Studies on an Arabidopsis MYB Transcription Factor Involved in Heat and Salt Tolerance

Scott Wilson Ramsay

College of Medical, Veterinary and Life Sciences
University of Glasgow

March 2014

Abstract

Heat stress has a significant impact on the productivity and yield of crops grown in the hot, arid zones of the world. There is mounting evidence that what has classically been termed 'drought stress' may in some cases be caused not by water stress *per se*, but rather by the uncontrolled elevation in leaf temperature that occurs when a plant loses its capacity for transpirational cooling. A previous screen of activation-tagged *Arabidopsis thaliana* seeds for novel halotolerant mutants implicated elevated levels of the transcription factor *MYB64* in mediating improved survival on high salt growth medium, and subsequent transcript profiling of this activation-tagged halotolerant line (HT5) revealed the upregulation of several members of the heat shock protein family.

Based on these preliminary findings, expression of two of the small heat shock proteins reported to be among the most highly upregulated in the HT5 line was investigated under various stress conditions in wild type Arabidopsis. Transcript and protein levels were measured in response to heat; their subcellular localisation was observed; and the phenotype of various knockout mutants was recorded. These studies have contributed to an understanding of how these might function in relation to one another and to the rest of the heat shock protein family.

This thesis also reports on the investigations of a transgenic line created to constitutively overexpress the *MYB64* transcription factor. Transcript profiling produced a list of 'upregulated' sequences, of which a significant proportion were previously shown to play key roles in abiotic (and, to an extent, biotic) stress responses. The robustness of these responses in the transgenic lines was investigated by qPCR under heat stress, and the phenotype of the plants was characterised in response to various stress regimes. The findings implicate *MYB64* in the regulation of a wide range of stress responses, and as plants are unlikely to encounter stress factors individually outside of the controlled conditions of a laboratory, these findings highlight the importance of considering such stresses in concert rather than isolation.

Author's Declaration

I declare that this thesis, submitted for the degree of Doctor of Philosophy, has been composed entirely by me and that the work presented herein was performed by me unless stated otherwise.

Acknowledgement

I would like to extend my eternal gratitude to my friends who have come and gone through the lab door in the time that I was lucky enough to have studied with them. Dr. Jillian Price, who taught me everything I needed to know; Jim Jardine, who taught me everything I wanted to know; Dr. Mai-Britt Jensen, my lab wife and partner in crime, Christine Merrick for being a superstar since iGEM back in 2007 and for proof-reading every single minute detail of this thesis; Dr. Louise Horsfall for all the coffee and our road trip across the western United States; Prof. Susan Rosser for letting me muscle in on all of her research group's outings; Tamar, Connie, Ema, Fabian, Karin, Simon, Dave, and especially Lynsey McLeay for reminding me that it's possible to have a normal life outside science (though it doesn't necessarily always have to be conventional…).

The staff at the Graduate and Undergraduate Schools of MVLS have been a reliable source of smiling faces and plenty of demonstrating & invigilating opportunities when I needed extra cash, so thanks to Leah Anderson, Lyndsay Ferry, Catherine & Joyce.

To all of the teaching staff and fellow demonstrators with whom I've had the pleasure to work – a big thank you for making it so much fun. Mary, Anne, and Chris in the Level 1 Biology Teaching Centre get a special mention, as do Zara, Daria and Elizabeth for helping to establish the Graduate Teaching Assistant Group.

Thanks to my parents, Moira and William, and my brother, Craig, for their encouragement and support – even if they have never been *quite* sure what I've been doing for these last four years.

Finally to my supervisor, Dr. Peter Dominy, for his unfailing support, advice, good humour, attention to detail (and for reminding me during the write-up process that plates do not, in fact, grow).

Table of Contents

List of Tables

List of Figures

List of Accompanying Material

Supplementary material supplied on DVD:

1. Supplementary File 1 – Transcriptome Analysis – Transgenic 35Spro:MYB64 Line 141 vs Wild Type

    a. ***Infinitely Upregulated*** Transcripts in 35Spro-MYB64 Line 141 vs Wild Type.csv
    b. ***Measurably Upregulated*** Transcripts in 35Spro-MYB64 Line 141 vs Wild Type.csv
    c. ***Transcripts With No Change*** in Regulation Between 35Spro-MYB64 Line 141 vs Wild Type.csv
    d. ***Measurably Downregulated*** Transcripts in 35Spro-MYB64 Line 141 vs Wild Type.csv
    e. ***Infinitely Downregulated*** Transcripts in 35Spro-MYB64 Line 141 vs Wild Type.csv
    f. ***Incalculable*** - sequences annotated at TAIR with coding regions but with reads aligned only with noncoding regions.csv
    g. ***Incalculable*** - Non-coding genomic sequences - no exons on which to calculate expression ratios.csv

    (accompanies section 4.3 - Transcript Profiling of 35Spro:MYB64 Line 141*)*

2. Supplementary File 2 - GO clustering - 100 most highly upregulated transcripts in 35spro-MYB64 line 141.xlsx

    (accompanies section 4.3 - Transcript Profiling of 35Spro:MYB64 Line 141*)*

3. Video Files 1, 2, 3, and 4.

    (accompanies section 3.3.6 - Sub-Cellular Localisation of Constitutively Expressed smHSPs; *for Video File legends, see section 3.3.6.5 on page 105)*

Definitions/Abbreviations

Ac-Tag	Activation tagging
ABA	Abscissic acid
ACD	α-crystallin domain
AGI	Arabidopsis Genome Initiative
ATP	Adenosine triphosphate
bHLH	Basic helix-loop-helix
CaMV	Cauliflower mosaic virus
cDNA	Complementary DNA
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dH_2O	Distilled water
gDNA	Genomic DNA
GFP	Green fluorescent protein
GO	Gene ontology
Gt	Gigatonnes
ER	Endoplasmic reticulum
HSF	Heat shock factor
HSP	Heat shock protein
HT5	Halotolerant line 5
MHa	Megahectares
MPa	Megapascals
mORF	Main ORF
mRNA	Messenger RNA
NASC	Nottingham Arabidopsis Stock Centre
nESI-MS	Nano-electrospray ionisation mass spectrometry
NGS	Next generation sequencing
ORF	Open reading frame
PCR	Polymerase chain reaction
qPCR	Quantitative RT-PCR
RFP	Red fluorescent protein
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPKM	Reads per kilobase, per million mapped reads

RT-PCR	Reverse transcriptase polymerase chain reaction
RuBisCO	Ribulose-1,5-bisphospate carboxylase/oxygenase
SDS-PAGE	Sodium dodecyl sulphate – polyacrylamide gel electrophoresis
SE	Standard error
smHSP	Small HSP
T-DNA	Transfer DNA
TAIL-PCR	Thermal asymmetric interlaced polymerase chain reaction
TAIR	The Arabidopsis Information Resource – www.arabidopsis.org
T_{air}	Air temperature
Ti plasmid	Tumour-inducing plasmid; contains T-DNA
T_K	Killing temperature
T_{leaf}	Leaf temperature
TSS	Transcriptional start site
uORF	Upstream ORF
USD	United States dollars
UTR	Untranslated region
YFP	Yellow fluorescent protein

1 Introduction

1.1 Crops in a Changing Climate: The Problem

Climate change is predicted to manifest in two main effects: an increase in global average temperature, and an alteration in the hydrological cycle. Temperatures will rise and cause increased evapotranspiration, and as this water returns precipitation rates will rise. As reviewed by Lobell & Gourdji (2012) this has the potential to affect crops in five ways. The first of these could be called the 'productivity race' in which high temperatures are known to accelerate development towards the seed-setting stage, and in doing so a plant produces fewer grains, thus reducing agricultural output (Stone, 2001). This can be understood in evolutionary terms as a trade-off between the goal of producing more seed in times of plenty and simply producing *enough* seed to ensure genetic continuity is assured in times of hardship. Second, higher temperatures and atmospheric carbon dioxide (CO_2) concentrations could prove favourable for many crop pathogens including weeds, insects and microbial pathogens (Ziska et al., 2011). Third, photosynthetic rate is negatively correlated with an increase in average growing temperatures. Even plants such as maize (a C4 plant, adapted to high temperature and its inactivating effect on RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase, a critical enzyme involved in the initial steps of carbon fixation from atmospheric CO_2)), suffer photosynthetic decline at leaf temperatures greater than 38 °C (Crafts-Brandner and Salvucci, 2002). Fourth, the overall effect of alterations in the pattern of rainfall and temperature will affect the water vapour pressure deficit between the atmosphere and leaves (Dai, 2011, Lobell et al., 2013b), and the rate at which

water is taken up by the roots and subsequently lost to transpiration will quicken. Loss of turgor pressure at a cellular level eventually leads to wilting of the entire plant, the natural response to which is closure of stomata to reduce water loss, but the knock-on effect of this is to reduce the influx of CO_2. When the increased internal ratio of O_2:CO_2 is paired with the temperature-dependent increase in RuBisCO photorespiration (Crafts-Brandner and Salvucci, 2002) it becomes clear how drought can lead to inhibition of photosynthesis and growth. This is an effect that leads (in the short term) to growth stasis, however, rather than lethality, and at normal temperatures plants can recover from wilting if irrigated. The fifth, related, effect of climate change is heat stress, which occurs when a plant is unable to counteract heating effects with sufficient transpirational cooling. Although the number of frost-free days in the growing season would increase, so would the number of days where the temperature reaches critical levels high enough to damage crops at a cellular level (Teixeira et al., 2013). Battisti and Naylor (2009) combined 23 different climate change models and predicted that average growing season temperatures in the tropics and the subtropics will be higher in 2050 than any maximum seasonal temperature recorded between 1900 and 2006. Gourdji et al. (2013) measured the percentage of farmland area subjected to five consecutive days of temperatures higher than the critical temperature during the reproductive days of the crop planted there; by 2050 those percentages are projected to increase from 5% to 18% for wheat, 8% to 27% for rice, and 15% to 44% for maize. Without effective management this represents a major threat to global food security.

1.1.1 Changes in Productivity and Population

It has been projected by the International Food Production Research Institute (IFPRI) that, without intervention, the global production of wheat in the year 2050 will have declined from 2000 levels by 27.4%, rice by 13.5%, and maize by 0.4% (focussing only on developing countries the equivalent figures are 33.5%, 13.6%, and 10%) (Nelson et al., 2009, Table 1). Following this worst-case scenario, daily per-capita calorie availability is expected to drop by 7.6% in developed countries and by 10.7% in developing countries (Nelson et al., 2009, Table 5). In the same period the world's population is projected to reach 9 billion. Accordingly, food prices will rise consistent with the principles of falling supply and rising demand (Nelson et al., 2009, Table 2). In short, from a figure of 881 million in 2005, the number of people at risk of food shortage in developing nations is projected to rise to more than 1 billion (Rosegrant et al., 2014).

However, with the instigation of a programme of 'adaptive investments' (*i.e.* those designed to allow the population to cope with such a change, as opposed to 'mitigation investments', which would reduce the extent of the change (United Nations Environmental Protection Agency; UNEP, 2014)) the best-case scenario for developing countries projected by the IFPRI forecasts a 5% *increase* in calorie availability (Nelson et al., 2009, Table 8). Such adaptive investments range from improvements to infrastructure such as roads and irrigation mechanisms, to low-tech local farming solutions such as the expansion of fertiliser use, to research & development (R&D) type solutions such as selective breeding or genetic engineering (Rosegrant et al., 2014).

Lobell et al. (2013a) performed a cost/benefit analysis of adaptation to climate change as opposed to attempting to mitigate the problem. They predict that an investment of 225bn USD (at year 2000 value) by 2050 would prevent the clearing and conversion of 61 megahectares (MHa) of land for arable use (261 MHa converted instead of 322 MHa), and in doing so would reduce the release of greenhouse gases over that period from 87 to 72 gigatonnes (Gt) of CO_2 equivalents; a saving of 15 Gt, or 17 %.

1.1.2 Management Strategies

The agricultural literature relating to climate change focuses heavily on strategies of adaptation rather than mitigation, the latter perhaps being seen as the domain of economic policy makers and energy industries, and of those involved in the emerging field of geoengineering. Solar radiation management, for example, would involve the dissemination of reflective particles into various levels of the atmosphere, and one of the current challenges is to design a system capable of managing the effects on a regional rather than a global scale. However, a synchronised planetary perturbation of temperature and precipitation levels would endanger food security just as much as current projections of unmanaged climate change (Pongratz et al., 2012). Historical evidence suggests that even while we have been aware of the threat of climate change, and while nations have signed accords to reduce the anthropogenic effects, mitigation strategies have simply not been effective. In the United States, one of the largest contributors to greenhouse gas emissions, CO_2 output increased 10% in the two decades to 2011 (United Nations Environmental Protection Agency, n.d.) and are projected to rise by a further 5-11% between

2011 and 2030 (depending on the effectiveness of carbon sequestration policies) (U.S. Dept. of State, 2014).

With the overwhelming impact of climate change expected to be higher temperatures and the related dehydration and salinity effects, perhaps it is necessary for plant biologists to change from the historical type of thinking that characterises heat and drought stresses as distinct, and instead to characterise the responses to these in terms of their interrelated effectiveness at dealing with heat *and* dehydration. Ultimately the principal damage caused by drought in a changing climate would be elevation of leaf temperatures to unacceptable levels, causing the loss of cellular and sub-cellular integrity before water shortage *per se* causes an irreversible effect, and the combined message from a number of projections on climate change over the last decade suggest that heat stress (and related) responses should be one of the most urgent focuses of plant biotechnology in the coming years (Battisti and Naylor, 2009, Ciais et al., 2005, Lobell et al., 2012, Lobell et al., 2013b).

1.2 Crop Responses to Climate Change-Related Stresses: Prospects for Adaptation

If we are to move towards an integrated model of the stress responses elicited by a changing climate, it is important to first of all recognise what has been discovered about the individual components thus far.

1.2.1 Heat Stress Responses

As described above, one response to elevated temperatures is the rapid acceleration of development towards the seed-setting stage (the 'productivity race'). Kumar et al. (2012) describe the involvement of the bHLH transcription factor PIF4 (phytochrome interacting factor 4) in accelerated flowering in response to elevated temperatures. PIF4 activates expression of Flowering Locus T (FT) at 27 °C in short day conditions, demonstrating that this mechanism is independent of photoperiod and thus crops may be at risk of engaging in this productivity race even before the seasonal transition from short to long days.

Heat shock proteins (HSPs) are a ubiquitous class of proteins that help cells cope with the physical stresses of increasing temperature, drought and salinity (Waters et al., 1996) and have been reported in, and have homology across, bacteria, yeast, and higher eukaryotes (Kotak et al., 2007a, Vierling, 1991). HSPs can be broadly distinguished into two subfamilies: those which have molecular masses grouped around 70 kDa, 90 kDa and 100 kDa; and small heat shock proteins (smHSPs) which have molecular masses of ≤30 kDa.

