtHSFA3* is controlled by AtHSFA2 and they cited two published works (Ogawa, et al., 2007; Schramm, et al., 2006). However, none of the studies they cited mentioned any information about *AtHSFA3* being controlled by AtHSFA2. Moreover, looking at the microarray data from plants overexpressing *AtHSFA2*, the list of genes upregulated by overexpression of *AtHSFA2* does not include *AtHSFA3* (Bechtold et al., 2013; Ogawa et al., 2007). It is still not clear how some of the links in the model were generated.

The second suggested model for the regulation of HSR by group-A1 AtHSFs was by Yoshida et al., (2011). The model suggests that group-A1 AtHSFs directly control the expression of heat inducible *AtHSFs* and *AtHSPs* under heat stress. Furthermore, they showed what seems to be an evidence of the ability of AtHSFA1b to induce the expression of *AtDREB2A* under heat stress based on ChIP-PCR experiment. They also showed some evidence that the expression of *AtHSFA3* was reduced in the *AthsfA1a/AthsfA1b/AthsfA1d* triple knockout mutant, which suggests that the expression of *AtHSFA3* might be controlled by one of those members of group-A1 AtHSFs. The final model suggested by the study is shown in figure 7.5.

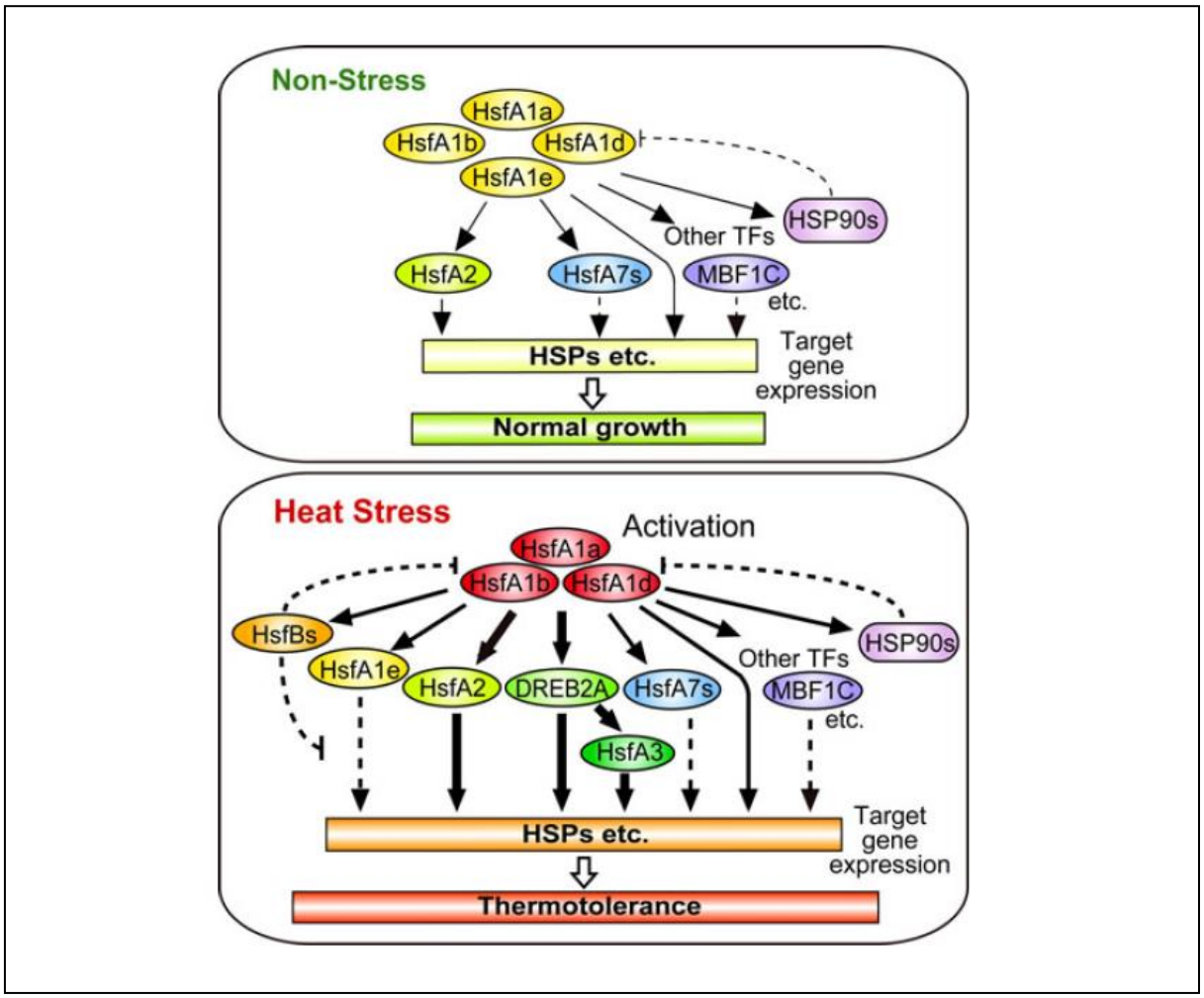


Fig.7.5. Model suggested for the transcriptional regulation cascade of HSR by group-A1 AtHSFs. This model was suggested by Yoshida et al., (2011).

There are, however, few issues in that study that weakens the credibility of the suggested model for the transcriptional cascade by AtHSFs during heat stress. First, the ChIP-PCR experiment they carried out and showed that AtHSFA1b binds to the promoters of *AtDREB2A*, *AtHSFA2*, and *AtHSP18.2* lacks a negative control which is an absolutely crucial part for the success of any ChIP-PCR/ChIP-SEQ experiment. It is not possible to tell whether the PCR signal generated from ChIP-PCR experiment is a real result or an artefact from PCR or non-specific antibody interaction with other proteins without including a negative control in the experiment (Massie and Mills, 2008; Buck and Lieb, 2003). Furthermore, the ChIP-SEQ experiment carried out in this research showed that AtHSFA1b binds to the promoter of *AtDREB2A* only under no stress conditions then

loses the binding under heat stress (See appendix B).

The study also showed that AtHSFA1b binds to the promoter of *AtHSFA2* under heat stress suggesting that the expression of *AtHSFA2* is controlled by AtHSFA1b under heat stress. This assumption was again based on the ChIP-PCR result and the loss of expression of *AtHSFA2* under heat stress in *AthsfA1a/AthsfA1b/AthsfA1d* triple knockout mutant. However, there is no evidence yet that this is the case. First, ChIP-SEQ data generated in this research showed that AtHSFA1b does not bind to the promoter of *AtHSFA2* under both no stress and heat stress conditions. Second, the RNA-SEQ did not show any increase in the transcript levels of *AtHSFA2* in 35S-AtHSFA1b::mRFP plants under no stress. The induction of *AtHSFA2* under heat stress in 35S-AtHSFA1b::mRFP plants could be by any other AtHSF but not AtHSFA1b. Furthermore, Nishizawa-Yokoi et al., (2011) showed strong evidence that the expression of *AtHSFA2* is controlled by either AtHSFA1d and/or AtHSFA1e. The model also did not include results from previous published work by the same group (Yoshida et al., 2008) where they showed evidence from microarray study that the overexpression of *AtHSFA3* leads to upregulation of *AtHSFA1e*. Finally, this model suggests an indirect regulation of *AtMBF1c* by AtHSFs which is a conclusion that lacks accuracy. The results obtained from ChIP-PCR, ChIP-SEQ and RNA-SEQ show strong evidence that *AtMBF1c* is directly controlled by AtHSFA1b under no stress conditions and possibly by other AtHSFs under heat stress (Chapter 3, Appendix B and C; Bechtold et al., 2013).

7.4 The AtHSFA1b transcriptional regulatory network:

Perhaps the most interesting finding in this study is the association between the loss of AtHSFA1b binding events and the general downregulation of the expression of developmental genes under heat stress (Section 5.2.5). It is not possible to conclude that the downregulation of AtHSFA1b target genes under heat stress is solely due to the loss of AtHSFA1b binding. However, this association between the loss of binding events and the downregulation of the expression of developmental genes under heat stress that were bound by AtHSFA1b under no stress strongly suggests that AtHSFA1b might be part of the regulation of plant development more than regulation of HSR. Based on all of the information obtained from the experiments carried out throughout this study and using published transcriptomics data of TFs controlled by AtHSFA1b, two models of the AtHSFA1b transcriptional regulatory network were generated. An intact binary model which shows the static architecture of the AtHSFA1b network under no stress conditions and a collapsed model which shows the diminished role of AtHSFA1b transcriptional network under heat stress. The intact model was divided into two parts; the first part is an AtHSFA1b local network which shows only TFs that are directly controlled by AtHSFA1b. The second part shows and extended network which shows the links between AtHSFA1b and its indirect genes (Figures 7.6, 7.7 and 7.8).

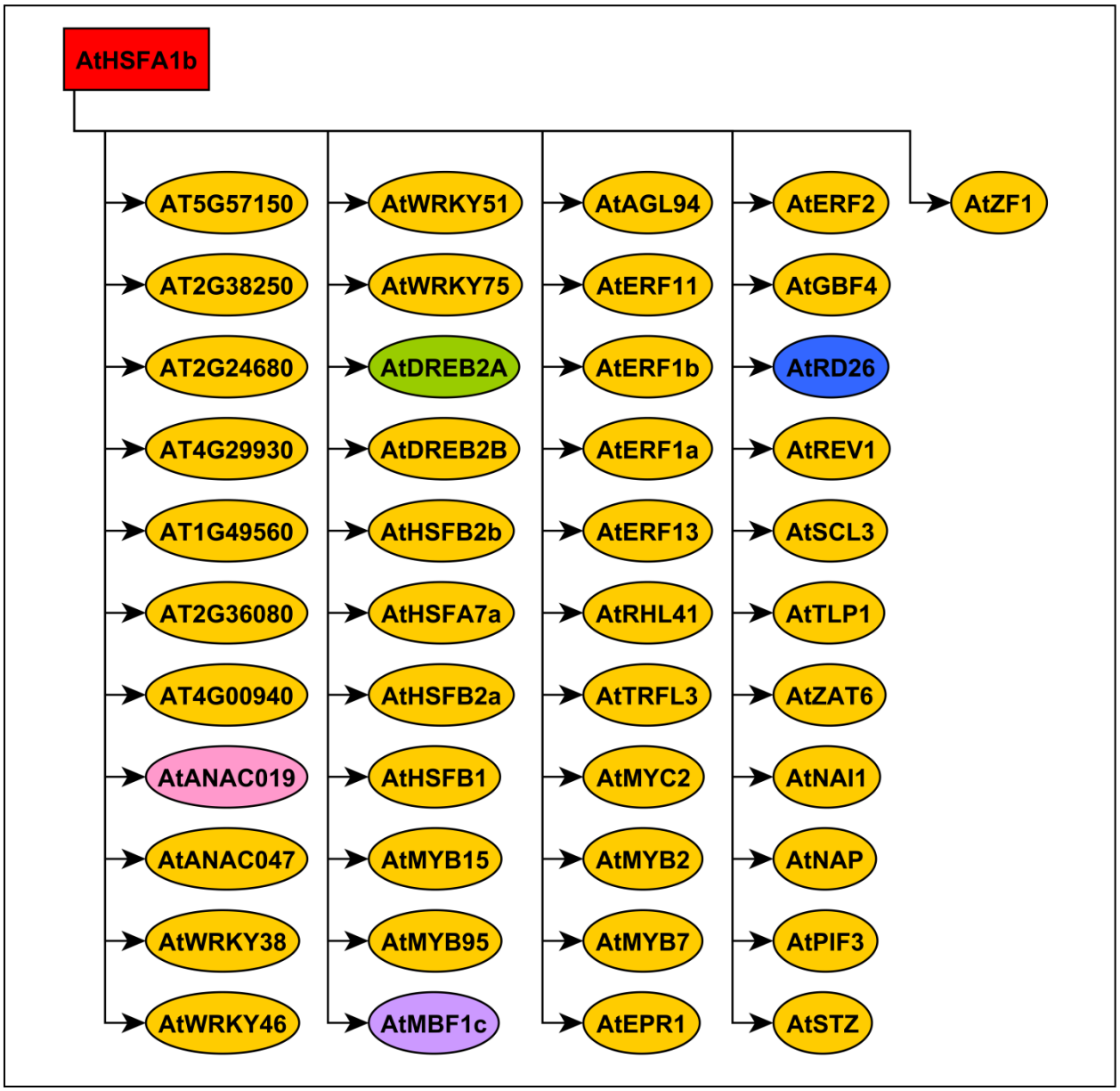


Fig.7.6. A model of the AtHSFA1b local transcriptional network. The diagram shows illustration of the AtHSFA1b local network consisting of all the identified TFs that were bound by AtHSFA1b from ChIP-SEQ and upregulated in 35S-AtHSFA1b::mRFP under no stress conditions. Unique colours indicate TFs with known transcriptomics data which are used to construct an extended network of AtHSFA1b.

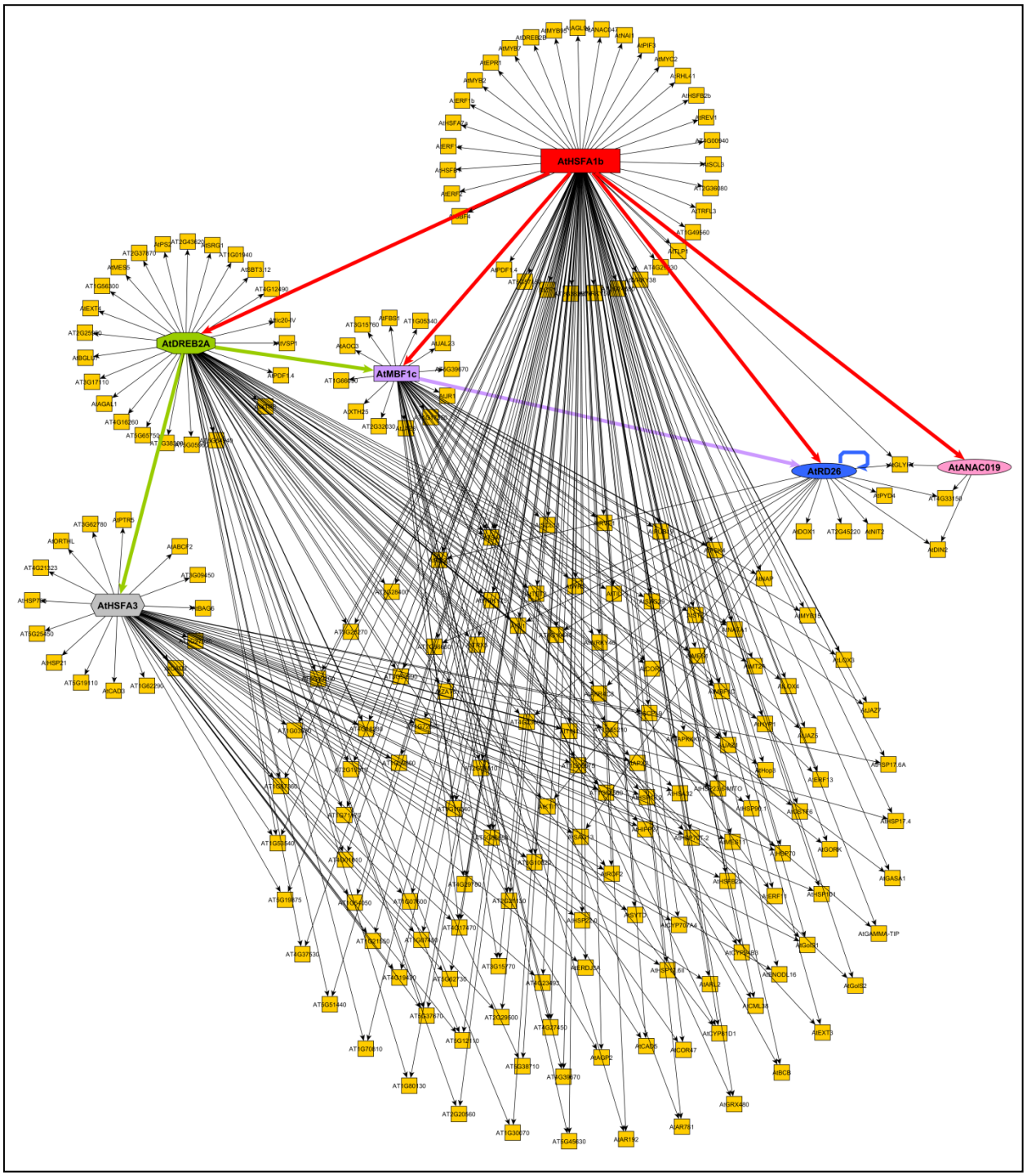


Fig.7.7. A model of an extended AtHSFA1b transcriptional network. The network consist of AtHSFA1b direct TF targets and indirect genes linked to AtHSFA1b via other TFs directly controlled by AtHSFA1b in 35S-AtHSFA1b::mRFP plants under non-stress conditions.

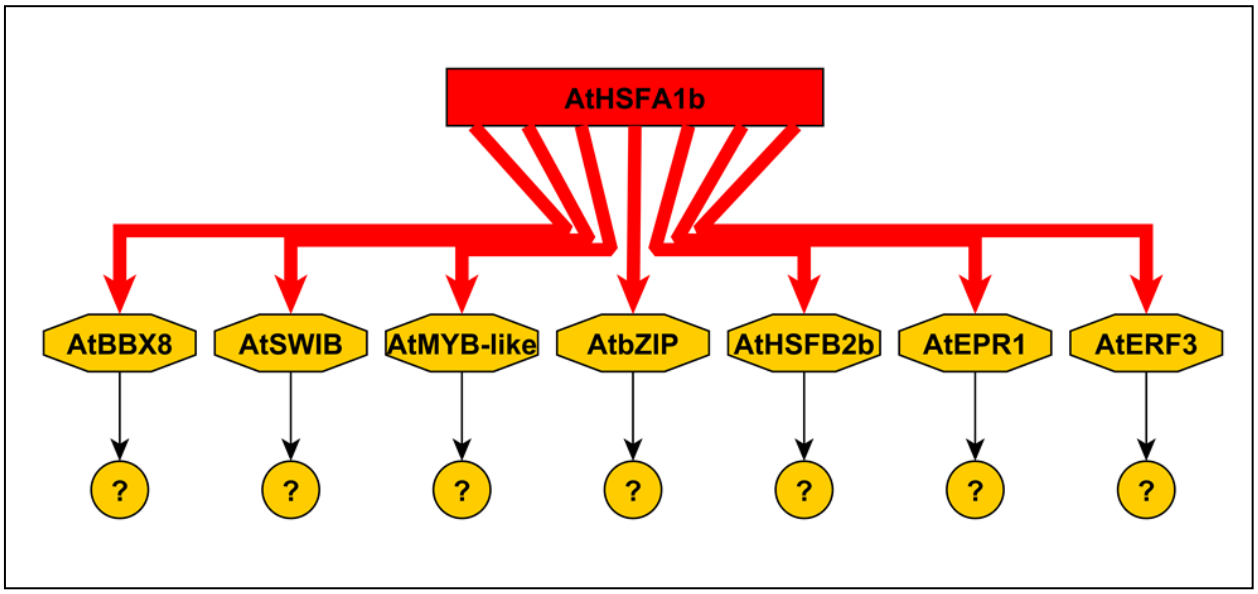


Fig.7.8. A model of the AtHSFA1b collapsed transcriptional network. The model shows the TFs directly controlled by AtHSFA1b in 35S-AtHSFA1b::mRFP plants under heat stress.

The network models generated clearly show that the architecture of the AtHSFA1b network is far more complex under non-stress conditions than under heat stress. This shows evidence of the possible limited involvement of AtHSFA1b in the regulation of HSR in *A. thaliana* and also shows that AtHSFA1b might be involved in the regulation of various cellular processes under non-stress conditions. However, it is important to state that the network models generated are based mainly on the experimental conditions used in this study and the fate of AtHSFA1b is still not known under heats stress durations less and more than 30 minutes.

The outcomes of heat stress experiments were based on a relatively short duration of heat stress and it is not known whether AtHSFA1b re-engages and binds again under prolonged heat stress or not. Therefore, a further improvement is to examine whether there is an involvement of AtHSFA1b in the regulation of HSR under prolonged heat stress conditions by carrying out ChIP-SEQ experiment under longer durations of heat stress. It is also important to state that the AtHSFA1b transcriptional regulatory network models generated in this study are static and do not show the dynamics of the AtHSFA1b transcriptional regulatory network. The next step is to

study the dynamics of the AtHSFA1b network and build a two dimensional network model using space and time. However, the constitutive expression nature of AtHSFA1b may not allow for the construction of such network as it would make it not possible to track early and late inducible target genes of AtHSFA1b. Therefore, engineering an AtHSFA1b construct under control of an inducible promoter would be ideal to build a dynamic network. This can be achieved by using chemically inducible promoters. Examples of such promoters include the dexamethasone inducible promoter and the estradiol promoter which allow the expression of genes controlled by those promoters to be induced by applying those chemicals to the plant. Studying the dynamics of the AtHSFA1b network would be a move to the next level to analyse the AtHSFA1b network in depth which might provide more valuable information of the AtHSFA1b involvement and the possible changes in the topology of its network that might take place in reponse to difference growth conditions.

References:

Abravaya, K., Myers, M. P., Murphy, S. P., and Morimoto, R. I. (1992). The human heat shock protein hsp70 interacts with HSF, the transcription factor that regulates heat shock gene expression. Genes Dev., 6(7): 1153-1164.

Ahn, S. G., Liu, P. C., Klyachko, K., Morimoto, R. I., and Thiele, D. J. (2001). The loop domain of heat shock transcription factor 1 dictates DNA-binding specificity and responses to heat stress. Genes & dev., 15(16), 2134-2145.

Åkerfelt, M., Morimoto, R. I., and Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. Nat. Rev. Mol. cell biol., 11(8), 545-555.

Alberti, S., Gitler, A. D. and Lindquist, S. (2007) A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast 24(10): 913-919.

Apel, K. and Hirt, H. (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol., 55: 373-399.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., and Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: Transcription regulation and beyond. FEMS microbial. Rev., 29(2), 231-262.

Babu, M. M. (2008). Evolutionary and Temporal Dynamics of Transcriptional Regulatory Networks. In Bio-Inspired Computing and Communication (pp. 174-183). Springer Berlin Heidelberg.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res., 37(Web Server issue): W202-208.

Bardou, P., Mariette, J., Escudié, F., Djemiel, C., and Klopp, C. (2014). jvenn: an interactive Venn diagram viewer. BMC bioinformatics, 15(1): 293.

Batista, R., Saibo, N., Lourenco, T., and Oliveira, M. M., (2008). Microarray analyses reveal that plant mutagenesis may induce more transcriptomic changes than transgene insertion. Proc. Natl. Acad. Sci. 105(9): 3640-3645.

Batista-Nascimento, L., Neef, D. W., Liu, P. C., Rodrigues-Pousada, C., and Thiele, D. J. (2011). Deciphering human heat shock transcription factor 1 regulation via post-translational modification in yeast. PLoS One 6(1): e15976.

Baxter, A., Mittler, R., and Suzuki, N. (2014) ROS as key players in plant stress signalling. J. Exp. Bot. 65(5): 1229-1240.

Bechtold, U., Albihlal, W. S., Lawson, T., Fryer, M. J., Sparrow, P. A., Richard, F., Persad, R., Bowden, L. Hickman, R., Martin, C., Beynon, J. L., Buchanan-Wollaston, V., Baker, N. R., Morison, J. I., Schoffl, F., Ott, S., and Mullineaux, P. M., (2013). Arabidopsis HEAT SHOCK TRANSCRIPTION FACTORA1b overexpression enhances water productivity, resistance to drought, and infection. J. Exp. Bot., 64(11): 3467-3481.

Begum, T., Reuter, R., and Schoffl, F. (2013) Overexpression of AtHsfB4 induces specific effects on root development of Arabidopsis. Mech. Dev., 130(1): 54-60.

Bhardwaj, N., Kim, P. M., and Gerstein, M. B. (2010). Rewiring of transcriptional regulatory networks: hierarchy, rather than connectivity, better reflects the importance of regulators. Sci. Signal., 3(146): 79.

Bonner, J. J., Carlson, T., Fackenthal, D. L., Paddock, D., Storey, K., and Lea, K. (2000). Complex regulation of the yeast heat shock transcription factor. Mol. Biol. Cell., 11(5): 1739-1751.

Bookout, A. L., Jeong, Y., Downes, M., Yu, R. T., Evans, R. M., and Mangelsdorf, D. J. (2006). Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. Cell 126(4): 789-799.

Boscheinen, O., Lyck, R., Queitsch, C., Treuter, E., Zimarino, V., and Scharf, K. D. (1997). Heat stress transcription factors from tomato can functionally replace HSF1 in the yeast Saccharomyces cerevisiaet. MGG., 255(3): 322-331.

Boulikas, T. (1994). Putative nuclear localization signals (NLS) in protein transcription factors. J. cell. biochem., 55(1): 32-58.

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M., and Smyth, D. R. (1993). Control of flower development in Arabidopsis thaliana by APETALA 1 and interacting genes. Development., 119: 721-721.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., ... and Kovalchuk, I. (2010). Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. PLoS one,5(3), e9514.

Bray, E. A. (2004) Genes commonly regulated by water-deficit stress in Arabidopsis thaliana. J. Exp. Bot., 55(407): 2331-2341.

Brenner, W. G., Romanov, G. A., Kollmer, I., Burkle, L. and Schmulling, T. (2005) Immediate-early and delayed cytokinin response genes of Arabidopsis thaliana identified by genome-wide expression profiling reveal novel cytokinin-sensitive processes and suggest cytokinin action through transcriptional cascades. Plant J. 44(2): 314-333.

Brisson, L. F., Tenhaken, R., and Lamb, C. (1994). Function of oxidative cross-linking of cell wall structural proteins in plant disease resistance. Plant Cell. 6(12): 1703-1712.

Brusslan, J. A., Rus Alvarez-Canterbury, A. M., Nair, N. U., Rice, J. C., Hitchler, M. J. and
Buck, M., and J. D. Lieb (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. Genomics 83(3): 349-360.

Busch, W., Wunderlich, M., and Schoffl, F. (2005) Identification of novel heat shock factor-dependent genes and biochemical pathways in Arabidopsis thaliana. Plant J. 41(1): 1-14.

Capuano, F., Mulleder, M., Kok, R., Blom, H. J., and Ralser, M. (2014) Cytosine DNA methylation is found in Drosophila melanogaster but absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe, and other yeast species. Anal. Chem., 86(8): 3697-3702.

Carper, S. W., Duffy, J. J., and Gerner, E. W. (1987). Heat shock proteins in thermotolerance and other cellular processes. Cancer Research, 47(20), 5249-5255.