In most organisms, the larger HSPs are the main effectors of thermotolerance and this is reflected in both their abundance at the protein level and the number of genes (Lindquist, 1986). They are ubiquitously expressed upon heat shock and serve to chaperone denatured proteins back into a re-folded conformation. Homologs in bacterial systems include ClpB (homolog of plant HSP100), HtpG (HSP90), DnaK (HSP70) and GroEL (HSP60). Small HSPs are comparatively over-represented in plants and have a different mode of action; their role is to interact with and protect native proteins in danger of denaturing as a result of

the increased kinetic energy conferred by a heat stress (Waters, 2013). Transgenic studies have repeatedly shown that overexpression of various plant and fungal smHSPs and HSPs has improved the thermotolerance of the host plants at critical temperatures (Cho and Choi, 2009, Sun et al., 2012, Katiyar-Agarwal et al., 2003, Kim et al., 2012, Liu et al., 2009, Neta-Sharir et al., 2005, Ogawa et al., 2007, Ono et al., 2001, Qi et al., 2011, Queitsch et al., 2000, Uchida et al., 2008, Xin et al., 2010, Xue et al., 2010, Yoshida et al., 2008, Zhou et al., 2012, Zhu et al., 2009). Notably, overexpression of HSP70 from *Trichoderma harzianum* (a fungus) in Arabidopsis caused an upregulation of *SOS1* (salt overly sensitive 1; a Na+/H+ antiporter (Wu et al., 1996) – see section 1.2.3 "Salt Stress Responses" for more on this gene) and *APX1* (ascorbate peroxidase 1; a heat-shock-activated scavenger of reactive oxygen species (ROS) (Panchuk et al., 2002)), and overexpression of HSP70 and of Arabidopsis HSP26 both independently caused a *downregulation* of HSFs and other HSPs (Montero-Barrientos et al., 2010, Rhoads et al., 2005). This demonstrates a mechanism of feedback between an HSP and the regulatory network which directs its expression as well as the expression of other effector components.

The main activators of the smHSP response are termed heat shock factors (HSFs), of which there are 21 in Arabidopsis (Kotak et al., 2007a). They share a common basic structure of an N-terminal DNA binding domain, a hydrophobic multimerisation domain, a nuclear localisation signal (NLS) and an activation domain (Kotak et al., 2007b), and are separated into 3 classes (A, B, and C) on the basis of structural differences. Knockout studies have implicated HSFA1a and HSFA1b in the initial phase of the heat-stress response, and HSFA2 in the recovery phase or in responses to extended periods of high temperature (Lohmann et al., 2004, Busch et al., 2005, Schramm et al., 2006).

The HSF response also appears to be activated by stimuli other than heat. In one study by (Nishizawa et al., 2006) expression of HSFA2 was inducible by H_2O_2 and high light, implying a role in oxidative stress responses. In recent years, several studies have reported the involvement of oxidative stress in abiotic stress tolerances (Huang et al., 2012, Atkin and Macherel, 2009) and levels of H_2O_2 have indeed been shown to rapidly accumulate shortly after a period of high temperature (Vacca et al., 2004). It has also been noted that HSFA3 is inducible by DREB2A, one of the classic signalling intermediates of drought stress responses (Sakuma et al., 2006, Yoshida et al., 2008). Overexpression of HSFA1b has also been shown to improve drought tolerance in Arabidopsis by preventing wilting and decreasing the rate of water loss of detached rosettes, and also to improve pathogen resistance via an alteration in the transcription levels of several hundred genes (Bechtold et al., 2013).

At least one HSF has been demonstrated to mediate some exclusively developmental processes despite belonging to the HSF class. HSP17.6 and HSP17.6a have been demonstrated to be inducible, at least under experimentally-induced conditions, by HSFA9 (Kotak et al., 2007b), a seed-specific transcription factor which is in turn under the control of ABI3 (abscissic acid (ABA) insensitive), a transcription factor which responds to ABA. Constitutive expression of HSFA9 from the 35S promoter increased levels of HSP17.6 and HSP17.6a. HSFA9 under the control of its own promoter was inducible upon application of ABA, and yet there was no evidence of HSFA9 expression under heat-shock, nor does knocking out ABI3 have any effect on thermotolerance. It appears, then, that the HSPs may also play an important role in maturing seeds and it can be hypothesised that their role in this situation is to protect proteins from loss of conformation during desiccation, a process by

which solutes would become more concentrated and thus ionic strength would be perturbed.

1.2.2 Hydration Stress Responses

As already discussed, the distinction between drought, salinity and heat stress responses is not as clear-cut as some of the literature would suggest. This is particularly evident in the range of responses noted to confer an advantage upon naturally drought-tolerant and naturally salt-tolerant crops.

Such responses to drought may include overt phenotypes like growth arrest; this is advantageous as the cessation of growth means that any metabolic activity can be focussed on avoiding or ameliorating the effects of the stress. One example is the inhibition of lateral root development observed in Arabidopsis in favour of extension of the main tap root, which would be able to extract water from deeper in the growth substrate (Xiong et al., 2006). Cessation of growth is perhaps a normal and active response to drought stress, but it is also partly imposed by necessity when stomatal closure occurs in order to prevent further loss of water. Stomatal closure is partly dependent on ABA signalling and partly dependent on turgor pressure, and it represents a temporary measure, reversible when the hormone signal is relieved or the water status of the plant improves (Chaves et al., 2009, Wilkinson and Davies, 2010).

On a molecular level, with the loss of transpirational pull through the xylem and a state of early dehydration, cells can act to increase turgor pressure by increasing the osmotic potential of the vacuole to draw in water from the roots (for more detail see section 1.2.3). In order to maintain osmotic potential balance between the vacuole and the cytoplasm, the plant also produces

compatible solutes (*i.e.* solutes that are not toxic to metabolism) using carbon which is now available, having been diverted from overall growth mechanisms (Blum, 2005, DaCosta and Huang, 2006, Lei et al., 2006, Xue et al., 2008). Carbon fixation mechanisms can also be enhanced to make more efficient use of the reduced CO_2 levels in the leaf. In one crop-based study, a naturally drought-resistant variety of Sorghum (*Sorghum bicolor* (L) Moensch) was shown to have increased levels of RuBisCO after drought stress, accompanied by an increase in the level of HSP60 (Jagtap et al., 1998), demonstrating the importance of molecular chaperones in a variety of stress conditions that threaten protein structure (in this case, solute concentration and thus ionic strength).

1.2.3 Salt Stress Responses

Soil salinisation is a problem for agriculture and is associated with several human activities, the main one of which is irrigation. Fifty-five percent of all the food produced globally is estimated to come from irrigated land (Food and Agriculture Organisation of the United Nations; FAO, 2013). As the water applied to a soil is lost by evapotranspiration, it leaves behind solutes which accumulate over time. While farmers might know this, shortages of water in areas where subsistence farming is practiced might mean there is no choice in the short term but to irrigate with brackish water. This is relatively common in, for example, Bangladesh, China, Egypt, India, Iran, Pakistan, Syria, Spain and parts of the United States (FAO, 2014). Forestry clearing is another human activity that impacts indirectly on salinity levels. When deep-rooted trees are removed and replaced with short-rooted crops, the transpiration activity that would have normally drawn up water from deep water tables via long taproots is lost.

Without this water sink the water table begins to rise, bringing with it the same solutes that lead to increased salinisation.

It is important to note here that seawater has a concentration of approximately 500 mM Na^+ and 550 mM Cl^-. Arabidopsis is a glycophyte, and as such is able to complete a life cycle at salt concentrations up to approximately 100 mM (Flowers, 2004). It is the Na^+ rather than the Cl^- ions that are responsible for the toxicity, as evidenced by tolerance to equivalent concentrations of KCl (Wu et al., 1996). Plants deal with this Na^+ toxicity in one of three main ways: efflux, sequestration, or exclusion. The main structural components of these strategies are membrane-bound proteins which act as channels through which excess Na^+ ions are shuttled out of the cytoplasm.

The classical Na^+ efflux pathway is the salt overly-sensitive (SOS) pathway comprising three parts. SOS3 is a Ca^{2+}-binding protein, activated by a calcium wave initiated when intracellular Na^+ levels reach a critical threshold (Liu and Zhu, 1998). The Ca^{2+}-bound form then interacts with SOS2, a serine/threonine kinase (Liu et al., 2000) which becomes myristoylated and translocates to the plasma membrane where it phosphorylates SOS1 (Ishitani et al., 2000), a Na^+/H^+ antiporter (Shi et al., 2000). This has the effect of shuttling excess Na^+ out of the cytoplasm into the apoplast. Another antiporter, HKT1 (H^+/K^+ Transporter 1), operates in shoots in a similar way to SOS1 by depositing excess Na^+ into the phloem where it is recirculated to the roots (Berthomieu et al., 2003).

There are at least 27 Na^+/H^+ antiporters in the Arabidopsis genome (Ward, 2001) and they do not all perform efflux at the plasma membrane. Others include NHX1 (Na^+/H^+ exchanger 1), a tonoplast antiporter, which sequesters excess Na^+ in the vacuole (Gaxiola et al., 1999). The primary reduction in cytoplasmic salt

concentration is an obvious advantage, but a secondary benefit is conferred by increasing vacuolar osmotic potential. This encourages water to enter, increasing the vacuolar volume and thus increasing turgor of the plant, without which the stomata would close and thus CO_2 assimilation rates would drop.

Homologs of all these types of Arabidopsis membrane channels have been identified in the major crops and research is underway to capitalise on their functions in order to improve halotolerance in the field (for a review, see Zhang and Shi (2013)).

1.3 Aims of This Research

This thesis focuses primarily on investigation of thermotolerance phenotypes of Arabidopsis mutants, touches on other related abiotic stress phenotypes, and ultimately leads to a discussion of cross-talk between abiotic stress tolerances more generally, and it is based on a previous research project aimed at finding novel regulators of halotolerance. That previous research (Price, 2005) led to the identification of a series of distinct and previously uncharacterised mutants, one of which was shown to have altered levels of the transcription factor *MYB64*. A microarray experiment performed on mature, long-term salt-stressed Arabidopsis plants comparing the *MYB64* overexpressing line with wild type revealed an upregulation of stress response transcripts, the highest of which encoded several members of the small heat shock protein family. An understanding of that previous characterisation underpins any understanding of the work contributing to this thesis, so it will be briefly summarised here.

1.4 Project History – *MYB64*: A MYB Transcription Factor Conferring Salt Tolerance

1.4.1 Mutant Search Strategy: Activation Tagging

Price aimed to identify and characterise novel mutations that led to an increase in survival rates among Arabidopsis seedlings grown in the presence of toxic levels of NaCl. This was done by obtaining a large collection of seed from parents that had been subjected to *Agrobacterium*-mediated insertional mutagenesis (Weigel et al., 2000) and sowing this seed on a screening medium – in this case, agar supplemented with 80 mM NaCl. Those seedlings that survived the screening were transferred to soil and grown for further characterisation.

One particular halotolerant mutant, sequentially designated Halotolerant 5 (HT5, Figure 1.1), carried a T-DNA insertion in the intergenic region between genes At5g11040, encoding an uncharacterised protein, and At5g11050, encoding the transcription factor MYB64. It has been shown experimentally that the 35S enhancer elements in this insertion are able to affect transcription patterns from insertion loci up to 3.6 kB upstream from the transcriptional start site (TSS) of a gene, as well as from loci downstream of the transcriptional stop site (Weigel et al., 2000) so both genes were considered to be candidates for altered expression mediated by the activation tag and leading to the phenotype. Semi-quantitative RT-PCR showed no difference in the levels of At5g11040 between wild type and the HT5 insertional mutant, but did show a clear increase in the level of At5g11050 transcript (Figure 1.1, **panel D**).

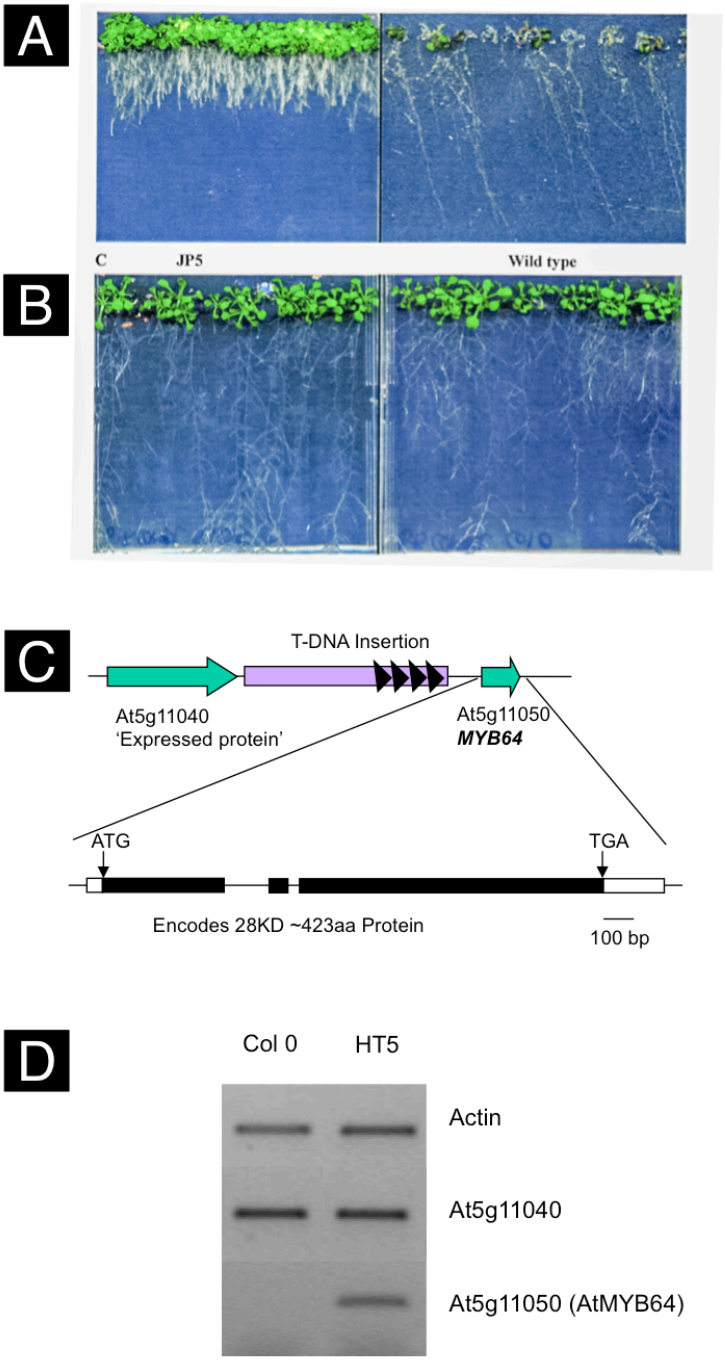


Figure 1.1 Characterisation of Halotolerant Activation-Tagged Mutant HT5

A HT5 (left) and background wild type line (Col 0; right) were grown on agar plates supplemented with NaCl to a final concentration of 80 mM. HT5 seedlings are more halotolerant and have a short-root phenotype. **B** There is no observable difference between growth of HT5 and wild type in the absence of NaCl. **C** TAIL-PCR was used to identify the genomic location of the T-DNA insertion in line HT5 as the intergenic region between At5g11040 and At5g11050. The gene annotations available from The Arabidopsis Information Resource (Swarbreck et al., 2008) identify these as 'Expressed protein' and *MYB64*, respectively. **D** RT-PCR was carried out to measure levels of At5g11040 and At5g11050 transcripts in the halotolerant activation-tagged line HT5, and on the background line Col 0 as a wild type control. Actin was used as an internal loading control.

(Figure adapted from Price, 2005)

1.4.2 MYB64 Functional Characterisation

MYB64 was largely uncharacterised in the literature. It has since been implicated in embryo sac development by whole-transcriptome expression profiling experiments (Johnston et al., 2007, Wuest et al., 2010) though not by functional studies at a gene product level. Very recently it has been implicated in the cellular development of the female gametophyte in Arabidopsis, working redundantly with *MYB119* (Rabiger and Drews, 2013). Since it is a transcription factor, it was of interest to determine which genes might be under its regulatory control and might therefore be responsible for effecting the halotolerant phenotype. A microarray experiment was conducted to compare the global transcript profile of the HT5 mutant with that of wild type, both grown with a long-term salt stress to replicate the conditions under which the phenotype was first identified (80 mM NaCl). The most 45 highly upregulated genes in line HT5 are presented in Table 1.1. It was perhaps surprising that the transcript with the highest increase in measured abundance encoded a smHSP gene rather than one associated explicitly with halotolerance. Several other genes close to the top of the list also encode members of the smHSP and the larger HSP70/HSP101 families. In addition, several genes with unknown functions at that time have since been updated in the database at TAIR with annotations describing a variety of stress-response functions (not shown).

AGI	Name	Avg FC	RP UP	FDR (%)
At1g53540	17.6 kDa heat shock protein (HSP17.6b)	5.08	12	1.44E-06
At1g74310	heat shock protein 101 (HSP101)	3.97	36	1.44E-06
At3g12580	Heat shock protein 70 (HSP70)	4.14	51	1.53E-06
At5g12030	17.6 kDa heat shock protein (HSP17.6a)	3.83	66	1.59E-06
At1g51420	unknown protein (putative sucrose phosphatase)	3.34	172	3.45E-06
At4g12470	pEARLI 1-like protein / Azelaic Acid Induced 1 (SA response)	2.78	209	3.59E-06
At5g48570	peptidylprolyl isomerase	2.43	476	7.17E-06
At4g12490	pEARLI 1-like protein / Defense response to fungus	2.38	528	7.07E-06
At1g26680	hypothetical protein	2.31	572	6.89E-06
At1g61800	g1ucose-6-phosphate/phosphate translocator precursor	2.41	611	6.13E-06
At5g10840	endomembrane protein 70, putative	2.81	728	6.75E-06
At1g36060	AP2 domain transcription factor RAP2, putative / member of DREB subfamily A-6 of ERF/AP2 TF family	2.60	980	8.43E-06
At1g76900	Tub family protein, putative	2.67	1088	8.74E-06
At3g50970	dehydrin Xero2 (low temp induced protein)	2.20	1104	8.31E-06
At4g12480	pEARLI 1 / Cold induced	2.42	1116	7.91E-06
At1g56600	water stress-induced protein, putative / Galactinol synthase – water and salinity stress	2.16	1140	7.63E-06
At2g05510	putative glycine-rich protein	2.20	1440	9.13E-06
At5g24110	WRKY family transcription factor (WRKY30) / response to SA	2.07	1508	9.09E-06
At3g47420	glycerol-3-phosphate transporter, putative	2.12	1624	9.32E-06
At3g61830	auxin response factor-like protein	2.22	1785	9.78E-06
At4g12500	pEARLI 1-like protein	2.02	1848	9.28E-06
At3g25830	myrcene/ocimene synthase, putative	2.12	2640	1.27E-05
At1g27730	salt-tolerance zinc finger protein (ZAT10)	2.00	2646	1.22E-05
At3g16460	putative lectin	2.03	2688	1.20E-05
At1g59860	17.6 kDa heat shock protein (HSPI 7.6A-CI)	2.11	2821	1.21E-05
At3g10020	expressed protein	2.06	2884	1.19E-05
At5g66690	UTP-glucose glucosyltransferase	2.03	3034	1.21E-05
At3g09260	glycosyl hydrolase family 1	2.02	3132	1.21E-05
At5g26280	low similarity to ubiquitin-specific Arabidopsis protease 12	1.95	3151	1.18E-05
At3g17790	acid phosphatase type 5	2.06	3375	1.23E-05
At1g07400	17.8 kDa class I heat shock protein (HSPI 7.8-CI)	2.08	4200	1.48E-05
At4g22610	putative protein	2.07	4294	1.47E-05
At1g17710	expressed protein	1.90	4400	1.47E-05
At1g45210	expressed protein	2.13	4755	1.54E-05
At5g06320	hairpin-induced protein-like	1.85	5109	1.62E-05
At2g11810	putative monogalactosyldiacylglycerol synthase	1.86	5239	1.61E-05
At1g01470	late embryogenesis abundant protein, putative	1.85	5400	1.62E-05
At2g03760	putative steroid sulfotransferase	1.85	5430	1.59E-05
At1g17330	hypothetical protein	1.89	6440	1.84E-05
At4g30280	xyloglucan endotransglycosylase, putative	1.95	6650	1.86E-05
At3g46230	17.4 kDa class I heat shock protein (HSPI7.4-C1)	1.83	6674	1.82E-05
At2g46240	hypothetical protein	1.93	6942	1.85E-05
At3g25250	kinase	1.83	7138	1.87E-05
At5g01220	putative protein	1.84	7200	1.84E-05
At5g52060	putative protein contains BAG domain	1.94	9362	2.35E-05

Table 1.1 HT5 Microarray Analysis – Most Highly Upregulated Transcripts

Wild type (Col 0) and HT5 plants were grown for 2 weeks on agar plates with MS media (modified to contain 530μM Ca^{2+} and 200μM K^{+}) supplemented with 80 mM NaCl. Whole-seedling tissue was harvested, RNA was extracted and microarray analysis performed to compare expression levels. The most highly upregulated transcripts are presented here. AGI = Arabidopsis Genome Initiative gene number. Avg FC = Average fold change. RP UP = Rank Product (Upregulated). FDR = False Discovery Rate.

1.4.3 *MYB64* Transgenic Overexpression Line

The activation-tagging technique used to create the HT5 mutant, in which the tetramerised enhancer elements are inserted at random, did not necessarily lead to maximal overexpression of *MYB64*. To confirm the connection between *MYB64* and any phenotype, a stronger overexpressing line was created by directly fusing the 35S promoter to the transcriptional start site of the gene. *MYB64* genomic DNA was cloned into the T-DNA region of Ti plasmid pMN19 directly 3' to the tetramerised 35S enhancer elements (Weigel et al., 2000), then wild type Arabidopsis (Col 0) was transformed, selected on appropriate media, and insertion was confirmed by PCR on genomic DNA using primers designed to detect the transgene. These transformants (referred to hereafter as 35Spro:MYB64 lines) were shown to express *MYB64* at significantly higher levels than wild type.

These lines were used for all subsequent investigations of *MYB64* expression described in this thesis.

1.5 Project History – The Small Heat Shock Proteins

The discovery that 15% of the 45 most highly upregulated transcripts in line HT5 belong to the heat shock protein gene families made thermotolerance an appealing avenue of investigation. If *MYB64* activates the expression of smHSPs, it would be important to gain an understanding of how they function before any phenotypic characterisation was undertaken.

1.5.1 smHSPs in Plants

Small heat shock proteins are important components of the heat stress response in every organism where they have yet been investigated, whether archaea, bacteria, or eukarya (Waters and Rioflorido, 2007, Waters, 2013). They function as ATP-independent chaperones to stop denaturation caused by molecular vibrations resulting from an increase in kinetic energy, and thus prevent any subsequent irreversible polypeptide aggregation. The larger HSPs (which are evolutionarily unrelated and are named based on size: HSP100, HSP90, HSP70, HSP60 families), in contrast, are ATP-dependent and their function is to re-fold proteins which have already denatured (Haslbeck et al., 2005). As those are unrelated to the smHSPs by sequence and as they operate in a fundamentally different way, the focus of this introduction will remain solely with the smHSPs.

The Arabidopsis genome database currently holds records for 21 smHSPs and 3 'smHSP-like proteins' (Kotak et al., 2007b). These are classified into 11 subfamilies based on sub-cellular localisation, 6 of which are thought to be cytosolic and 5 of which are localised to organelles (Scharf et al., 2001, Waters, 2013). They are all measurably expressed, suggesting that none are pseudogenes, while at the same time there appears to be a difficulty in ascribing individual chaperoning roles as there have been no publications documenting differences in chaperone clients to date. The large number of smHSPs and high degree of sequence similarity suggests there might be a high degree of functional redundancy. On the other hand, the fact that the family is so large, and that this is true across plant species as divergent as Arabidopsis and rice, reflects distinct functional roles and suggests conservation of all of them (for comparison, the human genome only encodes 10 smHSPs (Arrigo, 2013) while

rice (*Oryza sativa*) has 23 and poplar (*Populus trichocarpa*) has 36 (Waters et al., 2008)). One aim of this thesis, therefore, was to investigate the functions and expression patterns of some of the smHSPs found to be upregulated in the HT5 line and to determine if they perform different roles.

1.5.1.1 smHSP Sequences, Structures and Functions

The 21 smHSPs in *Arabidopsis thaliana* are, as already mentioned, classed based on their phylogenetic relationships and on their putative sub-cellular localisations (Scharf et al., 2001). Classes CI - CVI are thought to be localised in the cytosol. The two Class M subfamilies are associated with the mitochondrion (Lenne and Douce, 1994), Class CP with the chloroplast (Osteryoung and Vierling, 1994), Class E with endomembranous structures such as the endoplasmic reticulum (ER), and Class P with the peroxisome (Helm et al., 1995, Waters, 2013).

The sequences of smHSPs show a high degree of conservation both between and within genomes. At the heart of their classification as smHSPs is the 'α-crystallin domain' (ACD) found towards the C- terminal end. This domain is also found in proteins of the vertebrate eye lens whose function is to prevent aggregation of denatured polypeptides that might lead to the formation of cataracts (Clark et al., 2012). This duality of expression in different tissues in different organisms where denatured proteins would be problematic supported the development of the theory that this conserved domain contributes to a chaperone activity.

Early studies on the structure of smHSPs includes the work performed by Kim *et al* (1998) on the single smHSP from the archaeon *Methanococcus jannaschii* (Figure 1.2, panel **A**). Shaped into a hollow 24-mer octahedral shell, the C-

termini lie on the surface and make hydrophobic and backbone contacts with the ß-strands of neighbouring subunits. Only half of the N-termini are resolvable, and they are present in the centre of the sphere and interact with each other in a similar fashion to linking arms. Kim *et al*. suggested that the large percentage of surface contacts given to the dimer interface (48% for this one interaction) make the dimer likely to be the most stable, long-lived structure.

The only smHSP from the plant kingdom currently to have had a crystallographic structure published is the 16.9 kDa protein from wheat (*Triticum aestivum;* TaHSP16.9). It forms a dodecameric double-doughnut 'stack' with one 6-mer ring laid directly on top of the other, forming a 3-fold crystallographic axis through the hole and a 2-fold axis at 90° to this (Figure 1.2, panel **B**) (van Montfort et al., 2001). An examination of the N-terminal arm reveals several hydrophobic contacts between both helical domains and α-crystallin domain ß-strands of neighbouring subunits. The subunit compositions of several other Arabidopsis smHSPs have also been reported as dodecameric (Painter et al., 2008). Using nano-electrospray ionisation mass spectrometry (nESI-MS) to investigate the various quaternary organisations adopted by purified Arabidopsis HSP18.1-CI and HSP17.6-CI (two proteins closely related to those which are the subject of this thesis, but belonging to a separate class) Painter *et al*. demonstrated that the building block of those dodecamers is indeed the dimer. They report masses of 216,301 kDa and 210,258 kDa respectively, corresponding to the predicted masses of dodecamers, plus very small peaks corresponding to solution-phase dimers which had not yet been assimilated into oligomers. Although not proven experimentally, a model of the Arabidopsis chloroplast

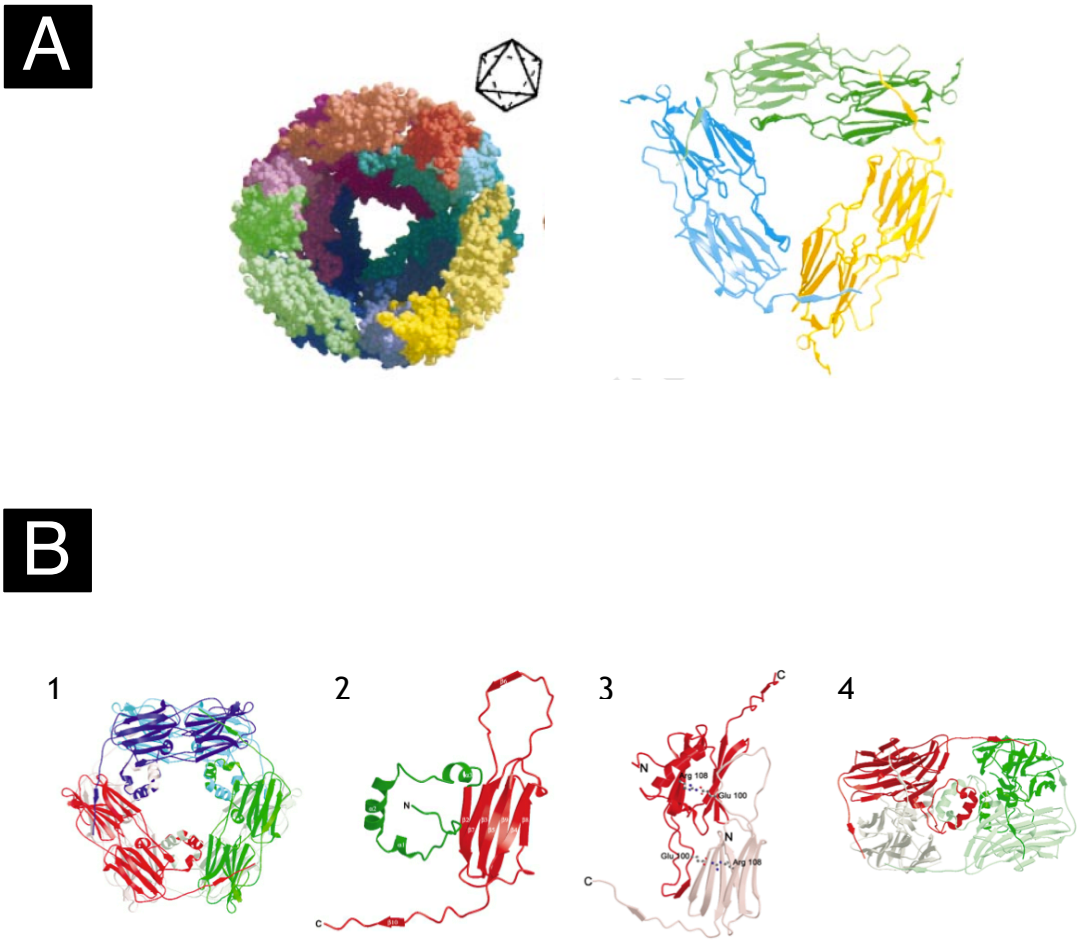


Figure 1.2 Crystal Structures of Two Previously Characterised smHSPs

A (Left) The single smHSP (16.5 kDa) from the archaeon *Methanococcus janaschii* forms a hollow, spherical 24-mer *in vitro*. (Right) An isolated view of three dimers within the sphere reveals C-terminal interactions between dimers.

(Adapted from Kim et al., 1998)

B 1: 16.9 kDa small heat shock protein from wheat (*Triticum* aestivum, TaHSP16.9) forms dodecameric double-doughnut structures, *i.e.* one doughnut on top of the other *in vitro*. 2: Monomers have a C-terminal extension which protrudes from the body of the subunit. 3: The basic unit of the dodecamer is the dimer. 4: When viewed in cross-section, the C-terminal extensions can be seen to facilitate dimer-dimer interactions as in the *M. janaschii* homolog, both within a ring, and between rings (four dimers shown, each dimer coloured differently).

(Adapted from van Montfort et al., 2001)

family member HSP21 also predicts a similar dodecamer of two doughnut-shaped discs (Lambert et al., 2011).