Charng, Y. Y., Liu, H. C., Liu, N. Y., Hsu, F. C., and Ko, S. S. (2006) Arabidopsis Hsa32, a novel heat shock protein, is essential for acquired thermotolerance during long recovery after acclimation. Plant Physiol. 140(4): 1297-1305.

Charng, Y. Y., Liu, H. C., Liu, N. Y., Chi, W. T., Wang, C. N., Chang, S. H., and Wang, T. T. (2007) A heat-inducible transcription factor, HsfA2, is required for extension of acquired thermotolerance in Arabidopsis. Plant Physiol. 143(1): 251-262.

Chary, S. N., Hicks, G. R., Choi, Y. G., Carter, D., and Raikhel, N. V. (2008). Trehalose-6-phosphate synthase/phosphatase regulates cell shape and plant architecture in Arabidopsis. Plant Physiol. 146(1): 97-107.

Chauhan, H., Khurana, N., Agarwal, P., and Khurana, P. (2011) Heat shock factors in rice (Oryza sativa L.): genome-wide expression analysis during reproductive development and abiotic stress. Mol. Genet. Genomics, 286(2): 171-187.

Chen, F. E., and Ghosh, G. (1999). Regulation of DNA binding by Rel/NF-kappaB transcription factors: structural views. Oncogene, 18(49), 6845-6852.

Chen, W. (2002) Expression Profile Matrix of Arabidopsis Transcription Factor Genes Suggests Their Putative Functions in Response to Environmental Stresses. Plant Cell 14(3): 559-574.

Chen, W., Provart, N. J., Glazebrook, J., Katagiri, F., Chang, H. S., Eulgem, T., ... and Zhu, T. (2002). Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell, 14(3), 559-574.

Chen, Y. F., Matsubayashi, Y., and Sakagami, Y. (2000). Peptide growth factor phytosulfokine-α contributes to the pollen population effect. Planta, 211(5), 752-755.

Cheong, Y. H., Chang, H. S., Gupta, R., Wang, X., Zhu, T. and Luan, S. (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. Plant Physiol. 129(2): 661-677.

Choi, D. S., Hwang, I. S., and Hwang, B. K. (2012). Requirement of the cytosolic interaction between PATHOGENESIS-RELATED PROTEIN10 and LEUCINE-RICH REPEAT PROTEIN1 for cell death and defense signaling in pepper. Plant Cell, 24(4), 1675-1690.

Choi, D. S., Hwang, I. S., and Hwang, B. K. (2012) Requirement of the cytosolic interaction between PATHOGENESIS-RELATED PROTEIN10 and LEUCINE-RICH REPEAT PROTEIN1 for cell death and defense signaling in pepper. Plant Cell 24(4): 1675-1690.

Collas, P. (2010). The current state of chromatin immunoprecipitation.Mol. biotech., 45(1), 87-100.

Cosgrove, D. J. (2005). Growth of the plant cell wall. Nat. Rev. Mol. cell biol., 6(11), 850-861.

Cubas, P., Lauter, N., Doebley, J., and Coen, E. (1999). The TCP domain: a motif found in proteins regulating plant growth and development. Plant J., 18(2): 215-222.

Cusanovich, D. A., Pavlovic, B., Pritchard, J. K., and Gilad, Y. (2014) The functional consequences of variation in transcription factor binding. PLoS Genet., 10(3): e1004226.

Dallas, P. B., Gottardo, N. G., Firth, M. J., Beesley, A. H., Hoffmann, K., Terry, P. A., Freitas, J. R., Boag, J. M., Cummings, A. J., and Kees, U. R. (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? BMC Genomics 6: 59.

Dalman, M. R., Deeter, A., Nimishakavi, G., and Duan, Z. H. (2012) Fold change and p-value cutoffs significantly alter microarray interpretations. BMC Bioinformatics 13 Suppl 2: S11.

Darnell, J. E., Jr. (2002). Transcription factors as targets for cancer therapy. Nat. Rev. Cancer., 2(10): 740-749.

de Carvalho, M. H. C. (2008). Drought stress and reactive oxygen species.Plant Signal Behav., 3(3), 156-165.

DeGregori, J., Leone, G., Miron, A., Jakoi, L., and Nevins, J. R. (1997). Distinct roles for E2F proteins in cell growth control and apoptosis. Proc. Natl. Acad. Sci., 94(14): 7245-7250.

Delgado, I. J., Wang, Z., de Rocher, A., Keegstra, K., & Raikhel, N. V. (1998). Cloning and Characterization of AtRGP1 A Reversibly Autoglycosylated Arabidopsis Protein Implicated in Cell Wall Biosynthesis. Plant Physiol., 116(4), 1339-1350.

Deplancke, B., Dupuy, D., Vidal, M., & Walhout, A. J. (2004). A gateway-compatible yeast one-hybrid system. Genome res., 14(10b), 2093-2101.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., ... & Gascuel, O. (2008). Phylogeny. fr: robust phylogenetic analysis for the non-specialist. *Nucleic acids research*, *36*(suppl 2), 465-469.

Desikan, R., Soheila, A. H., Hancock, J. T., and Neill, S. J. (2001). Regulation of the Arabidopsis transcriptome by oxidative stress. Plant physiol., 127(1): 159-172.

Desprez, T., Juraniec, M., Crowell, E. F., Jouy, H., Pochylova, Z., Parcy, F., Hofte, H., Gonneau, M. and Vernhettes, S. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc. Natl. Acad. Sci.104(39): 15572-15577.

Dhugga, K. S., Tiwari, S. C., and Ray, P. M. (1997) A reversibly glycosylated polypeptide (RGP1) possibly involved in plant cell wall synthesis: purification, gene cloning, and trans-Golgi localization. Proc. Natl. Acad. Sci. 94(14): 7679-7684.

Dietrich, R., Ploss, K., and Heil, M. (2005). Growth responses and fitness costs after induction of pathogen resistance depend on environmental conditions. Plant, Cell & Environ., 28(2): 211-222.

Doerks, T., Copley, R. R., Schultz, J., Ponting, C. P., and Bork, P. (2002) Systematic identification of novel protein domain families associated with nuclear functions. Genome Res. 12(1): 47-56.

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., Dixon, J. E., and Ecker, J. R. (2012) Widespread dynamic DNA methylation in response to biotic stress. Proc. Natl. Acad. Sci. 109(32): E2183-2191.

Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. Nucleic acids res., gkq310

Erkina, T. Y. and Erkine, A. M. (2006). Displacement of histones at promoters of Saccharomyces cerevisiae heat shock genes is differentially associated with histone H3 acetylation. Mol. Cell. Biol., 26(20): 7587-7600.

Erkine, A. M., Magrogan, S. F., Sekinger, E. A., and Gross, D. S. (1999). Cooperative binding of heat shock factor to the yeast HSP82 promoter in vivo and in vitro. Mol. Cell. biol., 19(3): 1627-1639.

Ezer, D., Zabet, N. R., and Adryan, B. (2014). Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression. Comput. struct. Biotech. J., 10(17), 63-69.

Fabian, M. R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. Annu. Rev. Biochem., 79: 351-379.

Farr, S. B., and Kogoma, T. (1991). Oxidative stress responses in Escherichia coli and Salmonella typhimurium. Microbiol. Rev., 55(4), 561-585.

Farzadfard, F., Perli, S. D., and Lu, T. K. (2013) Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. ACS Synth. Biol., 2(10): 604-613.

Feder, M. E., and Hofmann, G. E. (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annu.rev. of physiol., 61(1): 243-282.

Fields, S. and Song, O.-k (1989) A novel genetic system to detect protein–protein interactions. Nature 340(6230): 245-246.

Flicek, P., I. Ahmed, M. R. Amode, D. Barrell, K. Beal, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fairley, S. Fitzgerald, L. Gil, C. Garcia-Giron, L. Gordon, T. Hourlier, S. Hunt, T. Juettemann, A. K. Kahari, S. Keenan, M. Komorowska, E. Kulesha, I. Longden, T. Maurel, W. M. McLaren, M. Muffato, R. Nag, B. Overduin, M. Pignatelli, B. Pritchard, E. Pritchard, H. S. Riat, G. R. Ritchie, M. Ruffier, M. Schuster, D. Sheppard, D. Sobral, K. Taylor, A. Thormann, S. Trevanion, S. White, S. P. Wilder, B. L. Aken, E. Birney, F. Cunningham, I. Dunham, J. Harrow, J. Herrero, T. J. Hubbard, N. Johnson, R. Kinsella, A. Parker, G. Spudich, A. Yates, A. Zadissa and S. M. Searle (2013). Ensembl 2013. Nucleic Acids Res. 41(Database issue): D48-55.

Flick, K. E., Gonzalez, L., Harrison, C. J., and Nelson, H. C. (1994). Yeast heat shock transcription factor contains a flexible linker between the DNA-binding and trimerization domains. Implications for DNA binding by trimeric proteins. J. Biol. Chem., 269(17): 12475-12481.

Folsom, J. P., Richards, L., Pitts, B., Roe, F., Ehrlich, G. D., Parker, A., Mazurie, A., and Stewart, P. S. (2010) Physiology of Pseudomonas aeruginosa in biofilms as revealed by transcriptome analysis. BMC Microbiol., 10: 294.

Fritsch, M., and Wu, C. (1999). Phosphorylation of Drosophila heat shock transcription factor. Cell stress chaperones, 4(2): 102.

Fu, X., N. Fu, S. Guo, Z. Yan, Y. Xu, H. Hu, C. Menzel, W. Chen, Y. Li, R. Zeng and P. Khaitovich (2009). Estimating accuracy of RNA-Seq and microarrays with proteomics. BMC Genomics 10: 161.

Fujimoto, M. and A. Nakai (2010) The heat shock factor family and adaptation to proteotoxic stress. FEBS J. 277(20): 4112-4125.
Fujimoto, S. Y., Ohta, M., Usui, A., Shinshi, H., & Ohme-Takagi, M. (2000). Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box–mediated gene expression. Plant Cell, 12(3): 393-404.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003). Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. J. Biol. Chem., 278(26): 24132-24138.

Gao, F., Foat, B. C., and Bussemaker, H. J. (2004). Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. BMC bioinformatics, 5(1): 31.

Gécz, J., Shaw, M. A., Bellon, J. R., and de Barros Lopes, M. (2003) Human wild-type SEDL protein functionally complements yeast Trs20p but some naturally occurring SEDL mutants do not. Gene., 320(0): 137-144.

Gerber, H. P., Seipel, K., Georgiev, O., Hofferer, M., Hug, M., Rusconi, S., and Schaffner, W. (1994). Transcriptional activation modulated by homopolymeric glutamine and proline stretches. Science, 263(5148): 808-811.

Gill, S. S. and Tuteja, N. (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol. Biochem., 48(12): 909-930.

Giraud, E., Ng, S., Carrie, C., Duncan, O., Low, J., Lee, C. P., ... and Whelan, J. (2010). TCP transcription factors link the regulation of genes encoding mitochondrial proteins with the circadian clock in Arabidopsis thaliana. Plant Cell, 22(12), 3921-3934.

Gómez-Gómez, L., and Boller, T. (2000) FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis, Mol. Cell, Volume 5(6): 1003-1011.

Gómez-Gómez, L., Bauer, Z., and Boller, T. (2001). Both the Extracellular Leucine-Rich Repeat Domain and the Kinase Activity of FLS2 Are Required for Flagellin Binding and Signaling in Arabidopsis. Plant Cell, 13(5): 1155–1164.

Gonsalves, S. E., Moses, Razak, Z., Robert, F., and Westwood, J. T. (2011) Whole-genome analysis reveals that active heat shock factor binding sites are mostly associated with non-heat shock genes in Drosophila melanogaster. PLoS One 6(1): e15934.

Gotea, V., Visel, A., Westlund, J. M., Nobrega, M. A., Pennacchio, L. A., and Ovcharenko, I. (2010). Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. Genome res., 20(5), 565-577.

Guertin, M. J., and Lis, J. T. (2010). Chromatin landscape dictates HSF binding to target DNA elements. PLoS genet., 6(9), e1001114.

Guettouche, T., Boellmann, F., Lane, W. S. and Voellmy, R. (2005) Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. BMC Biochem., 6: 4.

Guhaniyogi, J. and Brewer, G. (2001) Regulation of mRNA stability in mammalian cells. Gene 265(1–2): 11-23.

Hahn, J. S., Hu, Z., Thiele, D. J., and Iyer, V. R. (2004). Genome-wide analysis of the biology of stress responses through heat shock transcription factor.Mol. cell. Biol., 24(12), 5249-5256.

Hahn, S. and Young, E. T., (2011). Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics 189(3): 705-736.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature, 405(6785): 486-489.

Harper, J., Adami, G. R., Wei, N., Keyomarsi, K., and Elledge, S. J. (1993) The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell, 75(4): 805-816.

Hartl, F. U. and Hayer-Hartl, M. (2009) Converging concepts of protein folding in vitro and in vivo. Nat. Struct. Mol. Biol., 16(6): 574-581.

Hartmann, U., Höhmann, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000). Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. Plant J., 21(4): 351-360.

Herdegen, T., and Leah, J. (1998) Inducible and constitutive transcription factors in the mammalian nervous system: control of gene expression by Jun, Fos and Krox, and CREB/ATF proteins, Brain Res. Rev., 28(3): 370-490.

Hietakangas, V., Ahlskog, J. K., Jakobsson, A. M., Hellesuo, M., Sahlberg, N. M., Holmberg, C. I., ... and Sistonen, L. (2003). Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. Mol. Cell. Biol., 23(8), 2953-2968.