Painter *et al.* also explored what happens when HSP18.1-CI and HSP17.6-CI are mixed *in vitro.* The detail of the mass spectra for each of the two proteins reveal several peaks each representing multiple charge states. Treating these as fingerprints for each species, a simulated 'mixture' fingerprint can be obtained by overlaying the two, however when the purified proteins are mixed at a 1:1 ratio they give rise to a spectrum with an intermediate pattern with peaks corresponding to mass/charge ratios that could only be explained by the exchange of dimers to form hetero-dodecamers (see Figure 1.3). Following this equilibration in real-time it became evident that this occurs on a timescale of minutes and does so more rapidly with increasing temperature. This further supported the notion that the dimer is the building block of the dodecamer and also possibly the functional form of the protein. Similar results were obtained in other *in vitro* studies of quaternary dynamics (Baldwin et al., 2011a, Stengel et al., 2010)

It is interesting to speculate on the significance of this promiscuity between members of the smHSP family. If these proteins, which belong to the same class and are highly homologous, readily exchange subunits *in vitro,* does this imply that the same would happen *in vivo?* If so, this would seem to contradict the hypothesis supporting the existence of, and implied evolutionary selection for, so many smHSPs in plants; that each must have a distinct target range, temporal expression pattern, or tissue-specific pattern. Why is the family so large and distinct if the proteins so easily form hetero-dodecamers? Sobott et al. (2002) provide evidence that promiscuity is even wider – they report subunit exchange

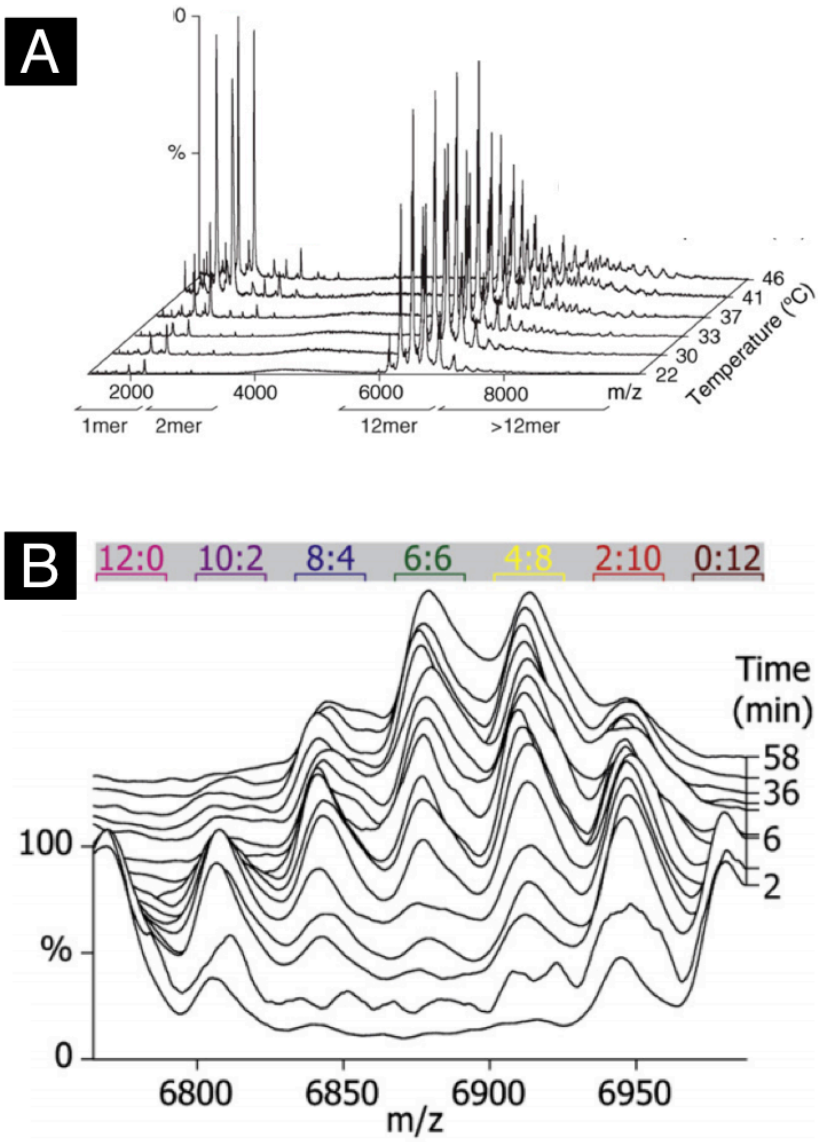


Figure 1.3 Temperature-dependent and Time-Dependent Dissociation of smHSP Dodecamers into Dimers and Monomers in vivo

A As temperature increased, the mass spectrometry profile of Arabidopsis HSP18.1 changed from that of a solely dodecameric species to a mixture of dodecamers, monomers and dimers.

(Figure adapted from Stengel et al., 2010)

B As time passed, the mass spectrometry profile of Arabidopsis HSP17.6 and HSP18.1 changed from that of two distinct groups of homodimers into that of a heterodimeric species, apparently proportional to the molar ratio of input dimers.

(Figure adapted from Painter et al., 2008)

between different organisms – an 18.1 kDa HSP from pea (*Pisum sativum;* PsHSP18.1) and TaHSP16.9. Again, dimers appear to be the currency of exchange and the authors showed that the final bias of one species in an oligomer over the other is determined simply by the molar ratio of monomers in the solution, and that homocomplexes fall to 50% of their original abundance within 3-4 minutes of mixing.

It will be crucial to understanding the role of the smHSPs in plants to resolve this apparent conflict. The distinct sub-cellular localisations of proteins on different branches of the phylogenetic tree go some way to explaining why so many genes are required, and yet, if these proteins are tasked with ensuring the stability of the whole gamut of proteins that exist in each cell, they must surely have a low specificity for their targets. This would explain the ease of exchange, but not the requirement for the large number of unique smHSP proteins.

The localisation signals that have been found so far are encoded in the variable N-terminal region, while the hydrophobic C-terminal extension remains relatively conserved in hydrophobic composition and size across classes. Kim et al. (1998) also expressed rice (*Oryza sativa)* HSP16.9 in *E. coli* with a deletion in the C-terminal two-thirds of the α-crystallin domain. Cells carrying this mutant protein were protected from heat-shock, controversially suggesting that the most highly-conserved domain is not necessary for function. Yeh et al. (1997) substituted several phenylalanine residues N-terminal of the α-crystallin domain and abolished thermotolerance but not oligomerisation. It has been suggested that the N-terminal extension plays a role specifically in client-binding (Siddique et al., 2008, Waters and Vierling, 1999). This then leads us to the conclusion that the thermoprotective function is encoded partly in the N-terminal region

and partly in the C-terminal extension which is most exposed and most easily accessible by proteins outside of an oligomer.

1.5.1.2 Regulation of smHSP Activation

The precise expression and localisation patterns of the full complement of smHSPs in *Arabidopsis* has yet to be published at a protein level, mainly due to high homology within classes which causes difficulties in raising antibodies specific to any particular one (Jinn et al., 1993) and it is apparent that much work has instead been done *in vitro* (as discussed above – see section 1.5.1.1 smHSP Sequences, Structures and Functions). This could be resolved by epitope tagging but this has yet to be reported in the literature; something the research reported in this thesis aimed to address. Much of the analysis published to date instead concerns expression at the transcript level.

The main initiators of the heat shock response with respect to HSPs are a family of transcriptional activators termed heat shock factors (HSFs), of which there are 21 in *Arabidopsis* (Kotak et al., 2007a). They share a common basic structure of an N-terminal DNA binding domain, a hydrophobic multimerisation domain, a nuclear localisation signal (NLS) and an activation domain (Kotak et al., 2007b). The fact that there are several HSFs which activate different components of the heat stress response suggests that they may play divergent roles within the plant, or that subsets of the smHSP family are responsible for discrete, non-redundant functions (Waters et al., 1996).

In general the HSFs are directly or indirectly responsible for the expression of smHSPs (Kotak et al., 2007a). As described in section 1.2.1, one notable exception is HSFA9 which induces HSP17.6 and HSP17.6a transcription in seeds in

a manner that is independent of heat shock (Kotak et al., 2007b). This HSF therefore appears to have adapted to a role in seed development rather than heat shock, perhaps mitigating the ionic-strength-induced protein conformation changes that would take place as a maturing seed desiccates.

1.5.2 MYB Genes in Plants

The heat shock proteins represent one avenue of research. A second aim of this study was to add depth to our understanding of the role of the transcription factor MYB64 in stress tolerance responses by further investigating the other genes that it might activate.

The first MYB gene characterised was the oncogenic v-Myb, a gene found in avian myeloblastosis virus (Klempnauer et al., 1982). Sequence analysis suggests that it was inherited from a genetic transfer event between the virus and a vertebrate host. The defining feature of a MYB transcription factor is its N-terminal DNA-binding domain; a 53 residue stretch comprising two main alpha-helices joined by a shorter helix in a 'helix-turn-helix' motif (Ogata et al., 1992). This domain can occur on its own, but usually it is accompanied by 1 or 2 imperfect repeats, all of which share a conserved tryptophan–rich hydrophobic core. MYB proteins, therefore, are classified based on the number of imperfect repeats: 'MYB1R' factors, 'R2R3-type MYB' factors, and 'MYB3R' factors; R2R3-type MYB factors are the largest of the three families in *Arabidopsis,* with the most recent count placing the figure at 198 genes (Yanhui et al., 2013). The functions attributed to the members of the R2R3-family (for a phylogenetic tree of this family see Figure 1.4) mainly include phenylpropanoid biosynthesis, cell development & identity (e.g. glabrous GL1), responses to plant hormones, and

responses to environmental stimuli such as drought, temperature, and salt (Urao et al., 1993, Abe et al., 2003, Ambawat et al., 2013). A striking feature of these roles is that they represent mainly plant-specific processes, especially when compared with the more general roles of MYB1R factors (telomeric DNA binding (Yu et al., 2000), circadian clock control (Schaffer et al., 2001)) and MYB3R factors (cell cycle control (Ito et al., 2001)). Combined with the evidence that there are so many more R2R3-type MYB factors in higher plants than in other eukaryotes, there is speculation that this family has contributed to the evolution of land plants (Martin and Paz-Ares, 1997, Stracke et al., 2001). These authors also report evidence that although they are closely related at the protein sequence level, many of them are expressed differentially so that functional redundancy is ruled out by spatial or temporal displacement.

In the context of responses relevant to tolerance of climate change, Seo and Park (2009) described the action of MYB96, an R2R3-type identified from an activation-tagging screen where its overexpression caused a reduction in lateral root growth in response to drought conditions (consistent with typical drought responses noted in section 1.2.2 “Hydration Stress Responses”, above). They also showed that ABA signaling was important in the activation of this MYB TF. Other

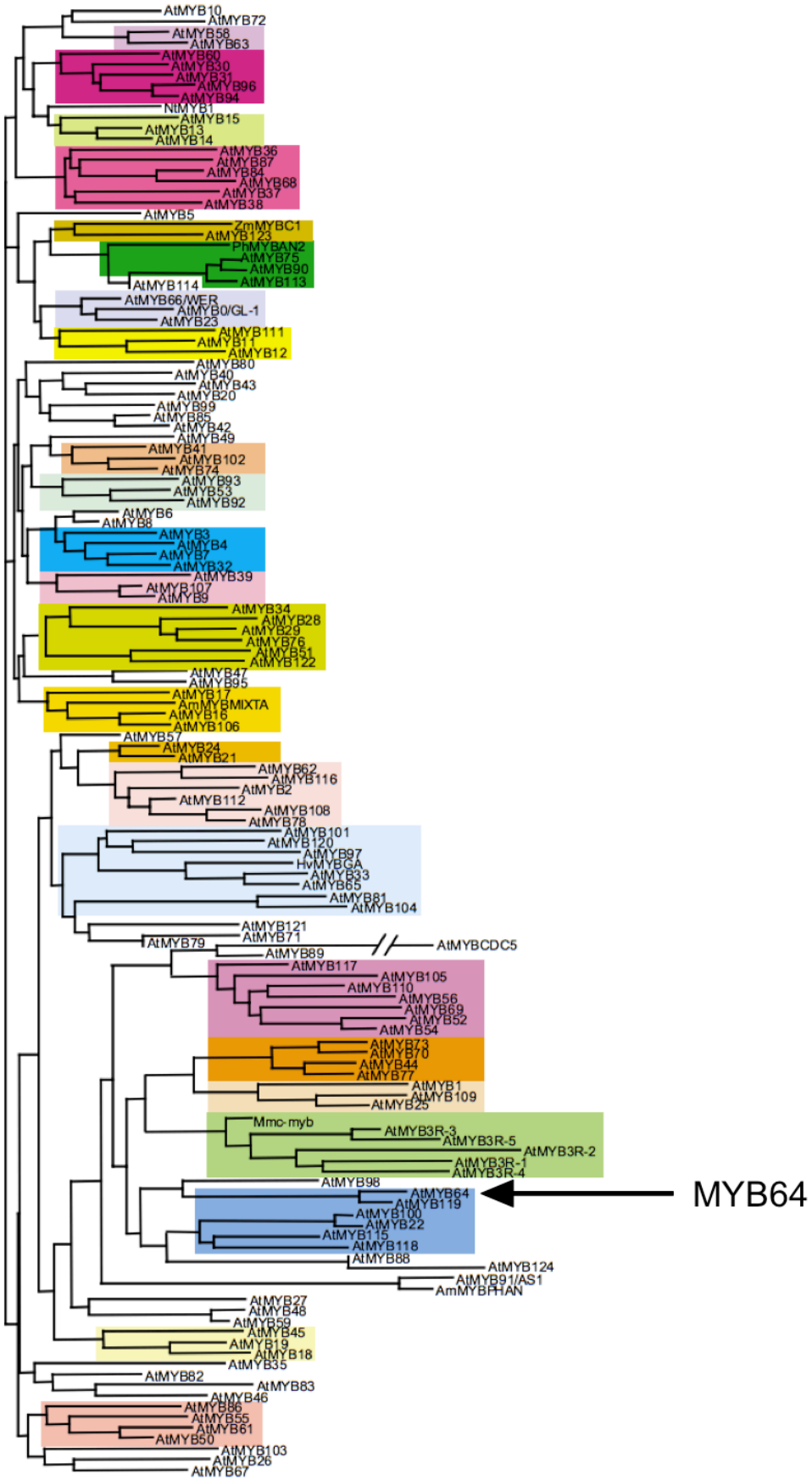


Figure 1.4 Phylogenetic Tree of the R2R3 Family of MYB Transcription Factors in Arabidopsis

MYB64 is most closely related to *MYB119*. Groups (designated by coloured boxes) are determined on the basis of short motifs (4-20 residues) outside of the conserved R2R3 domains. *(Figure adapted from Stracke et al., 2001)*

ABA-inducible, abiotic stress-responsive MYBs include MYB13, MYB15, MYB33 MYB70, MYB73, MYB77, and MYB101 in conditions of drought, osmotic stress, and other environmental factors, and work is underway to translate this research on the model species Arabidopsis into crops (for a review, see Ambawat et al., 2013). To date, outside of the research conducted in our lab, there has been no reported involvement of MYB64 (the R2R3 MYB on which the research presented in this thesis is based) in the halo- or thermotolerant response in plants.

There is controversy in the literature over the existence of a consensus MYB DNA binding domain. Martin & Paz-Ares (1997) report that MYB transcription factors are highly divergent in their target sequences, quoting the wide variation between specificities found in various plant species, none of which are similar to the consensus sequence from vertebrates. Martin & Paz-Arez go on to detail three possible consensus sequences found in various plant contexts. A year later, however, Romero *et al* (1998) rather more confidently published a different series of three consensus binding sites.