Hirt, H. (1999). Transcriptional upregulation of signaling pathways: more complex than anticipated?. TIPS., 4(1): 7-8.

Ho, J. W., Bishop, E., Karchenko, P. V., Nègre, N., White, K. P., & Park, P. J. (2011). ChIP-chip versus ChIP-seq: lessons for experimental design and data analysis. BMC genomics, 12(1), 134.

Holmberg, C. I., Hietakangas, V., Mikhailov, A., Rantanen, J. O., Kallio, M., Meinander, A., ... and Sistonen, L. (2001). Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. EMBO j., 20(14): 3800-3810.

Holmberg, C., Tran, S., Eriksson, J., and Sistonen, L. (2002) Multisite phosphorylation provides sophisticated regulation of transcription factors, TIBS., 27(12): 619-627.

Hu, W., Hu, G., and Han, B. (2009) Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed overlapped and stress specific response under abiotic stresses in rice. Plant Sci., 176(4): 583-590.

Hu, W., Hu, G., & Han, B. (2009). Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed overlapped and stress specific response under abiotic stresses in rice. Plant Sci., 176(4), 583-590.

Hübel, A., Lee, J. H., Wu, C., and Schöffl, F. (1995). Arabidopsis heat shock factor is constitutively active in Drosophila and human cells. MGG., 248(2): 136-141.

Hull, R. P., Srivastava, P. K., D'Souza, Z., Atanur, S. S., Mechta-Grigoriou, F., Game, L., ... and Behmoaras, J. (2013). Combined ChIP-Seq and transcriptome analysis identifies AP-1/JunD as a primary regulator of oxidative stress and IL-1β synthesis in macrophages. BMC genomics, 14(1), 92.

Hwang, S. M., Kim, D. W., Woo, M. S., Jeong, H. S., Son, Y. S., Akhter, S., ... and Bahk, J. D. (2014). Functional characterization of Arabidopsis HsfA6a as a heat-shock transcription factor under high salinity and dehydration conditions.Plant, cell & environment, 37(5), 1202-1222.

Hyodo, H., Yamakawa, S., Takeda, Y., Tsuduki, M., Yokota, A., Nishitani, K., and Kohchi, T. (2003). Active gene expression of a xyloglucan endotransglucosylase/hydrolase gene, XTH9, in inflorescence apices is related to cell elongation in Arabidopsis thaliana. Plant. Mol. Biol., 52(2): 473-482.

Irvine, R. A., Lin, I. G., and Hsieh, C. L. (2002). DNA Methylation Has a Local Effect on Transcription and Histone Acetylation. Mol. Cell. Biol., 22(19): 6689-6696.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci., 98(8): 4569-4574.

JACKSON, D. A., Pombo, A. N. A., and IBORRA, F. (2000). The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells. FASEB J., 14(2): 242-254.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol., 3(3): 318-356.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet., 33: 245-254.

Jakobsen, B. K., and Pelham, H. R. (1988). Constitutive binding of yeast heat shock factor to DNA in vivo. Mol. Cell. biol., 8(11): 5040-5042.

Jauniaux, E., Watson, A. L., Hempstock, J., Bao, Y. P., Skepper, J. N., and Burton, G. J. (2000). Onset of maternal arterial blood flow and placental oxidative stress: a possible factor in human early pregnancy failure. The American journal of pathology, 157(6): 2111-2122.

Jedlicka, P., Mortin, M. A., and Wu, C. (1997). Multiple functions of Drosophila heat shock transcription factor in vivo. EMBO j., 16(9): 2452-2462.

Jiang, C., Lamblin, A. F. J., Steller, H., and Thummel, C. S. (2000). A Steroid-Triggered Transcriptional Hierarchy Controls Salivary Gland Cell Death during Drosophila Metamorphosis. Mol. Cell, 5(3): 445-455.

Jiang, H., Wang, F., Dyer, N. P., and Wong, W. H. (2010). CisGenome Browser: a flexible tool for genomic data visualization. Bioinformatics, 26(14), 1781-1782.

Johansson, H., Jones, H. J., Foreman, J., Hemsted, J. R., Stewart, K., Grima, R., and Halliday, K. J., (2014) Arabidopsis cell expansion is controlled by a photothermal switch. Nat. Commun., 5: 4848.

Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science, 316(5830): 1497-1502.

Jung, H. S., Crisp, P. A., Estavillo, G. M., Cole, B., Hong, F., Mockler, T. C., ... and Chory, J. (2013) Subset of heat-shock transcription factors required for the early response of Arabidopsis to excess light. Proc. Natl. Acad. Sci., 110(35): 14474-14479.

Juven-Gershon, T., and Kadonaga, J. T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev. Biol., 339(2), 225-229.

Kass,S., Pruss, D., and Wolffe, A. (1997) How does DNA methylation repress transcription?, TIG., 13(11): 444-449.

Kaufmann, K., Muino, J. M., Jauregui, R., Airoldi, C. A., Smaczniak, C., Krajewski, P., and Angenent, G. C. (2009). Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the Arabidopsis flower. PLoS biol., 7(4), e1000090.

Kim, B. H., and Schöffl, F. (2002). Interaction between Arabidopsis heat shock transcription factor 1 and 70 kDa heat shock proteins. J. Exp. Bot., 53(367): 371-375.

Kim, J. M., To, T. K., Nishioka, T., and Seki, M. (2010) Chromatin regulation functions in plant abiotic stress responses. Plant Cell. Environ., 33(4): 604-611.

Kim, S. Y., Chung, H. J., and Thomas, T. L. (1997). Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. Plant J., 11(6): 1237-1251.

Kobayashi, T., Eun, C. H., Hanai, H., Matsubayashi, Y., Sakagami, Y., and Kamada, H. (1999). Phytosulphokine-α, a peptidyl plant growth factor, stimulates somatic embryogenesis in carrot. J. Exp. Bot.,50(336): 1123-1128.

Komori, T. (2002). Runx2, a multifunctional transcription factor in skeletal development. J. cell. Biochem., 87(1): 1-8.

Kondo, N., M. Katsuno, H. Adachi, M. Minamiyama, H. Doi, S. Matsumoto, Y. Miyazaki, M. Iida, G. Tohnai, H. Nakatsuji, S. Ishigaki, Y. Fujioka, H. Watanabe, F. Tanaka, A. Nakai and G. Sobue (2013) Heat shock factor-1 influences pathological lesion distribution of polyglutamine-induced neurodegeneration. Nat. Commun., 4: 1405.

Kotak, S., Port, M., Ganguli, A., Bicker, F., and Koskull-Döring, V. (2004). Characterization of C-terminal domains of Arabidopsis heat stress transcription factors (Hsfs) and identification of a new signature combination of plant class A Hsfs with AHA and NES motifs essential for activator function and intracellular localization. Plant J., 39(1): 98-112.

Kregel, K. C. (2002). Heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. J. Appl. Physiol. (1985) 92(5): 2177-2186.

Krishna, P., and Gloor, G. (2001). The Hsp90 family of proteins in Arabidopsis thaliana. Cell stress chaperones 6(3): 238.

Krizek, B. (2009) AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. Plant Physiol. 150(4): 1916-1929.

Kumar, M., Busch, W., Birke, H., Kemmerling, B., Nurnberger, T., and Schoffl, F. (2009) Heat shock factors HsfB1 and HsfB2b are involved in the regulation of Pdf1.2 expression and pathogen resistance in Arabidopsis. Mol. Plant., 2(1): 152-165.

Kumar, S. V. and Wigge, P. A. (2010). H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell, 140(1): 136-147.

Kuo, C. J., Conley, P. B., Chen, L., Sladek, F. M., Darnell, J. E., and Crabtree, G. R. (1992) A transcriptional hierarchy involved in mammalian cell-type specification. Nature, 355(6359): 457-461.

Kutschmar, A., Rzewuski, G., Stuhrwohldt, N., Beemster, G. T., Inze, D. and Sauter, M. (2009). PSK-alpha promotes root growth in Arabidopsis. New Phytol., 181(4): 820-831.

Laajala, T. D., Raghav, S., Tuomela, S., Lahesmaa, R., Aittokallio, T., and Elo, L. L. (2009). A practical comparison of methods for detecting transcription factor binding sites in ChIP-seq experiments. BMC genomics, 10(1), 618

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 40(Database issue): D1202-1210.

Landt, S. G., G. K. Marinov, A. Kundaje, P. Kheradpour, F. Pauli, S. Batzoglou, B. E. Bernstein, P. Bickel, J. B. Brown, P. Cayting, Y. Chen, G. DeSalvo, C. Epstein, K. I. Fisher-Aylor, G. Euskirchen, M. Gerstein, J. Gertz, A. J. Hartemink, M. M. Hoffman, V. R. Iyer, Y. L. Jung, S. Karmakar, M. Kellis, P. V. Kharchenko, Q. Li, T. Liu, X. S. Liu, L. Ma, A. Milosavljevic, R. M. Myers, P. J. Park, M. J. Pazin, M. D. Perry, D. Raha, T. E. Reddy, J. Rozowsky, N. Shoresh, A. Sidow, M. Slattery, J. A. Stamatoyannopoulos, M. Y. Tolstorukov, K. P. White, S. Xi, P. J. Farnham, J. D. Lieb, B. J. Wold and M. Snyder (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res., 22(9): 1813-1831.

Langmead, B. and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. Nat. Meth., 9(4): 357-359.

Lata, C., Yadav, A., and Prasad, M. (2011). Role of plant transcription factors in abiotic stress tolerance. Abiotic Stress Response in Plants. Physiological, Biochemical and Genetic Perspectives, 261-296.
Lee, J. H., Hübel, A., and Schöffl, F. (1995). Derepression of the activity of genetically engineered heat shock factor causes constitutive synthesis of heat shock proteins and increased thermotolerance in transgenic Arabidopsis. Plant J., 8(4): 603-612.

Lee, J. H., S. J. Yoo, S. H. Park, I. Hwang, J. S. Lee and J. H. Ahn (2007) Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. Genes Dev., 21(4): 397-402.

Lee, S., Carlson, T., Christian, N., Lea, K., Kedzie, J., Reilly, J. P., and Bonner, J. J. (2000). The yeast heat shock transcription factor changes conformation in response to superoxide and temperature. Mol. Biol. Cell., 11(5): 1753-1764.

Lee, T. I., and Young, R. A. (2013). Transcriptional regulation and its misregulation in disease. Cell, 152(6), 1237-1251.

Lelli, K. M., Slattery, M., and Mann, R. S. (2012). Disentangling the many layers of eukaryotic transcriptional regulation. Annu. Rev. Genet., 46: 43-68.

Lelli, K. M., Slattery, M., and Mann, R. S. (2012). Disentangling the many layers of eukaryotic transcriptional regulation. Annu. Rev. Genet., 46, 43-68.

Lerouxel, O., Mouille, G., Andeme-Onzighi, C., Bruyant, M. P., Seveno, M., Loutelier-Bourhis, C., Driouich, A., Hofte, H., and Lerouge, P. (2005) Mutants in DEFECTIVE GLYCOSYLATION, an Arabidopsis homolog of an oligosaccharyltransferase complex subunit, show protein underglycosylation and defects in cell differentiation and growth. Plant J., 42(4): 455-468.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16): 2078-2079.

Li, J. J., and Herskowitz, I. (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science, 262(5141): 1870-1874.

Li, J., Liu, Z. J., Pan, Y. C., Liu, Q., Fu, X., Cooper, N. G., ... & Shi, T. (2007). Regulatory module network of basic/helix-loop-helix transcription factors in mouse brain. Genome biol., 8(11), R244.

Li, M., J. Doll, K. Weckermann, C. Oecking, K. W. Berendzen and F. Schoffl (2010) Detection of in vivo interactions between Arabidopsis class A-HSFs, using a novel BiFC fragment, and identification of novel class B-HSF interacting proteins. Eur. J. Cell. Biol., 89(2-3): 126-132.

Li, Q., J. B. Brown, H. Huang and P. J. Bickel (2011) Measuring reproducibility of high-throughput experiments. Annals. of Applied Statistics 5(3): 1752-1779.

Li, Z., Zhang, L., Wang, A., Xu, X., & Li, J. (2013). Ectopic overexpression of SlHsfA3, a heat stress transcription factor from tomato, confers increased thermotolerance and salt hypersensitivity in germination in transgenic Arabidopsis. PloS one, 8(1), e54880

Liang, F., Cunningham, K. W., Harper, J. F., and Sze, H. (1997). ECA1 complements yeast mutants defective in Ca2+ pumps and encodes an endoplasmic reticulum-type Ca2+-ATPase in Arabidopsis thaliana. Proc. Natl. Acad. Sci., 94(16): 8579-8584.

Lin, B. L., Wang, J. S., Liu, H. C., Chen, R. W., Meyer, Y., Barakat, A., and Delseny, M. (2001). Genomic analysis of the Hsp70 superfamily in Arabidopsis thaliana. Cell stress chaperones, 6(3): 201.

Lindquist, S. (1986). The heat-shock response. Annu. Rev. Biochem.,55(1): 1151-1191.

Linster, C. L., T. A. Gomez, K. C. Christensen, L. N. Adler, B. D. Young, C. Brenner and S. G. Clarke (2007) Arabidopsis VTC2 encodes a GDP-L-galactose phosphorylase, the last unknown enzyme in the Smirnoff-Wheeler pathway to ascorbic acid in plants. J. Biol. Chem., 282(26): 18879-18885.