A preliminary search for all six sequences referred to above revealed no instances in the putative promoter regions of HSP17.6, HSP17.6a, or HSP17.6b (three of the smHSPs near the top of the list of the 45 most highly upregulated transcripts in line HT5 – see Table 1.1), which would leave open the possibility that perhaps there is an intermediate step between the activation of MYB64 and the expression of these two smHSPs mediated by an unidentified regulator in the pathway.

By investigating the relationship between *MYB64* overexpression, transcript profiles, and stress-response phenotypes, this latter part of this thesis aims to begin to elucidate the likely intermediate steps in such pathways.

2 Materials & Methods

2.1 Nomenclature

The Arabidopsis Information Resource (TAIR, available at www.arabidopsis.info, Swarbreck et al., 2008) is a central repository for Arabidopsis information. Their guidelines were followed when naming lines, vectors and genetic constructs (available at www.arabidopsis.org/portals/nomenclature/guidelines.jsp, TAIR, 2007). The TAIR guidelines are a development of the original community proposals set out by Meinke & Koornneef (1997).

2.1.1 Gene Numbers

After the completion of the project to sequence the Arabidopsis genome, gene numbers were assigned according to the following convention: 'At' represents the species name (*Arabidopsis thaliana*); 1, 2, 3, 4, or 5 represent the chromosome number (or M for Mitochondrial and C for Chloroplast); 'G' stands for Gene (other letters assigned to different features, such as repeats); and the final five digits were allocated sequentially to the genes identified by *in silico* sequence analysis along each piece of DNA. In the initial annotation exercise the final digit in the gene number was always 0, allowing space for 9 other gene numbers to be inserted should more be discovered by further experiments or improved *in silico* analysis. Example: locus At5g12020 – *Arabidopsis thaliana*; chromosome 5; gene; 1202nd feature found from the 'top' of the chromosome. Alternatively spliced products would be labelled as At5g12020.1, At5g12020.2 *etc*.

2.1.2 Gene Names / Symbols

Where possible, this thesis uses the standard 3-letter symbols for genes followed by a number, if necessary, to help distinguish one relative from another. Wild type genes are referred to in *UPPER CASE ITALICISED* text, while gene products are referred to in UPPER CASE NON-ITALICISED text. Traditional mutant alleles arising from single nucleotide polymorphisms (SNPs) are not discussed in this thesis.

The heat shock protein (HSP) gene family in Arabidopsis represents a historical exception to convention as it encodes many similar proteins which have been named and numbered to reflect their molecular masses rather than being numbered sequentially by date of discovery. The HSP gene family is broadly classified into two sub-groups based on size: those encoding proteins in the region of 70 – 100 kDa, and those in the region of 17 kDa (the latter termed "small HSPs" - smHSPs). The 3-letter functional symbol for the smHSPs is appended with their molecular mass in kDa, measured to 1 decimal place where necessary, *e.g. HSP70*, *HSP17.4*, and *HSP17.6*. A further letter is appended where two smHSPs are of equivalent mass at an accuracy of 1 decimal place, *e.g. HSP17.6*, *HSP17.6a*, *HSP17.6b*. In order to avoid redundancy in databases and in publications among the scientific community, the use of the letters "At" before the gene symbol has been discouraged (TAIR, 2007), so for the sake of brevity these prefixes have not been included for Arabidopsis gene names in this thesis either. Species prefixes have been used, however, for genes from species other than *A. thaliana* (*e.g. TaHSP16.9* for a homolog found in the wheat *Triticum aestivum*).

2.1.3 Genetic Constructs

Since TAIR gives no advice on naming genetic constructs, guidelines were sought from the journal with the highest impact factor in the field of plant biology that published primary research (Journal Citation Reports, Web of Knowledge, 2012) which, at the time of writing, was Plant Cell (Impact Factor 9.24; rank: 4th). Double-colons are to be reserved for inserted genetic elements, *e.g.* LFY::TAG1, and genetic fusions are to be written with a single colon between elements (Instructions for Authors, Plant Cell, 2013), so genetic constructs are written in this thesis with single colons. Where an exogenous promoter is incorporated into a construct, its source name is appended with "pro" to avoid confusion with the lower case "p" at the start of plasmid names, *e.g.* 35Spro:*HSP17.6*:*RFP*.

2.1.4 Transgenic Lines (Gain- and Loss-of-Function)

Lines generated by T-DNA insertion and obtained from stock centres are initially referred to in this thesis using the names given by their originators for consistency with the literature. Following the proposal of Young et al. (2001), those confirmed to have a knockout phenotype are thereafter referred to using the standard italicised gene name written in lower case to represent a loss-of-function.

Lines generated by T-DNA insertion in the laboratory during this course of work were all designed to overexpress a gene of interest. These are referred to using the terminology discussed in section 2.1.3, appended with any identifier assigned to distinguish it during the course of transgenesis *e.g.* line 35Spro:MYB64 line 127, or 35Spro:MYB64 line 141.

2.2 Materials

2.2.1 Chemicals

All chemicals were purchased from Sigma-Aldrich Chemical Co. ltd, Dorset UK, or Fisher Scientific Ltd., Loughborough UK.

2.2.2 Antibiotics

Antibiotics were made up as stock solutions described in Table 2.1, filter sterilised (Sartorius Minisart disc filter, 0.2 µm), then stored at -20 °C.

Table 2.1 Antibiotics

Antibiotic	Solvent	Stock concentration	Working concentration	Supplier
Carbenicillin	60 % EtOH	50 mg / ml	50 µg / ml	Fisher Scientific
Spectinomycin	H_2O	100 mg / ml	100 µg / ml	Sigma Aldrich
Kanamycin	H_2O	50 mg / ml	50 µg / ml	Melford
Gentamycin	H_2O	30 mg / ml	30 µg / ml	Melford
Rifampicin	DMSO	50 mg / ml	50 µg / ml	Sigma Aldrich

2.2.3 Bacterial Strains

2.2.3.1 General *Escherichia coli* Cloning Host

During cloning experiments, plasmids and the constructs generated therefrom were maintained in Top10 *Escherichia coli* cells (Invitrogen, Paisley, Scotland), genotype F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu) 7697 galU galK rpsL (StrR) endA1 nupG λ-.

2.2.3.2 *Agrobacterium tumefaciens* Strain for Plant Transformation

Agrobacterium tumefaciens strain AGL1 was used for all plant transformations and transfections. This strain was obtained from Craig Carr (University of Glasgow).

2.2.3.3 *Agrobacterium tumefaciens* Transformation-booster

Agrobacterium strain AGL1 carrying expression constructs of interest were co-infiltrated along with Agrobacterium strain AGL1 carrying a pair of transformation-boosting vectors, pGreen and pSoup, to improve transfection efficiency. The pGreen/pSoup transfection system consists of a binary Ti vector (pGreen) which is able to replicate independently in *E. coli*, but which requires a second vector (pSoup) to provide replication capability *in trans* when in Agrobacterium (Hellens et al., 2000). This binary system was used to express *VirGN54D* (an overactive variant of *VirG*; a transcriptional activator of the rest of the Agrobacterium virulence (*vir*) genes (Pazour et al., 1992)) from pRT18 (pGreen-based), which was used to transform Agrobacterium along with pSa-Rep (which is pSoup-based) (Vain et al., 2004). This transformation-boosting strain was obtained from Craig Carr (University of Glasgow).

2.2.4 Bacterial Growth Media

2.2.4.1 Lysogeny Broth (LB) and Plates for *E. coli* and Agrobacterium

Lysogeny broth (Bertani, 2004) was used as the culture medium for *E. coli*. The following was added to 900 ml dH2O:

- 5 g NaCl
- 10 g Tryptone
- 5 g Yeast extract

The pH was adjusted to 5.8 then volume was adjusted to 1 l before autoclaving.

For plates, agar was added at 15 g/l before autoclaving.

2.2.5 Plasmid Vectors

2.2.5.1 Cloning Vector

All cloning procedures were carried out using the vector pENTR-D-TOPO (Invitrogen, Paisley, Scotland). This is the 'entry vector' for the GATEWAY system, from which cloned DNA can be transferred into 'destination vectors' (Table 2.2) carrying a range of 5' and 3' sequences for various purposes.

2.2.5.2 Expression Vectors

The destination vectors in Table 2.2 encoding fluorescent tags were used to transfect *Nicotiana benthamiana* with GFP, YFP or RFP-tagged small heat shock proteins (smHSP), and the destination vectors encoding epitope tags were created for future experiments involving immunoprecipitation of smHSPs.

Table 2.2 Vectors for Plant Transfection

Vector	Construct	Tag size (amino acids)
	Fluorescent tags	
pH7FWG2[1]	35Spro :: *GATEWAY cassette* :: GFP	240
pH7YWG2[1]	35Spro :: *GATEWAY cassette* :: YFP	240
pB7RWG2[1]	35Spro :: *GATEWAY cassette* :: RFP	226
	Epitope tags	
pGWB10[2]	*GATEWAY cassette* :: FLAG	8
pGWB11[2]	35Spro :: *GATEWAY cassette* :: FLAG	8
pGWB13[2]	*GATEWAY cassette* :: HA	40
pGWB14[2]	35Spro :: *GATEWAY cassette* :: HA	40
pGWB16[2]	*GATEWAY cassette* :: 4xMYC	56
pGWB17[2]	35Spro :: *GATEWAY cassette* :: 4xMYC	56

[1] (Karimi et al., 2002)
[2] (Nakagawa et al., 2007)

2.2.6 Arabidopsis Lines

2.2.6.1 Wild-type *Arabidopsis*

Where segregated wild-type was not available as a control for screening experiments (for example, when lines acquired from outside sources were already homozygous and therefore the mutation was not segregating with each generation) stocks of the corresponding background line acquired from NASC were used (usually Col-0 or Col-7).

2.2.6.2 Small Heat Shock Protein (smHSP) Knockout Lines

The TAIR database was interrogated to find insertional mutations in and around the heat shock protein (HSP) genes of interest. Mutations were deemed to be of interest either if they were within the coding region of the gene, or if they were within 1 kb upstream or downstream of either end. These insertional mutants

were acquired and labelled as putative HSP knockout lines. The insertional lines were generated by Alonso et al. (2003) who subjected wild-type *Arabidopsis* (ecotype Columbia; Col-0) to Agrobacterium-mediated insertional mutagenesis and interrogated the genomic sequences of the resulting progeny to determine the precise locations of the insertions. Over 88,000 unique individuals were identified and seed was deposited at the Nottingham Arabidopsis Stock Centre (NASC, Scholl et al., 2000). Several lines from this collection were chosen based on the location of insertions either within, or near to, HSP genes of interest and acquired for further study. The details of each insertion are presented in Table 2.3.

Table 2.3 Knockout Arabidopsis Lines

Gene	AGI number	Original line name	NASC stock ID	Line details
HSP17.6	At5g12020	SALK_007510	N507510	Single exon gene; insertion 65bp upstream of stop codon.
HSP17.6a	At5g12030	SALK_072448	N572448	Single exon gene; insertion 152bp upstream of start codon.

2.2.6.3 Arabidopsis Activation-Tagged Lines

Activation-tagged lines of Arabidopsis used in screens for novel thermotolerant mutants were acquired from a collection initially generated by Weigel et al. (2000) and subsequently deposited at the NASC. The wild-type background for this collection was ecotype Columbia (Col-7). The entire Weigel collection comprises three sets deposited at the NASC at various times. Each set was available in various sizes of pools of individual transformant lines. Set 1 was acquired as 86 pools of 100 lines each (stock code N21995). Set 2 was acquired as 82 pools of 96 lines each (stock code N21991). Set 3 was acquired as 62 pools

of 100 lines each (stock code N23153). In total, these 3 sets represent 22,672 individual transformant lines for screening.

2.3 Tissue Culture Methods

2.3.1 Sowing Arabidopsis on Soil

Soil (Levington F2) was autoclaved and infused with a solution of the insecticide Intercept (Scotts, UK) at the working concentration according to the manufacturer's instructions to prevent the growth of fly larvae among plant roots.

Seeds were sown directly onto the surface of damp soil, then watered from above. Pots were covered with cling-film to prevent cross-contamination, seeds were stratified for 2-3 nights at 4°C in the dark, then pots were then transferred to the growth room at 22 °C to allow germination.

2.3.2 Sowing Arabidopsis on Agar Plates

2.3.2.1 Seed Surface Sterilisation

Seeds to be grown on agar medium were first surface-sterilised to prevent contamination of the agar with fungi and bacteria. Bleach solution was prepared as follows: 1 Covchlor 1000 Chlorine tablet was dissolved in 35ml dH2O; 5ml of this solution was added to 45ml ethanol with 1 drop (~500 µl) Tween-20; then the solution was allowed to sit at room temperature for 10 minutes before centrifugation at 800 x *g* for 10 minutes to pellet the white precipitate. This bleach was prepared fresh each time it was required.

Arabidopsis seeds were decanted into a 1.5 ml microfuge tube. At each step in the following procedure, silique (*i.e.* seed pod) debris was removed by pipetting at each change of solution to ensure the only material remaining was seed itself. (This was necessary as the presence of silique material correlated with subsequent contamination of the plates. It is hypothesised that this is related to the porosity of the dried silique material and the action of surface tension slowing the progress of the bleach solution at the openings of these pores, wherein microbiological contaminants would remain protected from the bleach.) Under sterile conditions, seeds were washed twice briefly with 70 % ethanol to remove as much excess material as possible, then 1 ml of bleach solution was added for no less than 5 minutes (to avoid incomplete sterilisation) and no longer than 7 (to avoid reduced viability). Five washes with 90 % ethanol were used to remove the chlorine, then the seeds were either pipetted onto an ethanol-soaked Whatman filter paper, grade 42 (Sigma-Aldrich, UK), allowed to dry, then sprinkled across an agar plate, or suspended in 0.1 % MSMO (Murashige and Skoog with Minimal Organics) agar (see below) to accurately pipette onto an agar plate.

Where required and noted in the text, borders were marked on plates in order to divide them into two or three sections (bipartite or tripartite split plates). This was so that lines of different genotypes could be sown and analysed on the same plate, thereby removing a source of experimental variation. This was especially important for heat stress experiments due to challenges in ensuring even heating conditions throughout an incubation chamber.

2.3.2.2 Preparation of Agar Plates and Top Agar

Agar for plates was prepared by dissolving MSMO powder (Sigma-Aldrich Company Ltd., Dorset, UK) in dH_2O to 1/10 normal strength and supplemented with 0.75 % sucrose as a carbon source. The pH was adjusted to 5.8, and 1.5 % agar (Fisher Scientific, UK) was added before autoclaving.

For experiments where positioning of seeds on the plate was important (*e.g.* on vertical plates where root length was to be measured) seeds were suspended in top agar after the final sterilisation wash step so that they could be accurately and easily positioned by pipette. Top agar was prepared as above but with a reduced concentration of 0.1 % agar.

2.3.3 Growth of Arabidopsis on Soil and Plates: 'Normal Conditions'

Seeds sown either on soil or on agar plates were grown in long day conditions: 16 hour photoperiod; light intensity approximately 100 – 150 $\mu M.m^{-2}.s^{-1}$; 22 °C/18 °C day/night temperature.

2.4 Salt Stress Methods

Salt stress experiments were carried out on plate-grown seedlings. A 5 M stock solution of NaCl was prepared in dH_2O and filter-sterilised (Sartorius Minisart disc filter, 0.2 µm). Agar was prepared as described above in section 2.3.2.2 and supplemented with a sufficient volume of the NaCl stock solution to reach the desired final concentration (40, 60 or 80 mM). Top agar was also supplemented to the same final concentration.