Lis, J. T., Simon, J. A., and Sutton, C. A. (1983) New heat shock puffs and β=galactosidase activity resulting from transformation of Drosophila with an hsp70-lacZ hybrid gene. Cell 35(2, Part 1): 403-410.
Littlefield, O., and Nelson, H. C. (1999). A new use for the'wing'of the'winged'helix-turn-helix motif in the HSF–DNA cocrystal. Nat. Struct. Mol. Biol., 6(5): 464-470.

Liu, H. C. and Y. Y. Charng (2013) Common and distinct functions of Arabidopsis class A1 and A2 heat shock factors in diverse abiotic stress responses and development. Plant Physiol., 163(1): 276-290.

Liu, J., Qin, Q., Zhang, Z., Peng, R., Xiong, A., Chen, J., and Quan- Yao, H. (2009) OsHSF7 gene in rice, Oryza sativa L., encodes a transcription factor that functions as a high temperature receptive and responsive factor. BMB Rep., 42(1): 16 – 21.

Liu, P. C. C. and D. J. Thiele (1999) Modulation of Human Heat Shock Factor Trimerization by the Linker Domain. J. Biol. Chem., 274(24): 17219-17225.

Liu, X. D., Liu, P. C., Santoro, N., & Thiele, D. J. (1997). Conservation of a stress response: human heat shock transcription factors functionally substitute for yeast HSF. EMBO j., 16(21): 6466-6477.

Liu, X., and Matsumura, P. (1994). The FlhD/FlhC complex, a transcriptional activator of the Escherichia coli flagellar class II operons. J. bacteriol., 176(23): 7345-7351.

Lohmann, C., G. Eggers-Schumacher, M. Wunderlich and F. Schoffl (2004) Two different heat shock transcription factors regulate immediate early expression of stress genes in Arabidopsis. Mol. Genet. Genomics, 271(1): 11-21.

Loivamäki, M., Stührwohldt, N., Deeken, R., Steffens, B., Roitsch, T., Hedrich, R., & Sauter, M. (2010). A role for PSK signaling in wounding and microbial interactions in Arabidopsis. Physiologia plant., 139(4), 348-357.

Loughran, Ö., & La Thangue, N. B. (2000). Apoptotic and growth-promoting activity of E2F modulated by MDM2. Molecular and cellular biology, 20(6), 2186-2197.

Luo, M., Y. Y. Wang, X. Liu, S. Yang, Q. Lu, Y. Cui and K. Wu (2012). HD2C interacts with HDA6 and is involved in ABA and salt stress response in Arabidopsis. J. Exp. Bot., 63(8): 3297-3306.

Luo, R. X., and Dean, D. C. (1999). Chromatin remodeling and transcriptional regulation. JNCI., 91(15): 1288-1294.

Malone, J. H. and Oliver, B. (2011) Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol. 9: 34.

Mandel, A., Gustafson-Brown, M., C., Savidge, B., and Yanofsky, M. F. (1992) Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. Nature 360(6401): 273-277.

Marchler, G., and Wu, C. (2001) Modulation of Drosophila heat shock transcription factor activity by the molecular chaperone DROJ1. EMBO j., 20(3): 499-509.

Marioni, J. C., C. E. Mason, S. M. Mane, M. Stephens and Y. Gilad (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 18(9): 1509-1517.

Massie, C. E. and I. G. Mills (2008) ChIPping away at gene regulation. EMBO Rep., 9(4): 337-343.

Matsubayashi, Y., and Sakagami, Y. (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of Asparagus officinalis L. Proc. Natl. Acad. Sci., 93(15): 7623-7627.

Matsubayashi, Y., Takagi, L., Omura, N., Morita, A., and Sakagami, Y. (1999). The endogenous sulfated pentapeptide phytosulfokine-α stimulates tracheary element differentiation of isolated mesophyll cells of Zinnia. Plant physiol.,120(4): 1043-1048.

Matthias, P. and A. G. Rolink (2005) Transcriptional networks in developing and mature B cells. Nat. Rev. Immunol. 5(6): 497-508.

McKenna, N. J., & O'Malley, B. W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. Cell, 108(4), 465-474.

Mendillo, M. L., Santagata, S., Koeva, M., Bell, G. W., Hu, R., Tamimi, R. M., ... & Lindquist, S. (2012). HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. Cell, 150(3), 549-562.

Mercier, E., A. Droit, L. Li, G. Robertson, X. Zhang and R. Gottardo (2011). "An integrated pipeline for the genome-wide analysis of transcription factor binding sites from ChIP-Seq. PLoS One 6(2): e16432.

Miller, G., and Mittler, R. (2006). Could heat shock transcription factors function as hydrogen peroxide sensors in plants?. Annals of Botany,98(2), 279-288.

Miller, G., V. Shulaev and Mittler, R. (2008) Reactive oxygen signaling and abiotic stress. Physiol. Plant. 133(3): 481-489.

Min, J. N., Zhang, Y., Moskophidis, D., and Mivechi, N. F. (2004) Unique contribution of heat shock transcription factor 4 in ocular lens development and fiber cell differentiation. Genesis 40(4): 205-217.

Mishra, S. K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., & Scharf, K. D. (2002). In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. Genes dev., 16(12), 1555-1567.

Mitsuda, N., M. Ikeda, S. Takada, Y. Takiguchi, Y. Kondou, T. Yoshizumi, M. Fujita, K. Shinozaki, M. Matsui and M. Ohme-Takagi (2010) Efficient yeast one-/two-hybrid screening using a library composed only of transcription factors in Arabidopsis thaliana. Plant Cell Physiol., 51(12): 2145-2151.

Mittler, R. (2006) Abiotic stress, the field environment and stress combination. Trends Plant Sci., 11(1): 15-19.

Morano, K. A., Santoro, N., Koch, K. A., and Thiele, D. J. (1999). A trans-activation domain in yeast heat shock transcription factor is essential for cell cycle progression during stress. Molecular and cellular biology, 19(1), 402-411.

Morgan, R. (2006) Engrailed: complexity and economy of a multi-functional transcription factor. FEBS Lett., 580(11): 2531-2533.

Morimoto, R. I. (1998). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. Genes dev., 12(24), 3788-3796.

Morozova, O., Hirst, M., and Marra, M. A. (2009) Applications of new sequencing technologies for transcriptome analysis. Annu. Rev. Genomics Hum. Genet., 10: 135-151.

Mosher, S., Seybold, H., Rodriguez, P., Stahl, M., Davies, K. A., Dayaratne, S., Morillo, S. A., Wierzba, M., Favery, B., Keller, H., Tax, F. E., and Kemmerling, B. (2013) The tyrosine-sulfated peptide receptors PSKR1 and PSY1R modify the immunity of Arabidopsis to biotrophic and necrotrophic pathogens in an antagonistic manner. Plant J., 73(3): 469-482.

Mumberg, D., Müller, R., and Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene 156(1): 119-122.

Mutch, D. M., Berger, A., Mansourian, R., Rytz, A., and Roberts, M. A. (2002). The limit fold change model: a practical approach for selecting differentially expressed genes from microarray data. BMC bioinformatics, 3(1): 17.

Neph, S., A. B. Stergachis, A. Reynolds, R. Sandstrom, E. Borenstein and J. A. Stamatoyannopoulos (2012) Circuitry and dynamics of human transcription factor regulatory networks. Cell 150(6): 1274-1286.

Nicol, J. W., Helt, G. A., Blanchard Jr, S. G., Raja, A. and Loraine, A. E. (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics 25(20): 2730-2731.

Niehus, E., Gressmann, H., Ye, F., Schlapbach, R., Dehio, M., Dehio, C., Stack, A., Meyer, T. F., Suerbaum, S., and Josenhans, C. (2004) Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of Helicobacter pylori. Mol. Microbiol. 52(4): 947-961.

Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K., and Shigeoka, S. (2006). Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. Plant J., 48(4), 535-547.

Nishizawa-Yokoi, A., Nosaka, R., Hayashi, H., Tainaka, H., Maruta, T., Tamoi, M., Ikeda, M., Ohme-Takagi, M., Yoshimura, K., Yabuta, Y., and Shigeoka, S. (2011) HsfA1d and HsfA1e involved in the transcriptional regulation of HsfA2 function as key regulators for the Hsf signaling network in response to environmental stress. Plant Cell Physiol., 52(5): 933-945.

Niu, M., Tabari, E. S., and Su, Z. (2014). De novo prediction of cis-regulatory elements and modules through integrative analysis of a large number of ChIP datasets. BMC genomics, 15(1): 1047.

Noel, L. D., G. Cagna, J. Stuttmann, L. Wirthmuller, S. Betsuyaku, C. P. Witte, R. Bhat, N. Pochon, T. Colby and J. E. Parker (2007) Interaction between SGT1 and cytosolic/nuclear HSC70 chaperones regulates Arabidopsis immune responses. Plant Cell 19(12): 4061-4076.

Nover, L., Bharti, K., Döring, P., Mishra, S. K., Ganguli, A., and Scharf, K. D. (2001). Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need?. Cell stress chaperones, 6(3): 177.

Nowak, S. J., & Corces, V. G. (2000). Phosphorylation of histone H3 correlates with transcriptionally active loci. Genes & development, 14(23), 3003-3013.

Ogawa, D., K. Yamaguchi and T. Nishiuchi (2007) High-level overexpression of the Arabidopsis HsfA2 gene confers not only increased themotolerance but also salt/osmotic stress tolerance and enhanced callus growth. J. Exp. Bot., 58(12): 3373-3383.

O'Geen, H., S. Frietze and P. J. Farnham (2010) Using ChIP-seq technology to identify targets of zinc finger transcription factors. Methods Mol. Biol. 649: 437-455.

Oh, Y. M., J. K. Kim, S. Choi and J. Y. Yoo (2012). Identification of co-occurring transcription factor binding sites from DNA sequence using clustered position weight matrices. Nucleic Acids Res., 40(5): e38.

Ohno, C. K., G. V. Reddy, M. G. Heisler and E. M. Meyerowitz (2004) The Arabidopsis JAGGED gene encodes a zinc finger protein that promotes leaf tissue development. Development 131(5): 1111-1122.

Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., and Tasaka, M. (2007) ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in Arabidopsis. Plant Cell 19(1): 118-130.

Östling, P., Bjork, J. K., Roos-Mattjus, P., Mezger, V., and Sistonen, L. (2007) Heat shock factor 2 (HSF2) contributes to inducible expression of hsp genes through interplay with HSF1. J. Biol. Chem., 282(10): 7077-7086.

Pabo, C. O. and Sauer, R. T. (1992) Transcription Factors: Structural Families and Principles of DNA Recognition. Annu. Rev. Biochem., 61(1): 1053-1095.

Pajerowska-Mukhtar, K. M., W. Wang, Y. Tada, N. Oka, C. L. Tucker, J. P. Fonseca and X. Dong (2012) The HSF-like transcription factor TBF1 is a major molecular switch for plant growth-to-defense transition. Curr. Biol., 22(2): 103-112.

Palomero, T., Lim, W. K., Odom, D. T., Sulis, M. L., Real, P. J., Margolin, A., ... & Ferrando, A. A. (2006). NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth.Proceedings of the National Academy of Sciences, 103(48), 18261-18266.

Panchuk, II, Volkov, R. A., and Schoffl, F. (2002) Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in Arabidopsis. Plant Physiol. 129(2): 838-853.

Panikulangara, T. J., Eggers-Schumacher, G., Wunderlich, M., Stransky, H., & Schöffl, F. (2004). Galactinol synthase1. A novel heat shock factor target gene responsible for heat-induced synthesis of raffinose family oligosaccharides in Arabidopsis. Plant physiol., 136(2), 3148-3158.

Parsell, D. A., and Lindquist, S. (1993) The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. Annu. Rev. genet., 27(1): 437-496.

Pavelka, N., G. Rancati, J. Zhu, W. D. Bradford, A. Saraf, L. Florens, B. W. Sanderson, G. L. Hattem and R. Li (2010) Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature 468(7321): 321-325.

Pawitan, Y., S. Michiels, S. Koscielny, A. Gusnanto and A. Ploner (2005) False discovery rate, sensitivity and sample size for microarray studies. Bioinformatics 21(13): 3017-3024.

Pellegrini (2012) Genome-wide evaluation of histone methylation changes associated with leaf senescence in Arabidopsis. PLoS One 7(3): e33151.

Peng, H. and Zhang, J. (2009) Plant genomic DNA methylation in response to stresses: Potential applications and challenges in plant breeding. Prog. Nat. Sci. 19(9): 1037-1045.

Pérez-Salamó, I., Papdi, C., Rigo, G., Zsigmond, L., Vilela, B., Lumbreras, V., Nagy, I., Horvath, B., Domoki, M., Darula, Z., Medzihradszky, K., Bogre, L., Koncz, C., and Szabados, L. (2014) The heat shock factor A4A confers salt tolerance and is regulated by oxidative stress and the mitogen-activated protein kinases MPK3 and MPK6. Plant Physiol., 165(1): 319-334.

Persson, S., Paredez, A., Carroll, A., Palsdottir, H., Doblin, M., Poindexter, P., Khitrov, N. Auer, M. and Somerville, C. R. (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in Arabidopsis. Proc. Natl. Acad. Sci. 104(39): 15566-15571.

Piao, H. L., Pih, K. T., Lim, J. H., Kang, S. G., Jin, J. B., Kim, S. H., and Hwang, I. (1999) An Arabidopsis GSK3/shaggy-like gene that complements yeast salt stress-sensitive mutants is induced by NaCl and abscisic acid. Plant physiol., 119(4): 1527-1534.

Pines, J. and Hunter, T. (1989) Isolation of a human cyclin cDNA: Evidence for cyclin mRNA and protein regulation in the cell cycle and for interaction with p34cdc2. Cell 58(5): 833-846.

Pirkkala, L., Nykänen, P., and Sistonen, L. E. A. (2001). Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. FASEB J., 15(7), 1118-1131.

Piterková, J., Luhová, L., Mieslerová, B., Lebeda, A., and Petřivalský, M. (2013) Nitric oxide and reactive oxygen species regulate the accumulation of heat shock proteins in tomato leaves in response to heat shock and pathogen infection. Plant Science 207(0): 57-65.

Prändl, R., Hinderhofer, K., Eggers-Schumacher, G., and Schöffl, F. (1998) HSF3, a new heat shock factor from Arabidopsis thaliana, derepresses the heat shock response and confers thermotolerance when overexpressed in transgenic plants. Mol. Gen. Genet., 258(3), 269-278.

Qian, J., Chen, J., Liu, Y. F., Yang, L. L., Li, W. P., and Zhang, L. M. (2014) Overexpression of Arabidopsis HsfA1a enhances diverse stress tolerance by promoting stress-induced Hsp expression. Genet. Mol. Res. 13(1): 1233-1243.

Qiu, D., Xiao, J., Xie, W., Cheng, H., Li, X., & Wang, S. (2009). Exploring transcriptional signalling mediated by OsWRKY13, a potential regulator of multiple physiological processes in rice. BMC plant biology, 9(1), 74.

Quinlan, A. R. and I. M. Hall (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6): 841-842.

Ramachandran, S., Hiratsuka, K., and Chua, N.-H. (1994) Transcription factors in plant growth and development. Curr. Opin. Genetics Dev., 4(5): 642-646.

Ray, D. K., Gerber, J. S., MacDonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, *6*.

Ray, A., Robinson-Beers, K., Ray, S., Baker, S. C., Lang, J. D., Preuss, D., ... and Gasser, C. S. (1994) Arabidopsis floral homeotic gene BELL (BEL1) controls ovule development through negative regulation of AGAMOUS gene (AG). Proc. Natl. Acad. Sci., 91(13): 5761-5765.

Reddy, S. and Finlayson, S. A. (2014) Phytochrome B promotes branching in Arabidopsis by suppressing auxin signaling. Plant Physiol. 164(3): 1542-1550.

Reddy, S. K., Liu, S. , Rudd, J. C., Xue, Q., Payton, P., Finlayson, S. A., Mahan, J., Akhunova, A., Holalu, S. V., and Lu, N. (2014) Physiology and transcriptomics of water-deficit stress responses in wheat cultivars TAM 111 and TAM 112. J. Plant Physiol., 171(14): 1289-1298.

Reid, J. E. and Wernisch, L. (2011) STEME: efficient EM to find motifs in large data sets. Nucleic Acids Res. 39(18): e126.

Reindl, A. and Schöffl, F., (1998) Interaction between the Arabidopsis thaliana heat shock transcription factor HSF1 and the TATA binding protein TBP, FEBS Lett., 436(3): 318 – 322.

Rerie, W. G., Feldmann, K. A., and Marks, M. D. (1994). The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. Genes & Development, 8(12), 1388-1399.

Rerie, W. G., Feldmann, K. A., and Marks, M. D. (1994). The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. Genes Dev., 8(12): 1388-1399.

Ritossa, F. (1962) A new puffing pattern induced by temperature shock and DNP in drosophila. Experientia 18(12): 571-573.

Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., and Mittler, R. (2004) When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. Plant Physiol., 134(4): 1683-1696.

Ross, J. (1995) mRNA stability in mammalian cells. Microbiol. Rev., 59(3), 423-450.

Rottger, R., Ruckert, U., Taubert, J., and Baumbach, J. (2012) How little do we actually know? On the size of gene regulatory networks.. IEEE/ACM Trans Comput. Biol. Bioinform. 9(5): 1293-1300.

Sabehat, A., Weiss, D., and Lurie, S. (1998). Heat-shock proteins and cross-tolerance in plants. Physiol. Plant., 103(3): 437-441.

Sakuma, Y., Maruyama, K., Qin, F., Osakabe, Y., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2006) Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. Proc. Natl. Acad. Sci. 103(49): 18822-18827.

Sakurai, H. and Enoki, Y. (2010) Novel aspects of heat shock factors: DNA recognition, chromatin modulation and gene expression. FEBS J., 277(20): 4140-4149.

Saleh, A., Alvarez-Venegas, R., and Avramova, Z. (2008). An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. Nat. Prot., 3(6), 1018-1025.

Santoro, N., Johansson, N., and Thiele, D. J. (1998) Heat shock element architecture is an important determinant in the temperature and transactivation domain requirements for heat shock transcription factor. Mol. Cell. Biol., 18(11): 6340-6352.

Sato, T., Hanada, M., Bodrug, S., Irie, S., Iwama, N., Boise, L. H., ... and Wang, H. G. (1994) Interactions among members of the Bcl-2 protein family analyzed with a yeast two-hybrid system. Proc. Natl. Acad. Sci., 91(20): 9238-9242.

Savic, D., Gertz, J., Jain, P., Cooper, G. M., & Myers, R. M. (2013). Mapping genome-wide transcription factor binding sites in frozen tissues. Epigenetics & chromatin, 6(1), 1-10.

Scharf, K. D., Heider, H., Höhfeld, I., Lyck, R., Schmidt, E., and Nover, L. (1998) The tomato Hsf system: HsfA2 needs interaction with HsfA1 for efficient nuclear import and may be localized in cytoplasmic heat stress granules. Mol. Cell. Biol., 18(4): 2240-2251.

Scharf, K. D., T. Berberich, I. Ebersberger and L. Nover (2012). The plant heat stress transcription factor (Hsf) family: structure, function and evolution. Biochim. Biophys. Acta., 1819(2): 104-119.

Schirmer, E. C., Lindquist, S., and Vierling, E. (1994) An Arabidopsis heat shock protein complements a thermotolerance defect in yeast. Plant Cell, 6(12), 1899-1909.

Schlesinger, M. J. (1990). Heat shock proteins. J. Biol. Chem., 265(21): 12111-12114.

Schmidt, R., J. H. Schippers, A. Welker, D. Mieulet, E. Guiderdoni and B. Mueller-Roeber (2012) Transcription factor OsHsfC1b regulates salt tolerance and development in Oryza sativa ssp. japonica. AoB Plants 2012: pls011.

Schöffl, F., Prändl, R., and Reindl, A. (1998) Regulation of the heat-shock response. Plant physiol., 117(4): 1135-1141.

Schramm, F., Ganguli, A., Kiehlmann, E., Englich, G., Walch, D. and von Koskull-Doring, P. (2006). The heat stress transcription factor HsfA2 serves as a regulatory amplifier of a subset of genes in the heat stress response in Arabidopsis. Plant Mol. Biol., 60(5): 759-772.

Seibel, J., Jördening, H. J., and Buchholz, K. (2006). Glycosylation with activated sugars using glycosyltransferases and transglycosidases. Biocatalysis and biotransformation, 24(5): 311-342.

Sellers, W. R., Rodgers, J. W., and Kaelin, W. G. (1995) A potent transrepression domain in the retinoblastoma protein induces a cell cycle arrest when bound to E2F sites. Proc. Natl. Acad. Sci.,92(25): 11544-11548.

Serrano, M., Hannon, G. J., and Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature, 366(6456): 704-707.

Shamimuzzaman, M., and Vodkin, L. (2013). Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by ChIP-Seq and RNA-Seq. BMC genomics, 14(1), 477.

Sharrock, R. A., and Quail, P. H. (1989). Novel phytochrome sequences in Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family. Genes Dev., 3(11): 1745-1757.

Shin, R., A. Y. Burch, K. A. Huppert, S. B. Tiwari, A. S. Murphy, T. J. Guilfoyle and D. P. Schachtman (2007) The Arabidopsis transcription factor MYB77 modulates auxin signal transduction. Plant Cell 19(8): 2440-2453.

Shin, R., Burch, A. Y., Huppert, K. A., Tiwari, S. B., Murphy, A. S., Guilfoyle, T. J., & Schachtman, D. P. (2007). The Arabidopsis transcription factor MYB77 modulates auxin signal transduction. Plant Cell, 19(8), 2440-2453.

Shiu, S.-H., Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., Hollander, J., Meins, F., and Kovalchuk, I. (2010) Transgenerational Adaptation of Arabidopsis to Stress Requires DNA Methylation and the Function of Dicer-Like Proteins. PLoS ONE 5(3): e9514.
Siberil, Y., Doireau, P., and Gantet, P. (2001). Plant bZIP G-box binding factors.European Journal of Biochemistry, 268(22), 5655-5666.

Siebenlist, U., Franzoso, G., and Brown, K. (1994) Structure, regulation and function of NF-kappaB. Annu. Rev. cell biol., 10(1): 405-455.

Singh, V. and Aballay, A., (2014) Heat Shock and Genetic Activation of HSF-1 Enhance Immunity to Bacteria. Cell Cycle 5(21): 2443-2446.

Smith, B. J., and Yaffe, M. P. (1991) A mutation in the yeast heat-shock factor gene causes temperature-sensitive defects in both mitochondrial protein import and the cell cycle. Mol. cell. biol., 11(5): 2647-2655.

Sorger, P. K. (1990) Yeast heat shock factor contains separable transient and sustained response transcriptional activators. Cell 62(4): 793-805.

Spitz, F. and E. E. Furlong (2012) Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet., 13(9): 613-626.

Stergachis, A. B., Haugen, E., Shafer, A., Fu, W., Vernot, B., Reynolds, A., ... & Stamatoyannopoulos, J. A. (2013). Exonic transcription factor binding directs codon choice and affects protein evolution. Science, 342(6164), 1367-1372.

Sugio, A., Dreos, R., Aparicio, F., and Maule, A. J. (2009). The cytosolic protein response as a subcomponent of the wider heat shock response in Arabidopsis.The Plant Cell Online, 21(2), 642-654.

Swindell, W. R., Huebner, M., and Weber, A. P. (2007). Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. BMC genomics,8(1), 125.

Tacchi, L., Secombes, C. J., Bickerdike, R., Adler, M. A., Venegas, C., Takle, H., and Martin, S. A. (2012) Transcriptomic and physiological responses to fishmeal substitution with plant proteins in formulated feed in farmed Atlantic salmon (Salmo salar). BMC genomics, 13(1): 363.

Tan, P. K., Downey, T. J., Spitznagel Jr, E. L., Xu, P., Fu, D., Dimitrov, D. S., ... and Cam, M. C. (2003) Evaluation of gene expression measurements from commercial microarray platforms. Nucleic acids res., 31(19): 5676-5684.

Tanabe M., Sasai N., Nagata K., Liu X. D., Liu P. C., Thiele D. J., and Nakai A. (1999) The mammalian HSF4 gene generates both an activator and a repressor of heat shock genes by alternative splicing. J. Biol. Chem. (274): 27845–27856

Tate, P. H. and Bird, A. P., (1993) Effects of DNA methylation on DNA-binding proteins and gene expression. Curr. Opin. Genetics Dev., 3(2): 226-231.

Tian D, Traw MB, Chen JQ, Kreitman M, Bergelson J. 2003. Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana.Nature 423: 74–77.

Tian, X., & Lei, Y. (2006). Nitric oxide treatment alleviates drought stress in wheat seedlings. Biol. Plant., 50(4), 775-778.

Tjian, R. and Maniatis, T. (1994) Transcriptional activation: A complex puzzle with few easy pieces. Cell 77(1): 5-8.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn J. L., and Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protocols, 7(3): 562-578.

Trémousaygue, D., Garnier, L., Bardet, C., Dabos, P., Hervé, C., & Lescure, B. (2003). Internal telomeric repeats and 'TCP domain'protein-binding sites co-operate to regulate gene expression in Arabidopsis thaliana cycling cells. The Plant Journal, 33(6), 957-966.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., ... & Rothberg, J. M. (2000). A comprehensive analysis of protein–protein interactions in Saccharomyces cerevisiae. Nature, 403(6770), 623-627.

Valouev, A., Johnson, D. S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., ... & Sidow, A. (2008). Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. Nature methods, 5(9), 829-834.

Varga-Weisz, P. D. (2010) Insights into how chromatin remodeling factors find their target in the nucleus. Proc. Natl. Acad. Sci. U S A 107(46): 19611-19612.

Venturi, C. B., Erkine, A. M., and Gross, D. S. (2000). Cell cycle-dependent binding of yeast heat shock factor to nucleosomes. Mol. cell. Biol., 20(17): 6435-6448.

Vierling, E. (1991) The Roles of Heat Shock Proteins in Plants." Annu.Rev. Plant Physiol. Plant Mol. Biol., 42(1): 579-620.

Vinocur, B. and A. Altman (2005). Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. Curr. Opin. Biotechnol., 16(2): 123-132.

von Koskull-Döring, P., Scharf, K.-D., and Nover, L. (2007) The diversity of plant heat stress transcription factors. TIPS., 12(10): 452-457.

Wagner, M., Hermanns, I., Bittinger, F., and Kirkpatrick, C. J. (1999) Induction of stress proteins in human endothelial cells by heavy metal ions and heat shock.American Journal of Physiology-Lung Cellular and Molecular Physiology,277(5): 1026-1033.