2.5 Heat Stress Methods

An optimum thermal acclimation and heat stress regime was empirically determined as part of this work, as described in section 3.5.1 ("Thermotolerance of Wild Type Arabidopsis"). Briefly, plants were acclimated by incubating at 37 ± 0.5 °C in the dark using a fan-circulated incubator for a period of 1 hour, returned to Normal Conditions for 3 hours, then exposed to a heat stress temperature of 44 ± 0.5 °C in the same incubator for a period of 3 hours (unless stated otherwise).

2.6 Molecular Methods

2.6.1 Isolation of Nucleic Acids from Plant Tissue

2.6.1.1 Genomic DNA Isolation

Approximately 200 mg tissue was removed from plants (when sampling leaves) or 200 mg of whole seedling tissue was transferred to a sterile microfuge tube. Tissue was frozen by submerging the tube in liquid nitrogen, a sterile ball bearing was added, and the tissue was ground in a TissueLyser automated grinder (Qiagen, West Sussex, UK) at 26 rpm for 30 seconds. 1 ml plant DNA extraction buffer (200 mM Tris (pH 7.5), 250 mM NaCl, 25 mM EDTA, 0.5 % SDS) was added and allowed to stand for 10 minutes at 4°C. Samples were pelleted by centrifugation at 13,000 rpm in a benchtop microcentrifuge at 4 °C for 15 minutes, then the supernatant was discarded. Pellets were washed repeatedly with DNA extraction buffer until the green colour had gone (~5 times). Supernatant was removed one final time then 50 µl dH_2O was added to extract

DNA from the pellet, mixed by vortexing briefly, then cell debris was pelleted once again at 13,000 x *g* for 15 minutes. The resulting supernatant with DNA in solution was removed to a fresh sterile microfuge tube.

2.6.1.2 Quantification of DNA

Where necessary, DNA was quantified by UV spectrophotometry. Absorbance was measured at 260 nm and 280 nm and a ratio of the two values calculated as an indicator of purity (optimal value 2.0; Sambrook and Russell, 2001) then satisfactorily pure samples were quantified according to the following formula:

$$A_{260} \times 50\ \mu g/\mu l \times \frac{\text{Final volume}}{\text{Sample volume}} = [\text{DNA}]\ (\mu g/\mu l)$$

2.6.1.3 Total RNA Isolation

Approximately 200 mg tissue was excised with a razor blade (individual leaves) or harvested from plates (whole seedlings), put into microfuge tubes and snap-frozen in liquid nitrogen. A sterile ball bearing was added to each tube to facilitate automated grinding in a TissueLyser automated grinder (Qiagen, West Sussex, UK) at 26 vibrations per second for 30 seconds.

RNA was extracted using Tri Reagent (Sigma-Aldrich Company Ltd., Dorset, UK): 1 ml of Tri Reagent was added to each microfuge tube of ground tissue and allowed to stand for 5 minutes at 4 °C. Samples were centrifuged for 5 minutes at 13,000 x *g* to pellet any cellular debris, the supernatant was removed to a fresh microfuge tube, then 0.2 ml chloroform was added and the sample was shaken vigorously for 15 seconds to create an evenly mixed emulsion, then placed back on ice for 10 minutes. Another centrifugation was performed for 10

minutes at 13,000 x *g* to encourage complete separation of the aqueous phase (containing nucleic acid) from the phenolic phase (containing protein and lipid), then the aqueous upper phase was removed to a fresh microfuge tube containing 0.5 ml isopropanol on ice to precipitate the nucleic acid. After 10 minutes the precipitate was pelleted by centrifugation for 15 minutes at 13,000 x *g* then washed twice in ice-cold 100% ethanol. The supernatant was discarded and the pellet was resuspended in 30 µl DEPC-treated (*i.e.* nuclease-free) water and stored at -20 °C.

2.6.1.4 DNase Treatment of RNA

Contaminating genomic DNA was removed using the TURBO DNA-free kit (Ambion, Texas, USA): 0.1 volumes of 10x TURBO DNase Buffer and 1 µl of TURBO DNase were added to the RNA and the mixture incubated at 37 °C for 20-30 minutes. 0.1 volumes of DNase Inactivation Reagent were added to remove divalent cations and thereby cease the reaction. The Inactivation Reagent was pelleted by centrifugation at 13,000 x *g* for 2 minutes, and the RNA removed to a fresh microfuge tube and stored at -20 °C.

2.6.1.5 Quantification of RNA

RNA extraction and purification, described above, were the first steps in comparing gene expression levels by the Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR; see section 2.6.3). In order to be able to standardise the quantity of RNA in each RT-PCR reaction, each sample was quantified by UV spectrophotometry. Absorbance was measured at 260 nm and 280 nm. The ratio of the two values was calculated as an indicator of purity (optimal value 1.8;

Sambrook and Russell, 2001) then satisfactorily pure samples were quantified according to the following formula:

$$A_{260} \times 40\ \mu g/\mu l \times \frac{\text{Final volume}}{\text{Sample volume}} = [RNA]\ (\mu g/\mu l)$$

In order to ensure that each complementary DNA (cDNA) sample within an experiment would be synthesised to an equal concentration, it was important to standardise the amount of RNA that would be added to the reverse transcription reaction. 1 µg was chosen. When the extracted RNA concentration was insufficient to provide 1 µg within the volume limits specified by the protocol, the maximum permitted volume of RNA was used and the mass of RNA contained therein calculated, so that all other samples could be adjusted proportionally downwards for equal input.

2.6.2 Polymerase Chain Reaction (PCR)

PCR was performed in an MJ Research DNA Engine PTC-200 Peltier Thermal Cycler (Genetic Research Instrumentation, Essex, UK) on genomic DNA using ReddyMix complete master mix (Catalog #AB0785B, ThermoScientific, Belgium) with constituents as listed in Table 2.4. Thermal cycling followed the pattern described in Table 2.5

Table 2.4 Components of PCR Reaction Mix

Component	Final Concentration
Primers	1.25 µM each
DNA template	Variable
ReddyMix Master Mix	1x
dH2O	To 20 µl

Table 2.5 Generic Thermal Cycle Used with ReddyMix for Standard PCR

Step	Temperature	Duration
1	95 °C	3 minutes
2	95 °C	30 seconds
3	(Primer-specific annealing temperature)	30 seconds
4	72 °C	1 min/kb
4	Go to step 2 for an empirically-determined number of cycles	N/A
5	72 °C	5 minutes
6	4 °C	Forever

2.6.3 Reverse-Transcriptase PCR (RT-PCR)

2.6.3.1 cDNA Synthesis

In the first step of reverse transcriptase PCR (RT-PCR) first-strand cDNA was synthesised using the Omniscript RT Kit (Qiagen, West Sussex, England). A master mix was prepared consisting of the following components per cDNA synthesis reaction: 2 µl 10 x Omniscript reverse transcriptase buffer; 1 µl (4 units) Omniscript reverse transcriptase; deoxyribonucleotide (dNTP) mix at a final concentration of 0.5 mM each; 1 µl (10 units) RNase inhibitor, oligodeoxythymidine (oligodT) primer at a final concentration of 1 µM, and RNase-free water to a final concentration of 20 µl (taking account of the volume of RNA template to be added). Master mix was briefly vortexed, aliquoted into reaction tubes, and RNA templates were added. These final reaction mixtures were mixed by pipetting, then incubated at 37 °C for 1 hour.

2.6.3.2 RT-PCR

From each RT reaction, an equal volume of first-strand cDNA was added to normal PCR mixes (see section 2.6.2) in place of genomic DNA. PCR was carried out as normal.

2.6.4 Agarose Gel Electrophoresis of DNA

Agarose gels were prepared by melting 1 % agarose (w/v) in 1 x TAE buffer (40 mM Tris-acetate, 1 mM EDTA) then pouring into gel casts. To stain DNA, SYBR Safe (Invitrogen, Paisley, UK) was added to molten Agarose at a dilution of 1:10,000 and mixed thoroughly. DNA samples were loaded alongside a molecular weight marker (1kb DNA Ladder, Promega, Southampton, UK) and separated by electrophoresis in 1 x TAE buffer at approximately 70 V and visualised by UV illumination.

2.6.5 Agarose Gel Electrophoresis of RNA

Where required, integrity of RNA was checked by agarose gel electrophoresis. 1 µg of RNA (see section 2.6.1.5) was separated on a 1.5 (w/v) agarose gel containing 10 % formaldehyde and 1 x MOPS buffer, pH 7.0 (20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA (Sambrook and Russell, 2001)). Before loading, the RNA was mixed with 1 % (v/v) formaldehyde, 30 % (v/v) formamide, 1 x MOPS pH 8.0, and 0.1 volumes of ethidium bromide as a staining agent. RNA mixtures were heated at 65 °C for 10 minutes, cooled on ice, then mixed with 0.2 volumes of loading dye (Promega UK, ltd., Southampton, UK) then loaded on the MOPS gel. Electrophoresis was performed in 1 x MOPS buffer pH 7.0 for 2 hours at 100 V and visualised by UV illumination. RNA integrity was assessed by the presence of defined bands.

2.6.6 Quantitative PCR (qPCR)

2.6.6.1 Reference Gene Selection

Actin 2 (At3g18780) was initially selected as an internal reference gene candidate because of its widely assumed nature as a housekeeping gene with a stable expression pattern, but this was proven not to be the case when *MYB64* was constitutively overexpressed and under heat stress. Various sources were consulted to find a more suitable reference gene candidate, and a protein phosphatase subunit was selected (PP2A; At1g13320) (for a full account see section 4.5.1). As primers had already been designed, tested and published, it was decided that these were the best reference genes against which to normalise during heat-stress qPCR investigations (for all primer sequences, see section 2.6.13).

2.6.6.2 Determination of qPCR Efficiency

All primer sets were tested to determine the reaction efficiency before use in qPCR analysis. Reaction efficiency was determined using a plasmid containing each gene of interest to generate a standard curve. Six 10-fold dilutions were made from a stock solution of 100 pg of plasmid containing the gene of interest (GOI) in order to generate a standard curve. Each dilution was mixed with 12.5 µl SYBR Green and 0.5 µl of each primer and qPCR was performed.

For calculation of reaction efficiency, the log of the RNA concentration was plotted on the X-axis and Ct (number of cycle) values on the Y-axis. A line of best fit was generated and reaction efficiencies (RE) were determined using the equation: $RE = 10^{(-1/m)}/2*100$, where m is the slope of the line. Each data point

was tested in duplicate (technical replication). Two negative control reactions were also carried out with no DNA template. Primer pairs that resulted in efficiencies of >90 % were deemed acceptable; those with efficiencies of <90 % were redesigned.

2.6.6.3 Reaction Setup Used in qPCR

Quantitative PCR reactions were set up using an automated liquid handling system (CAS- 3200; Corbett Robotics, Sydney, Australia). One microgram of total RNA was converted into cDNA and each cDNA was diluted 1:3. PCR reactions were performed using the Brilliant II SYBR Master Mix kit (Stratagene, UK) containing 5 µl of each cDNA sample in a reaction mixture (20 µl) containing 1 x SYBR Green and 0.5 µM of each primer. Reactions were set up in duplicate (technical replication). The thermal cycle was as described in Table 2.6. The fluorescence emitted during this final step allowed the specificity of the amplified PCR product to be determined; a single PCR product of the correct sequence should dissociate at one discrete temperature, and thus the fluorescence signal should exhibit a narrow peak somewhere within the range from 55 to 85 °C.

Table 2.6 Generic Thermal Cycle Used for qPCR

Step	Temperature	Duration
1	95 °C	10 minutes
2	95 °C	30 seconds
3	55 °C	1 minute
4	72 °C	1 minute
5	Go to step 2 for 40 cycles	N/A
6	95 °C	1 minute
7	55 °C to 85 °C	30 seconds

2.6.7 Cloning

2.6.7.1 Primer Design for Directional Cloning

The open reading frame (ORF) of each gene of interest was cloned into the GATEWAY system (Invitrogen, Paisley, UK). This comprises an 'entry vector' into which a gene of interest can be cloned, and from which the gene of interest can subsequently be transferred by means of an enzyme reaction to a range of expression vectors with the appropriate GATEWAY receiving cassette (see section 2.2.5). The entry vector used to clone all constructs, pENTR/D-TOPO, was supplied as a linearised plasmid with a molecule of topoisomerase covalently linked to the 3'OH of each end. This facilitates recombination and circularisation with a PCR product without the need to supply a ligase enzyme. The sequence of the plasmid allows for directional cloning; the end of the vector that is to be recombined adjacent to 5' end of the PCR product has a 3' overhang with the sequence *CACC*. By incorporating this short sequence as a 5' extension to the 5' cloning primer, base complementarity ensures that the PCR product anneals and recombines with the entry vector in only the intended direction.

Cloning primers for the three smHSP genes (At5g12020, At5g12030 and At1g53540) were designed to anneal to the extreme 5' and 3' ends (excluding the stop codon) of each gene's ORF for use in expression vectors with an endogenous promoter, or alternatively to the 5' end of the 1kb putative promoter for native-promoter versions (see Figure 2.1). The sequence upstream of At1g53540 contained part of the ORF of the adjacent gene. While it is not unknown for a coding region to also contain transcriptional regulation sequences

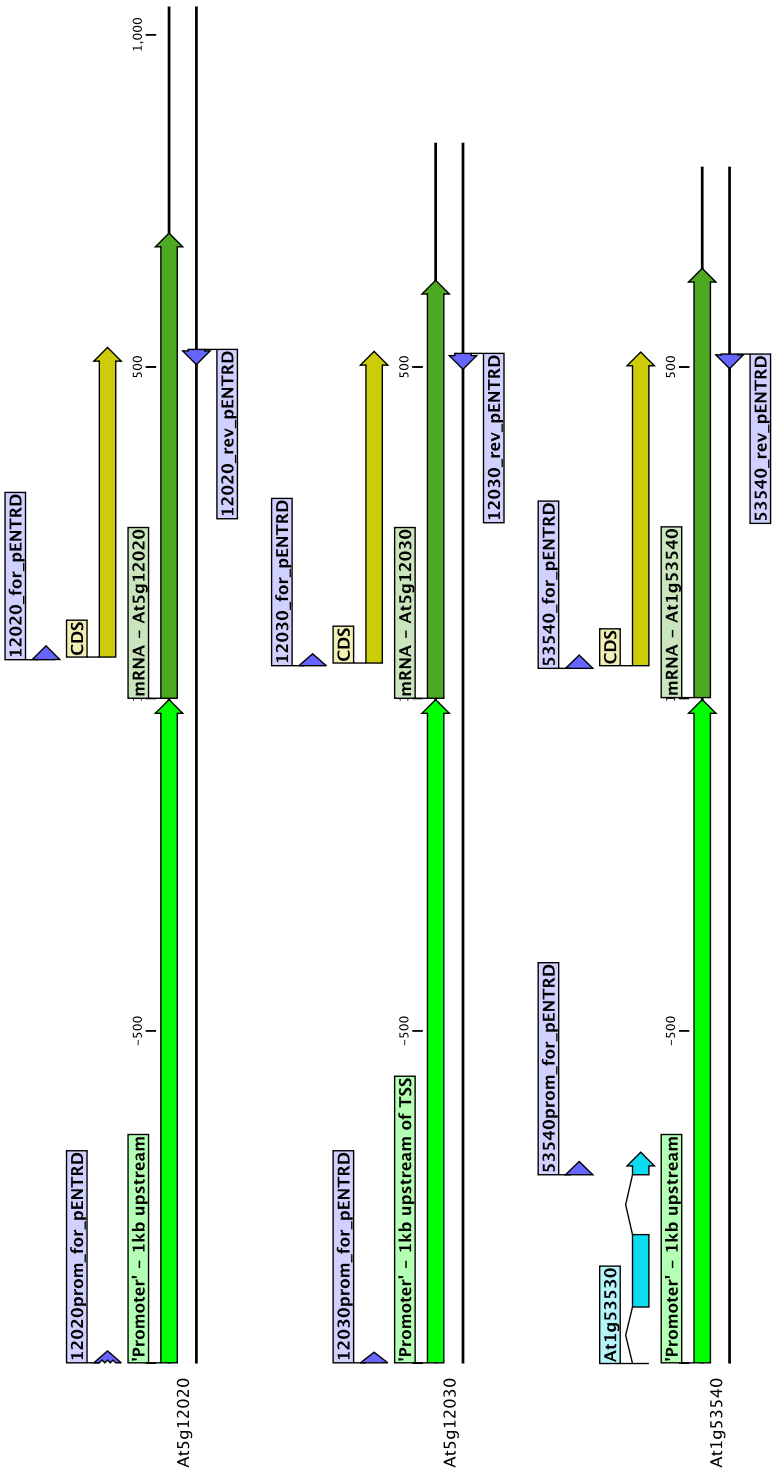


Figure 2.1 Small HSP Cloning Primer Positions.

All three smHSP ORFs were cloned using primers designed to anneal to the extreme 5' and 3' ends. Versions of the genes containing the putative promoter regions were cloned using 5' primers designed to anneal 1kb upstream of the transcription start sites, except At1g53540 which encoded the end of the adjacent gene. While it is not unprecedented for a coding region of one gene to also function as an *in cis* regulator of another, only the last exon of this adjacent gene was incorporated in order to keep the complexity of the construct low. The 5' primers carried a 5' extension with the sequence CACC to facilitate directional recombination with the entry vector pENTR/D-TOPO.

for neighbouring genes, only the final exon of this adjacent gene was incorporated in an effort to keep the complexity of the construct as low as possible.

PCR was performed on genomic DNA using the Expand High Fidelity PCR System (Roche, West Sussex, UK) (Section 2.6.7.2). The 5' cloning primer for each gene carried the required 5' CACC extension for directional recombination with the entry vector (Table 2.10).

2.6.7.2 High Fidelity PCR for Cloning

When cloning genes, PCR was performed using a high-fidelity proof-reading DNA polymerase to reduce the chance of incorporating mismatched bases. The Expand High Fidelity System (Roche, UK) was used and reaction mixes were prepared according to the manufacturer's instructions as described in Table 2.7 and Table 2.8.

Table 2.7 Expand High-Fidelity PCR System: Reaction Mix 1

Component	Final Concentration
Expand High Fidelity Buffer with $MgCl_2$	1 x ($MgCl_2$: 1.5 mM)
High fidelity Taq polymerase	2.6 Units / reaction
dH_2O	To 25 µl

Table 2.8 Expand High-Fidelity PCR System: Reaction Mix 2

Component	Final Concentration
Primers 1 and 2	300 nM each
dNTP mix	200 µM each
Template DNA	Variable
dH_2O	To 25 µl

These were combined into one reaction mixture just prior to incubation in the thermal cycler to prevent 3'-5' exonuclease activity of the proofreading polymerase upon the template. The thermal cycle was as described in Table 2.9.

Table 2.9 General Thermal Cycle Pattern Used for High-Fidelity PCR for Cloning

Step	Temperature	Duration
1	95 °C	3 minutes
2	95 °C	30 seconds
3	Primer-specific	30 seconds
4	72 °C	1 min/kb
4	Go to step 2 for an empirically-determined number of cycles	N/A
5	72 °C	5 minutes
6	4 °C	Forever

2.6.7.3 Isolation of PCR Products from Agarose Gel

After agarose gel electrophoresis of the PCR product to be cloned, the DNA was extracted from the agarose gel. The fragment was excised using a clean, sharp, razor blade and transferred to a sterile microfuge tube. The DNA fragment was purified using a QIAquick Gel Extraction Kit (Qiagen, West Sussex, UK) following the manufacturer's instructions.

2.6.7.4 Cloning into GATEWAY Entry Vectors

The isolated PCR product was then recombined with the pENTR D-TOPO entry vector according to the manufacturer's instructions. A cloning reaction was set up incorporating 5 ng PCR product, 20 ng pENTR D-TOPO vector (supplied in a solution with final working concentrations of: 8.33 % glycerol, 8.33 mM Tris-HCl pH 7.4, 0.17 mM EDTA, 0.33 mM DTT, 0.017% Triton X-100, 17 µg/mL BSA, and 5 µM bromophenol blue), 200 mM NaCl, and 10 mM $MgCl_2$, made up to 6 µl with dH_2O. The reaction mix was left at room temperature for 5-10 minutes then 2 µl of the finished reaction mixture was used to transform chemically competent

E.coli. Transformants were selected on LB agar plates containing kanamycin and analysed by miniprep, restriction digestion, and sequencing.

2.6.8 Transformation of Chemically Competent *E. coli* Cells

Half of the plasmid DNA from the cloning reaction described above was added to a 50 µl aliquot of chemically competent Top10 *E. coli* cells (Invitrogen, Paisley, UK) kept on ice; the other half was retained for repetition in the event of technical failure. This cell suspension was briefly mixed by flicking and immediately placed back on ice for 5 to 30 minutes to allow DNA uptake by the cells. The suspensions were heat-shocked in a water bath at 42 °C for 30 seconds and returned to ice immediately. After two minutes, 250 µl LB broth was added and suspensions were then incubated at 37 °C with at shaking at 200 rpm for one hour. Suspensions were spread on LB agar plates containing appropriate antibiotics for plasmid selection and incubated at 37 °C overnight.

2.6.8.1 Colony PCR

Transformant colonies were checked for the presence of the intended plasmid the following day using primers to the PCR product that was intended to be cloned. Master mixes for normal PCR were set up according to the protocol in section 2.6.2. A sterile pipette tip was used to pick cells from a bacterial colony, then the tip was then stirred into an aliquot of master mix in place of template DNA. The thermal cycle used was as described in Table 2.5, and the high temperatures were sufficient to lyse cells and allow the PCR mixture to gain access to the plasmid templates therein.

2.6.8.2 Plasmid DNA Isolation

A single colony that had been shown to be successfully transformed was used to inoculate 5 ml of LB broth supplemented with the appropriate antibiotic. The culture was grown overnight at 37 °C with constant shaking at 200 rpm. The plasmid DNA was isolated from the overnight culture using the QIAprep Spin Miniprep Kit (Qiagen, West Sussex, UK) according to the manufacturer's instructions.

2.6.8.3 Confirmation of Cloned Sequences

Sequencing was carried out by GATC Biotech (London, UK) in order to confirm that the clones carried the correct sequences. Sequencing results were analysed with CLC Genomics Workbench v3 using the Sequencing Data Analysis tool.

2.6.8.4 Transfer into GATEWAY Destination Vectors – The LR Reaction

After confirming that the GATEWAY entry vectors contained the correct sequences, they were 'flipped' into the appropriate destination vectors which include 5' and 3' sequences to express the genes of interest in various ways (see section 2.2.5.2 for a complete of vectors used). The destination vectors are supplied as empty plasmids carrying only the *ccdB* gene (used for negative selection) in the cloning site. The cloning site in the destination vectors are flanked by sequences utilised for recombination by bacteriophage λ (*attL* sites) and the entry vectors have corresponding sequences on either side of the cloned gene of interest (*attR* sites). When mixed, the recombination reaction is mediated by bacteriophage λ Int and Excisionase (Xis), as well as the E. coli

Integration Host Factor (IHF). These three enzymes are provided as a proprietary mix in the GATEWAY LR recombination kit (Invitrogen, Paisley, UK).

The kit protocol was carried out as follows: 300 ng of entry vector plasmid DNA carrying the gene of interest was added to 300 ng of destination vector, along with 4 µl of 5 x LR Clonase reaction buffer, and TE Buffer pH 8.0 to a total volume of 16 µl. The reaction was incubated at room temperature for 1 hour, then 4 µg of the supplied Proteinase K was added to terminate the reaction. Transformation of chemically competent *E. coli* cells with the LR reaction product was then carried out as described in section 2.6.8.

2.6.9 Transformation of Electrocompetent Agrobacterium Cells

2.6.9.1 Preparation of Electrocompetent Agrobacterium

An aliquot of the Agrobacterium strain AGL1 was used to inoculate 5 ml LB broth supplemented with gentamycin (to select for the T_i plasmid) and grown to culture in a 28 °C incubator shaking at 200 rpm for 16 hours. The 5 ml culture was used to inoculate 500 ml LB broth supplemented with gentamycin. This was then placed in a 28 °C incubator shaking at 200 rpm until the culture reached an OD_{600} of 0.5 to 0.8, after which it was centrifuged at 3000 x *g* for 5 minutes. The supernatant was removed and the cell pellet resuspended in 30 ml sterile, cold dH_2O and then transferred to a sterile 50 ml falcon tube. The sterile cold dH_2O wash was repeated three times, centrifuging 3000 x *g* for 5 minutes between each. After the final wash, cells were resuspended in 1 to 5 ml sterile 10% (v/v) glycerol and 100 µl aliquots were placed on dry ice before storing at -80 °C.

2.6.9.2 Transformation

500 ng of plasmid DNA was added to a 50 µl aliquot of electrocompetent Agrobacterium strain AGL1 thawed on ice. The cells were then transferred to an electroporation cuvette and pulsed at with electric current at 2500 V. The cells were then transferred to a 15 ml falcon tube, 1 ml of LB broth was added and the suspension was placed in a 28°C incubator shaking at 200 rpm for three hours. Finally, the cells were plated out on LB agar containing rifampicin to select for Agrobacterium, gentamycin (to select for the T_i-plasmid) and the appropriate antibiotic to which the plasmid construct carries resistance, and placed in a 28°C static incubator for two days.

2.6.10 Transient Transfection in *Nicotiana benthamiana*

Transformation was performed as described in Bazzini (2007): Agrobacterium carrying the construct of interest was streaked from a glycerol stock onto a freshly prepared LB agar plate supplemented with the appropriate antibiotics and grown for 2 days at 28 °C. A single colony was picked and used to inoculate a 10 ml culture of LB broth with the appropriate antibiotics, and this was incubated at 28 °C with shaking at 200 rpm overnight. The resulting culture was centrifuged 3000 x *g* for 15 minutes and the pellet was resuspended in 10 – 15 ml sterile 10 mM $MgCl_2.6H_2O$. The OD_{600} was measured and diluted down to a value of 0.2. A 1:1000 dilution of 200 mM acetosyringone was added then the culture was inverted several times and incubated at room temperature for 2 hours.

A small incision was made on the underside of a leaf of *N. benthamiana* using a razor blade, taking care to slice through only the lower epidermis and mesophyll layers. The Agrobacterium was taken up into a syringe, the syringe was pressed

over the incision, and the plunger was gently depressed to force the culture into the intercellular spaces of the leaf. Plants were then returned to the growth cabinet for 3 – 5 days to allow transfection to take place.

2.6.11 Stably Transformed Arabidopsis Lines

Wild type Arabidopsis (Col 0) was transformed using the floral dip method of (Clough and Bent, 1998). Pots containing six to eight evenly spaced *A. thaliana* Col-0 seeds were grown in long day conditions (see section 2.3.3) for five to six weeks, after which any inflorescence stems were removed to encourage growth of secondary inflorescences. The plants were ready to transform five to ten days later, when roughly one fifth of the flowers had opened. A single colony of a confirmed positive *A. tumefasciens* transformant was used to inoculate 10 ml LB broth containing appropriate antibiotics for plasmid selection and the culture was grown in a 28 °C incubator shaking at 200 rpm for 8 hours. The cultures were then transferred to 500 ml LB broth with appropriate antibiotics and left to grow in a 28 °C incubator shaking at 200 rpm overnight. When cell density reached an OD_{600} of 0.8 to 1.0, the cells were centrifuged at 4000 x *g* for 10 minutes and supernatant removed. The cell pellet was resuspended in a 5 % sucrose solution containing 0.03 % Silwet L-77 (a wetting agent), and the total volume made up to 500 ml. The cell suspension was then transferred to beakers approximately the diameter of the plant pots, and the *A. thaliana* Col-0 plants were upturned into the suspension for 1 min. Dipped plants were covered to increase humidity for 24 hours at room temperature and returned to growth room. Plants were allowed to set seed, and these were collected as normal.

2.6.12 Confocal Microscopy

The subcellular localisation of fluorescently tagged proteins was visualised using a confocal laser scanning microscope (Zeiss LSM 510). GFP tags were excited at 488 nm and emission was collected between 505 – 530 nm. RFP tags were excited at 543 nm and emission was collected between 560 – 615 nm. YFP tags were excited at 514 nm and emission was collected between 530 – 600 nm.

2.6.13 Primer Sequences

Unless otherwise stated, primers were designed in CLC Genomics Workbench v3 using the 'Design Primers' module.

2.6.13.1 Small HSP Cloning Primers

Small HSPs cloning was performed using the primers detailed in Table 2.10 in order to tag and express in *N. benthamiana*.

Table 2.10 Primers Used to Clone smHSPs

Gene	Sequence bound	Sequence
HSP17.6 (At5g12020)	5' end of promoter	caccCACATTGGCCATCAAATCC
	5' end of ORF	caccATGGATTTAGGAAGGTTTC
	3' end of ORF	AGCAACTTGAACTTGAATTGTCTTT
HSP17.6a (At5g12030)	5' end of promoter	caccAGACGATACCACTAGCTCATAC
	5' end of ORF	caccATGGATTTGGAGTTTG
	3' end of ORF	AGCGACTTGAACTTGTATAGTCTTTG
HSP17.6b (At1g53540)	5' end of promoter	caccGGCCACCAGAATACTTTG
	5' end of ORF	caccATGTCTCTAATTCCAAGC
	3' end of ORF	ACCAGAGATATCAATGGACTTAAC

2.6.13.2 Small HSP RT-PCR Primers

Levels of smHSP expression in Arabidopsis were assessed by RT-PCR using the primers detailed in Table 2.11.

Table 2.11 RT-PCR Primers Used to Distinguish Between HSP17.6 and HSP17.6a

Gene	Sequence bound	Sequence
HSP17.6 (At5g12020)	5' end of unique region (*i.e.* region with no homology to *HSP17.6a*)	GAAGACCACAACAACGAG
	3' end of ORF	AGCAACTTGAACTTGAATTGTCTTT
HSP17.6a (At5g12030)	5' end of unique region (*i.e.* region with no homology to *HSP17.6*)	CCTGAAGAACAAACCGAG
	3' end of ORF	AGCGACTTGAACTTGTATAGTCTTTG

2.6.13.3 Small HSP Knockout Genotyping Primers

Knockout lines of *hsp17.6* and *hsp17.6a* were obtained and PCR was carried out to determine genotyping using the primers detailed in Table 2.12.

Table 2.12 Primers Used to Genotype Putative smHSP Knockouts

Gene	Sequence bound	Sequence
hsp17.6 K/O (At5g12020)	Region 5' of T-DNA insertion in line N507510	GAAGACCACAACAACGAG
	Region 3' of T-DNA insertion in line N507510	AGCAACTTGAACTTGAATTGTCTTT
hsp17.6a K/O (At5g12030)	Region 5' of T-DNA insertion in line N572448	CCTGAAGAACAAACCGAG
	Region 3' of T-DNA insertion in line N572448	AGCGACTTGAACTTGTATAGTCTTTG

2.6.13.4 *MYB64* Regulon qPCR Primers

Table 2.13 details the primers used to validate the reported upregulation of a suite of genes in response to *MYB64* overexpression (the *MYB64* 'regulon') as reported by next generation sequencing.