Wahl, V., Ponnu, J., Schlereth, A., Arrivault, S., Langenecker, T., Franke, A., ... and Schmid, M. (2013). Regulation of flowering by trehalose-6-phosphate signaling in Arabidopsis thaliana. Science, 339(6120): 704-707.

Walhout, A. J. M. and Vidal, M. (2001) High-Throughput Yeast Two-Hybrid Assays for Large-Scale Protein Interaction Mapping. Methods 24(3): 297-306.

Wang, C., J. Xu, D. Zhang, Z. A. Wilson and D. Zhang (2010) An effective approach for identification of in vivo protein-DNA binding sites from paired-end ChIP-Seq data. BMC Bioinformatics 11: 81.

Wang, C., Xu, J., Zhang, D., Wilson, Z. A., & Zhang, D. (2010). An effective approach for identification of in vivo protein-DNA binding sites from paired-end ChIP-Seq data. BMC bioinformatics, 11(1), 81.

Wang, G., J. Zhang, D. Moskophidis and N. F. Mivechi (2003) Targeted disruption of the heat shock transcription factor (hsf)-2 gene results in increased embryonic lethality, neuronal defects, and reduced spermatogenesis. Genesis 36(1): 48-61.

Wang, H., Maurano, M. T., Qu, H., Varley, K. E., Gertz, J., Pauli, F., ... & Stamatoyannopoulos, J. A. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome research, 22(9), 1680-1688.

Wang, W., B. Vinocur, O. Shoseyov and A. Altman (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. TIPS., 9(5): 244-252.

Wang, X., M. A. Khaleque, M. J. Zhao, R. Zhong, M. Gaestel and S. K. Calderwood (2006) Phosphorylation of HSF1 by MAPK-activated protein kinase 2 on serine 121, inhibits transcriptional activity and promotes HSP90 binding. J. Biol. Chem. 281(2): 782-791.

Wang, Z., Gerstein, M., and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet., 10(1): 57-63.

Welchen, E., & Gonzalez, D. H. (2006). Overrepresentation of elements recognized by TCP-domain transcription factors in the upstream regions of nuclear genes encoding components of the mitochondrial oxidative phosphorylation machinery. Plant Physiology, 141(2), 540-545.

Weng, M., Y. Yang, H. Feng, Z. Pan, W. H. Shen, Y. Zhu and A. Dong (2014) Histone chaperone ASF1 is involved in gene transcription activation in response to heat stress in Arabidopsis thaliana. Plant Cell Environ., 37(9): 2128-2138.

Westerheide, S. D. and R. I. Morimoto (2005) Heat shock response modulators as therapeutic tools for diseases of protein conformation. J. Biol. Chem. 280(39): 33097-33100.

Westerheide, S. D., Anckar, J., Stevens, S. M., Sistonen, L., and Morimoto, R. I. (2009) Stress-inducible regulation of heat shock factor 1 by the deacetylase SIRT1. Science, 323(5917): 1063-1066.

Whitfield, T. W., Wang, J., Collins, P. J., Partridge, E. C., Aldred, S. F., Trinklein, N. D., ... & Weng, Z. (2012). Functional analysis of transcription factor binding sites in human promoters. Genome Biol, 13(9), R50.

Wiederrecht, G., Seto, D., and Parker, C. S. (1988) Isolation of the gene encoding the S. cerevisiae heat shock transcription factor. Cell 54(6): 841-853.

Wilhelms, M., R. Molero, J. G. Shaw, J. M. Tomas and S. Merino (2011) Transcriptional hierarchy of Aeromonas hydrophila polar-flagellum genes. J. Bacteriol., 193(19): 5179-5190.

Williams, M. E., Foster, R., and Chua, N. H. (1992). Sequences flanking the hexameric G-box core CACGTG affect the specificity of protein binding. The Plant Cell Online, 4(4), 485-496.

Wittkopp, P. J. and Kalay, G. (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet., 13(1): 59-69.

Wray, G. A., Hahn, M. W., Abouheif, E., Balhoff, J. P., Pizer, M., Rockman, M. V., and Romano, L. A. (2003) The evolution of transcriptional regulation in eukaryotes. Mol. Boil. Evol., 20(9): 1377-1419.

Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman and L. A. Romano (2003) The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20(9): 1377-1419.

Wu, C. (1995). Heat Shock Transcription Factors: Structure and Regulation. Annu. Rev. Cell Dev. Biol., 11(1): 441-469.

Wu, K., Tian, L., Zhou, C., Brown, D., and Miki, B. (2003) Repression of gene expression by Arabidopsis HD2 histone deacetylases. Plant J., 34(2): 241-247.

Wu, T. D. and Nacu, S. (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26(7): 873-881.

Wunderlich, M., R. Gross-Hardt and F. Schoffl (2014) Heat shock factor HSFB2a involved in gametophyte development of Arabidopsis thaliana and its expression is controlled by a heat-inducible long non-coding antisense RNA. Plant Mol. Biol., 85(6): 541-550.

Xiang, J., Ran, J., Zou, J., Zhou, X., Liu, A., Zhang, X., Peng, Y., Tang, N., Luo, G., and Chen, X. (2013) Heat shock factor OsHsfB2b negatively regulates drought and salt tolerance in rice. Plant Cell Rep., 32(11): 1795-1806.

Xue, G. P., Sadat, S., Drenth, J., & McIntyre, C. L. (2014). The heat shock factor family from Triticum aestivum in response to heat and other major abiotic stresses and their role in regulation of heat shock protein genes. Journal of experimental botany, 65(2), 539-557

Yamakawa, S., Sakuta, C., Matsubayashi, Y., Sakagami, Y., Kamada, H., and Satoh, S. (1998) The promotive effects of a peptidyl plant growth factor, phytosulfokine-α, on the formation of adventitious roots and expression of a gene for a root-specific cystatin in cucumber hypocotyls. J. Plant Res., 111(3): 453-458.

Yamasaki, K., Kigawa, T., Seki, M., Shinozaki, K., and Yokoyama, S. (2013) DNA-binding domains of plant-specific transcription factors: structure, function, and evolution. TIPS., 18(5): 267-276.

Yang, H., Matsubayashi, Y., Nakamura, K., and Sakagami, Y. (1999) Oryza sativa PSK gene encodes a precursor of phytosulfokine-α, a sulfated peptide growth factor found in plants. Proc. Natl. Acad. Sci., 96(23): 13560-13565.

Yang, H., Y. Matsubayashi, K. Nakamura and Y. Sakagami (2001) Diversity of Arabidopsis Genes Encoding Precursors for Phytosulfokine, a Peptide Growth Factor. Plant physiol., 127(3): 842-851.

Yant, L., Mathieu, J., Dinh, T. T., Ott, F., Lanz, C., Wollmann, H., ... & Schmid, M. (2010). Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor APETALA2. The Plant Cell Online, 22(7), 2156-2170.

Yokotani, N., Ichikawa, T., Kondou, Y., Matsui, M., Hirochika, H., Iwabuchi, M., & Oda, K. (2008). Expression of rice heat stress transcription factor OsHsfA2e enhances tolerance to environmental stresses in transgenic Arabidopsis. Planta, 227(5), 957-967.

Yokotani, N., Ichikawa, T., Kondou, Y., Matsui, M., Hirochika, H., Iwabuchi, M., and Oda, K. (2008) Expression of rice heat stress transcription factor OsHsfA2e enhances tolerance to environmental stresses in transgenic Arabidopsis. Planta 227(5): 957-967.

Yoshida, T., Ohama, N., Nakajima, J., Kidokoro, S., Mizoi, J., Nakashima, K., Maruyama, K., Kim, J. M., Seki, M., Todaka, D., Osakabe, Y., Sakuma, Y., Schoffl, F., Shinozaki, K., and Yamaguchi-Shinozaki, K., (2011) Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. Mol. Genet. Genomics., 286(5-6): 321-332.

Yoshida, T., Y. Sakuma, D. Todaka, K. Maruyama, F. Qin, J. Mizoi, S. Kidokoro, Y. Fujita, K. Shinozaki and K. Yamaguchi-Shinozaki (2008) Functional analysis of an Arabidopsis heat-shock transcription factor HsfA3 in the transcriptional cascade downstream of the DREB2A stress-regulatory system. Biochem. Biophys. Res. Commun., 368(3): 515-521.

Yu, H. D., Yang, X. F., Chen, S. T., Wang, Y. T., Li, J. K., Shen, Q., ... & Guo, F. Q. (2012). Downregulation of chloroplast RPS1 negatively modulates nuclear heat-responsive expression of HsfA2 and its target genes in Arabidopsis. PLoS genetics, 8(5), e1002669.

Yura, T., and Nakahigashi, K. (1999) Regulation of the heat-shock response, Curr. Opin. Microbiol., 2(2): 153-158.

Zaarur, N., V. L. Gabai, J. A. Porco, Jr., S. Calderwood and M. Y. Sherman (2006) Targeting heat shock response to sensitize cancer cells to proteasome and Hsp90 inhibitors." Cancer Res., 66(3): 1783-1791.
Zaret, K. S. and J. S. Carroll (2011) Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 25(21): 2227-2241.

Zenoni, S., A. Ferrarini, E. Giacomelli, L. Xumerle, M. Fasoli, G. Malerba, D. Bellin, M. Pezzotti and M. Delledonne (2010) Characterization of transcriptional complexity during berry development in Vitis vinifera using RNA-Seq. Plant Physiol. 152(4): 1787-1795.

Zhang, H., H. He, X. Wang, X. Wang, X. Yang, L. Li and X. W. Deng (2011). Genome-wide mapping of the HY5-mediated gene networks in Arabidopsis that involve both transcriptional and post-transcriptional regulation. Plant J., 65(3): 346-358.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., ... and Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol, 9(9), R137.

Zhong, M., Orosz, A., and Wu, C. (1998) Direct Sensing of Heat and Oxidation by Drosophila Heat Shock Transcription Factor, Mol. Cell., 2 (1): 101-108.

Zhou, C., Labbe, H., Sridha, S., Wang, L., Tian, L., Latoszek-Green, M., ... and Wu, K. (2004). Expression and function of HD2-type histone deacetylases in Arabidopsis development. The Plant J., 38(5): 715-724.

Zhou, L., CHEUNG, M. Y., Zhang, Q., LEI, C. L., ZHANG, S. H., SUN, S. S. M., and LAM, H. M. (2009). A novel simple extracellular leucine-rich repeat (eLRR) domain protein from rice (OsLRR1) enters the endosomal pathway and interacts with the hypersensitive-induced reaction protein 1 (OsHIR1). Plant, cell environ., 32(12), 1804-1820.

Zhu, L. J., C. Gazin, N. D. Lawson, H. Pages, S. M. Lin, D. S. Lapointe and M. R. Green (2010). "ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data." BMC Bioinformatics 11: 237.

Zinser, E., Sperka-Gottlieb, C. D., Fasch, E. V., Kohlwein, S. D., Paltauf, F., and Daum, G.. (1991). Phospholipid synthesis and lipid composition of subcellular membranes in the unicellular eukaryote Saccharomyces cerevisiae. J. Bacteriol., 173(6): 2026-2034.
Zou,J., Guo, Y., Guettouche, T.,Smith,D. F., and Voellmy, R. (1998) Repression of Heat Shock Transcription Factor HSF1 Activation by HSP90 (HSP90 Complex) that Forms a Stress-Sensitive Complex with HSF1, Cell, 94(4): 471-480.

Zuo, J., Baler, R., Dahl, G., and Voellmy, R. (1994) Activation of the DNA-binding ability of human heat shock transcription factor 1 may involve the transition from an intramolecular to an intermolecular triple-stranded coiled-coil structure. Mol. Cell. Biol., 14(11): 7557-7568.

Appendices:

Appendix A:

List of ChIP-PCR primers used

CHIP_MBF1c-F	GCGGAAACGATACTCCTCAG	AT3G24500
CHIP_MBF1c-R	GTTAGCCGAACCGAATACCA	AT3G24500
CHIP_WRKY58-F	TTCCATGTTTAGCCCGTAGA	AT3G01080
CHIP_WRKY58-R	AATTTTTGTAGATCTCTCTAGATATGG	AT3G01080
CHIP_TAF4b-F1	TTTGCAAAAGTCTGATGCTAATATG	AT1G27720
CHIP_TAF4b-R1	GCCGTTCTGTTGATCTTCTTTT	AT1G27720
CHIP_TAF4b-F2	GGCGCGCAGTGTTTTATC	AT1G27720
CHIP_TAF4b-R2	AGGAATCTCGCGGTACCAA	AT1G27720
CHIP_HSFB2a-F	GATAGCGTTTCACATATTCACAGG	AT5G62020
CHIP_HSFB2a-R	AAAAAGTGAAAAGGGGACCA	AT5G62020
CHIP_ZAT6-F	CTCCACTTGGTTGGTTGGTT	AT5G04340
CHIP_ZAT6-R	ACACGTGTTTGTGGAAGTCG	AT5G04340
CHIP_HSFA7a-F	AAAACACAAAAGGTGGGTCCT	AT3G51910
CHIP_HSFA7a-R	AGGAAGAGTAGAAAGTGAGTGATGA	AT3G51910
CHIP_TFIIS-F	CAAGATAAATGATCAAAACTATTACCA	AT3G10820
CHIP_TFIIS-R	TCTGCAGAAAGTCCACAAGAA	AT3G10820
CHIP_HSFB2b-F	TGGATGACACATCAAAGCAGA	AT4G11660
CHIP_HSFB2b-R	GGCTTCAAGAAACTTCCTATGG	AT4G11660

List of qRT-PCR primers used:

qPCR_NAC12-F	CGCGGAAGTATGGACGTTAT	AT3G12910
qPCR_NAC12-R	AGTTCCCTTTGCCTTCCCTA	AT3G12910
qPCR_NAC72-F	CCCAAAGGCACTAAAACCAA	AT4G27410
qPCR_NAC72-R	CACGACAAGCTTGAACAGGA	AT4G27410
qPCR_NAC25-F	CGACGGTATGGTCGTTTCTT	AT1G61110
qPCR_NAC25-R	AACCCCGAACTCTGAGGAAT	AT1G61110
qPCR_ATAFL-F	CCGGTCACGACTTCTCTCTC	AT3G15500
qPCR_ATAFL-R	TTAGGTCTTGACCCGTTTGG	AT3G15500

Appendix B:

AtHSFA1b binding sites under no stress and heat stress and all the related data files are found in the following links

https://drive.google.com/folderview?id=0B4jdOv6BYryaX0JXRzYzSER0X28&usp=sharing

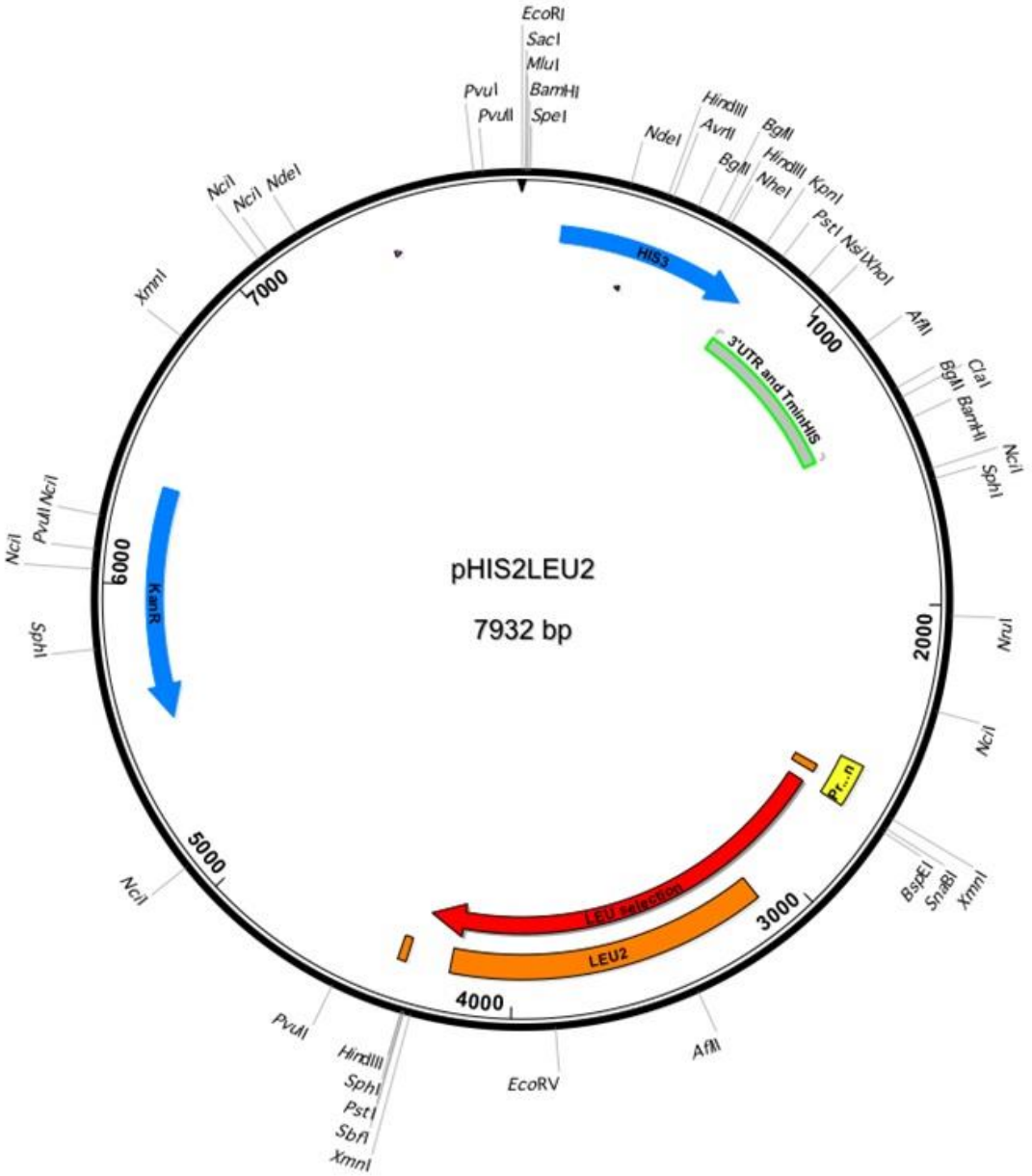
Appendix C:

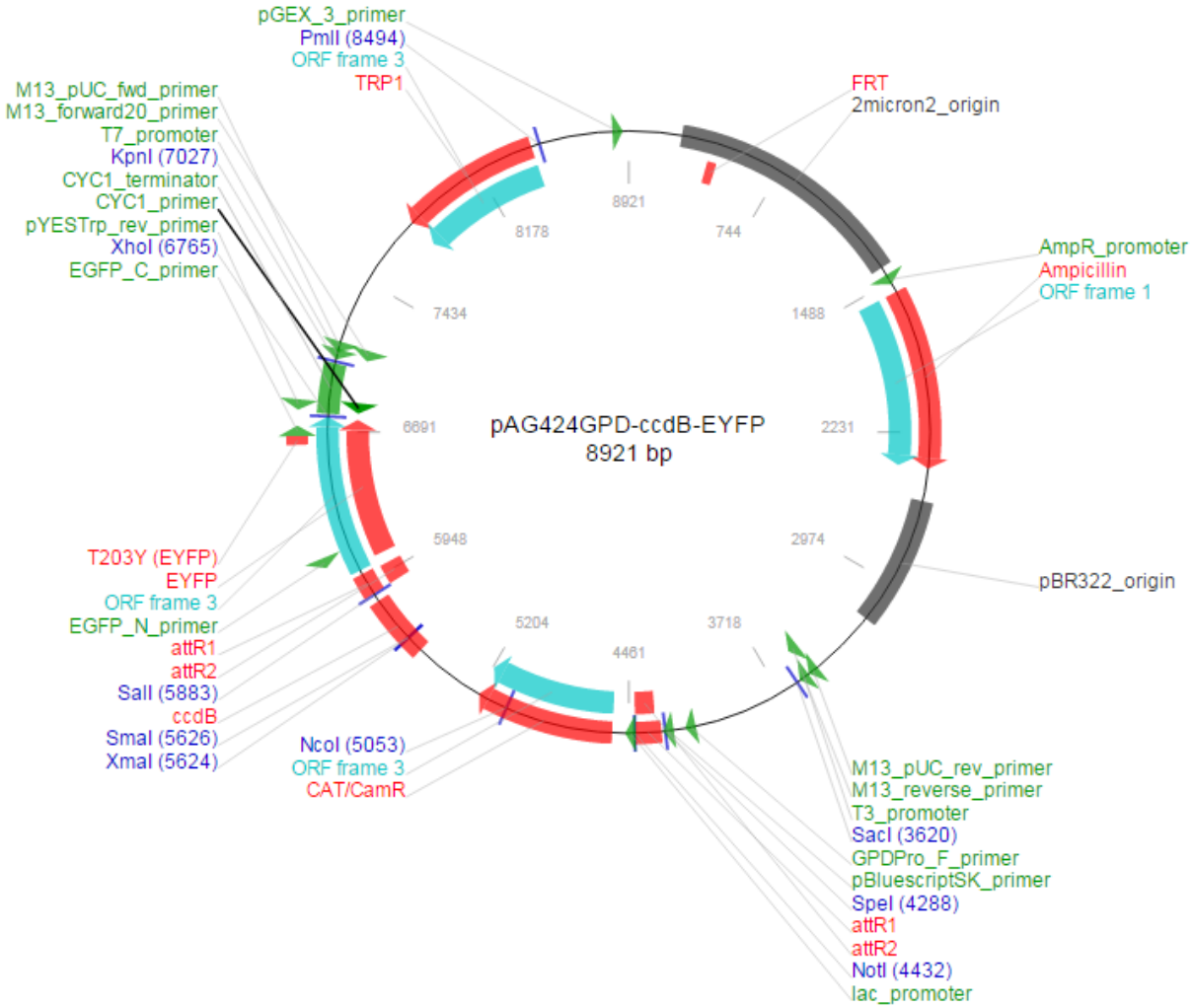
All the tables of differentially expressed genes in 35S-AtHSFA1b under no stress, wild type and 35S-AtHSFA1b::mRFP under heat stress are found in the following link

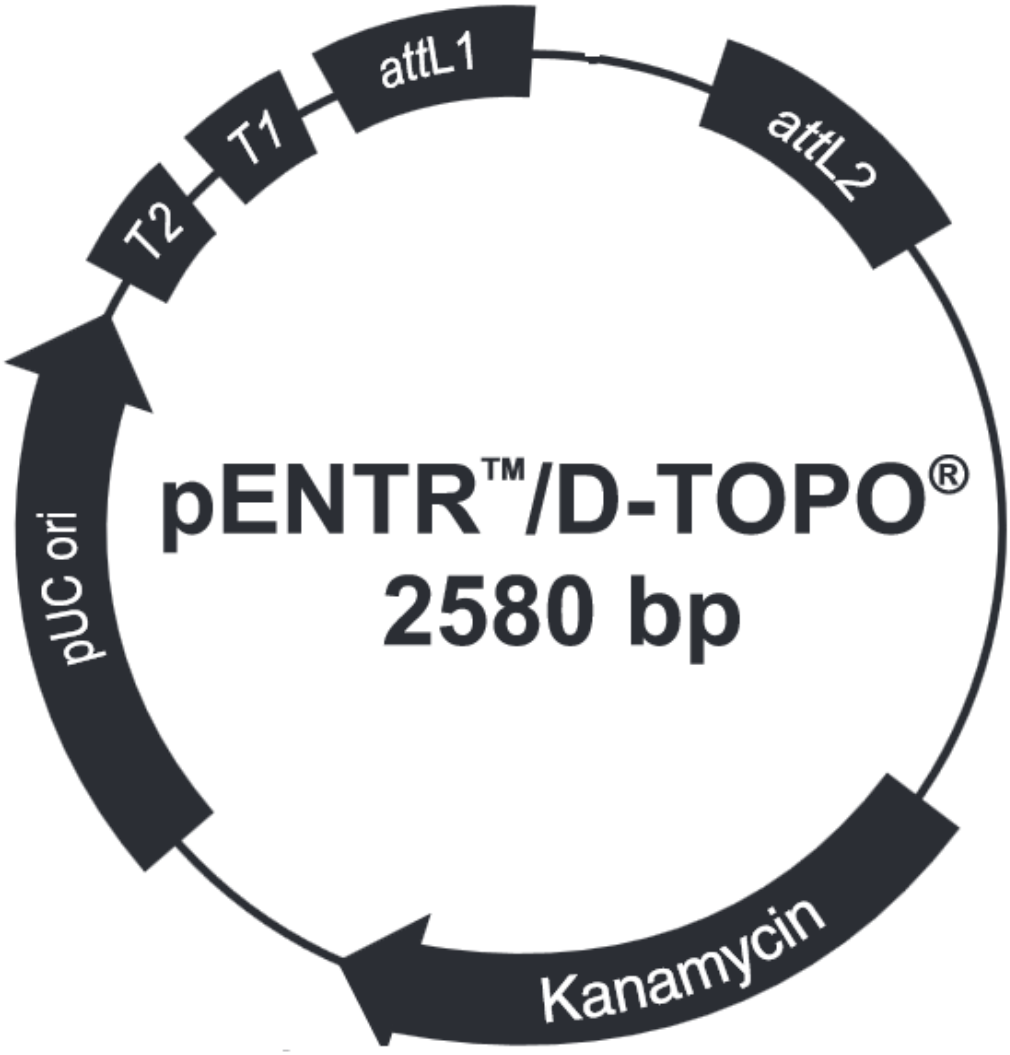
https://drive.google.com/folderview?id=0B4jdOv6BYryadGVYSkFIYk5nRmc&usp=sharing

Appendix D:

Maps of the plasmids used in this study:




pAG424GPD-ccdB-EYFP
8921 bp
pGEX_3_primer
PmlI (8494)
ORF frame 3
TRP1
FRT
2micron2_origin
M13_pUC_fwd_primer
M13_forward20_primer
T7_promoter
KpnI (7027)
CYC1_terminator
CYC1_primer
pYESTrp_rev_primer
XhoI (6765)
EGFP_C_primer
AmpR_promoter
Ampicillin
ORF frame 1
pBR322_origin
T203Y (EYFP)
EYFP
ORF frame 3
EGFP_N_primer
attR1
attR2
SalI (5883)
ccdB
SmaI (5626)
XmaI (5624)
NcoI (5053)
ORF frame 3
CAT/CamR
M13_pUC_rev_primer
M13_reverse_primer
T3_promoter
SacI (3620)
GPDPro_F_primer
pBluescriptSK_primer
SpeI (4288)
attR1
attR2
NotI (4432)
lac_promoter
8921
744
1488
2231
2974
3718
4461
5204
5948
6691
7434
8178


attL1
T1
T2
attL2
pENTR™/D-TOPO®
2580 bp
pUC ori
Kanamycin