Table 2.13 Primers Used for qPCR Confirmation of MYB64 Regulon Upregulation

Gene	Name	Sequence	Source
MYB64 (At5g11050)	5g11050 f q new	GGACAGCCGATGAAGACAGGAAG	*n/a*
	5g11050 r q new	AGTCTTCGAGAATGCGCGGC	
ERF11 (At1g28370)	At1g28370qPCR f	CCAACGAAGGTAACGGAG	*n/a*
	At1g28370qPCR f	CTCCACGAAACTCAATAGCAC	
MPK11 (Splice variant 1) (At1g01560.1)	01560.1 q f new	CATTCACGGTTCTTCCTG	*n/a*
	01560.1 q r new	GTGAAAAGTTTACCTCAGTGATGAGTC	
MPK11 (Splice variant 2) (At1g01560.2)	01560.2 f q new	AACAGACGCATTACAGTCGATGAAGCC	*n/a*
	At1g01560.2qPCR r	AATCAAAATGGAACGGTCTC	
CBF1 (At4g25490)	At4g25490qPCR f	AGAGCCAAACAAGAAAACCA	*n/a*
	At4g25490qPCR r	GATATCCTTGGCGCATGT	
PP2C (At5g59220)	12020_for_unique	GAAGACCACAACAACGAG	*n/a*
	12030_rev_pENTRD	AGCAACTTGAACTTGAATTGTCTTT	
CZF2 (At5g04340)	12030_for_unique	CCTGAAGAACAAACCGAG	*n/a*
	12030_rev_pENTRD	AGCGACTTGAACTTGTATAGTCTTTG	
PP2A-A3 (At1g13320)	At1g13320 qFwd	TAACGTGGCCAAAATGATGC	(Czechowski et al., 2005)
	At1g13320 qRev	GTTCTCCACAACCGCTTGGT	
Actin 2 (At3g18780)	Act2Sens4 fwd	CTAAGCTCTCAAGATCAAAGGCTTA	(Love et al., 2005)
	Act2Sens4 rev	ACTAAAACGCAAAACGAAAGCGGTT	
UBQ10 (At4g05320)	At4g05320 qFwd	GGCCTTGTATAATCCCTGATGAATAAG	(Czechowski et al., 2005)
	At4g05320 qFwd	AAAGAGATAACAGGAACGGAAACATAGT	

2.6.14 Isolation of Protein from Plant Tissue

A razor blade was used to excise a mature section of leaf approximately 3 inches long. This section was placed into a microfuge tube along with a sterile ball bearing and snap-frozen in liquid nitrogen. Tissue samples were ground using the TissueLyser automated grinder (Qiagen, West Sussex, UK) for 1 minute at 25 rpm. In a modified version of the extraction and precipitation protocol described by Jacobsen and Shaw (1989), samples were removed to ice and 1 ml of extraction buffer (20 mM Tes-KOH pH 8.0, 0.5 M NaCl, 1 protease inhibitor cocktail tablet (Roche catalog number 11836153001) per 10 ml of buffer) was added quickly before the tissue sample was able to thaw. Samples were mixed by vortexing and ball bearings were removed. Debris was pelleted by

centrifugation at 10,000 x *g* for 10 minutes at 4 °C, and the supernatant was decanted to a fresh 15 ml tube. 10 volumes of acetone were added to precipitate proteins. Samples were incubated at -20 °C for 8 – 12 hours to aid precipitation. Proteins were pelleted by centrifugation at 3000 x *g* for 5 minutes at 4 °C. Acetone was removed by pipetting and the pellets were allowed to almost completely dry at 50 °C for 10 minutes. Pellets were then resuspended in 500 µl dH_2O and stored at – 20 °C.

2.6.14.1 Quantification of Protein Samples

Protein samples were quantified using the Bradford assay (Bradford, 1976). Briefly, standard solutions of bovine serum albumin (BSA) were prepared to concentrations of 0, 0.25, 0.5, 0.75, 1.0, and 1.4 mg/ml to facilitate the construction of a standard curve. Samples of unknown concentration were prepared in a final volume adjusted to 33 µl, then this 33 µl was added to 1 ml Bradford Reagent (BioRad, USA) and mixed by gentle inversion. Absorbance was measured at 595 nm after an incubation of 5 minutes, blanked against a mixture of 33 µl of water and 1 ml of Bradford Reagent. If the concentration of the unknown samples was found to be outwith the range of the standard curve, the protein:dH_2O ratio in the 33 µl sample was adjusted and the sample reanalysed.

2.6.15 SDS-Polyacrylamide Gel Electrophoresis (PAGE) of Proteins

SDS-PAGE was carried out on the Protean Mini gel kit from BioRad (Hercules, USA). Resolving gels were made according to Sambrook and Russell (2001) to an acrylamide concentration of between 7.5 and 15 % (depending on the molecular

weight of the proteins to be resolved) (0.375 M Tris HCl pH 8.8, 0.1 % SDS, 0.1 % ammonium persulfate and 0.1 % TEMED) and immediately poured into casts. A thin layer of isopropanol was laid over the top to ensure an even surface. Once the resolving gel had set a stacking gel was made to a concentration of 5.7 % acrylamide (0.125 M Tris HCl pH 6.8, 0.1 % SDS, 0.1 % ammonium persulfate and 0.1 % TEMED). The isopropanol was removed and the stacking gel was poured on top, then a plastic comb was inserted to create wells. After the stacking gel had set, the plastic comb was removed and the wells washed out with water, before assembling the gel running kit, after which enough 1x running buffer (20 mM Tris base, 0.195 M glycine, 0.1 % SDS, pH 8.3) was added to the protein tank to cover both electrodes. Equal amounts of protein were added to each well and a broad range pre-stained protein marker was run alongside samples. The acrylamide gel was electrophoresed at 65 V until the proteins reached the end of the stacking gel, after which the voltage was increased to 135 V until the end of the gel was reached.

2.6.16 Silver Staining of SDS-PAGE Gels

SDS-PAGE gels were placed into a clean plastic dish and bathed in fixing solution (ethanol : glacial acetic acid : dH_2O at a ratio of 3:1:6) with gentle shaking action for 8 – 12 hours. The gel was rinsed five times with dH_2O. Freshly prepared silver stain was added (10.5 ml NaOH at 0.36% (w/v), 700 µl NH_4OH at 14.8 M, and 0.4 g silver nitrate) and shaking continued for 15 minutes. The silver stain was removed and the gel washed for 1 hour in dH_2O with frequent changes. The dH_2O was removed and developing solution was added (120 µl 1 % citric acid, 250 µl 38 % formaldehyde, 500 ml dH_2O). The gel was observed for the development of colour, then the reaction was quenched with 1 % acetic acid.

2.7 Illumina Next-Generation Sequencing (Transcript Profiling)

Total RNA was prepared from mature plant leaves (see sections 2.6.1.3 - 2.6.1.5) and supplied to the operators of the Sir Henry Wellcome Functional Genomics Facility (SHWFGC, University of Glasgow). Further mRNA quality assessment and quantification was carried out by the technicians of this facility, as well as preparation of cDNA and subsequent preparatory treatments. These included the fragmentation of cDNA, ligation of linkers containing restriction enzyme recognition sites which recruit enzymes that cut 35 bp from the site, and ligation of a second linker to the resulting sticky ends.

These linker-modified cDNAs were attached to the surface of a glass channel and subjected to several rounds of bridge-amplification facilitated by the addition of primers complementary to the linker sequences. This resulted in clusters of identical molecules spread throughout the channel, each representing a fragment of an RNA molecule. The chip was processed in an Illumina Genome Analyzer (Illumina, UK) which floods the channels with successive washes of dye-labelled nucleotide triphosphates in a process of reversible chain-termination. After each round, a computerised laser scans the channel and builds a picture of fluorescence at each randomly-positioned cluster and, by comparing the images from one round to the next, records the sequence of each original molecule.

The output of this process was a list of between 5 million and 10 million individual 35 bp-long reads from which the linker sequences were removed by the technicians at the SHWFGC. This output list was then aligned to an Arabidopsis reference genome in order to calculate expression values for each

gene to allow comparison of global expression between the wild type and the 35Spro:MYB64 sample. For a fuller account, see section 4.2.

2.7.1 Alignment of NGS Output Data to Reference Genome

The latest edition of the Arabidopsis genome sequence at the time of operation was the TAIR9 Genome Release available online at The Arabidopsis Information Resource (TAIR) website (available at www.arabidopsis.org). Alignment of the Next Generation Sequencing (NGS) output dataset to the reference genome was carried out using CLC Genomics Workbench v3. The reference genome was provided with annotations of known transcripts from which CLC Genomics Workbench was able to interpret those NGS reads which aligned partially to the end of one exon and partially to the start of another. The number of reads aligned with each gene model was used to calculate an expression score for that gene, expressed as Reads Per Kilobase per Million mapped reads (RPKM). The first calculation, 'reads per kilobase', compensates for differences in transcript level attributable purely to differences in gene length (*i.e.* mRNA representing longer genes would be expected to produce a greater number of cDNA fragments); the second calculation, 'per million' mapped reads, compensates for differences in input RNA abundance (*i.e.* more concentrated RNA preparations would lead to higher expression values in that sample). Reads that aligned to more than 3 different loci were treated as too ambiguous to interpret and were discarded. Those that aligned to 3 or fewer loci, which represented < 0.5 % of the total remaining viable reads, were included in the calculation of expression of all three genes.

3 Function of Three Small Heat Shock Proteins Regulated by *MYB64* Under Salt Conditions

3.1 Introduction

The microarray results obtained by Price (2005) and discussed in section 1.4 revealed that when halotolerant activation-tagged line HT5, overexpressing *MYB64*, was exposed to long-term moderate salt stress, 15 % of the 45 transcripts with the greatest increase in abundance relative to wild type had a functional annotation related to thermotolerance. Indeed, the first four in the list were members of the heat shock protein (HSP) families (Table 1.1). Two of the four, *HSP70* (At3g12580) and *HSP101/ClpB1* (At1g74310), are well-studied and are responsible for the ATP-dependent refolding of denatured proteins (Mogk et al., 2003, Tonsor et al., 2008). *HSP17.6a* (At5g12030) and *HSP17.6b* (At1g53540) are less well-studied and belong to a subgroup known as small HSPSs (smHSPs) characterised by high sequence similarity to each other and a molecular weight of approximately 17 – 20 kDa (Waters, 2013). HSP17.6a has a chromosomal neighbour, HSP17.6 (At5g12020), and while this was not among the most highly upregulated transcripts it was considered worth investigated as it was likely to have arisen as a gene duplication and might be a functional homolog.

Presented with the choice of investigating the roles of either *HSP70* / *HSP101* or the smHSPs, the latter offered the potential for a more original contribution to the field. This choice was made even more appealing by virtue of the fact that

in addition to these two smHSPs at the top of the rankings, the top 45 transcripts also included a further 3 members of the smHSP family.

3.2 Predicted Structures of smHSPs of Interest

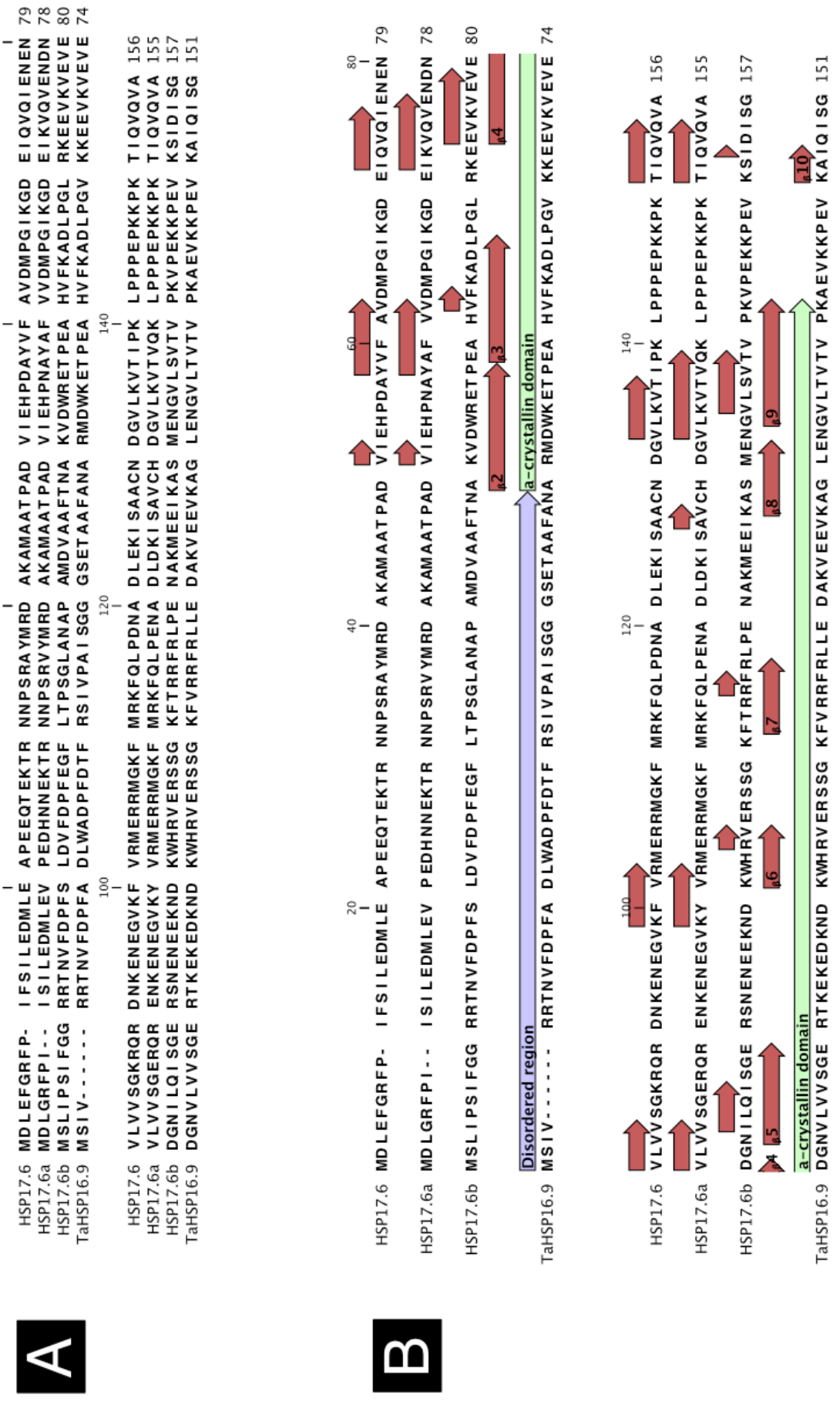
In order to make predictions about their 3D structure, the protein sequences of Arabidopsis *HSP17.6*, *HSP17.6a*, *HSP17.6b* were aligned with that of the resolved wheat homolog *TaHSP17.6* using the Protein Alignment tool in CLC Genomics Workbench v.3 using default parameters (Figure 3.1). Annotations above the wheat sequence represent ß-strands (red arrows) that have been resolved by X-ray crystallography. Equivalent annotations above the Arabidopsis sequences also represent predicted ß-strands.

The predicted and observed ß-strands occupy almost identical regions of each protein indicating that this secondary structure is highly conserved both within species and between these species. The α-crystallin domain of *HSP17.6b* and the ß-strands therein are noticeably shifted several residues to the right relative to the other three sequences. This may represent a small N-terminal insertion/duplication event during the evolutionary history of this family member.

Another measure of sequence similarity and of predicted structure and function is the degree of hydrophobicity of the residues at each point along a polypeptide. The hydrophobicity of the same four smHSPs were compared using the Kyte-Doolittle Hydropathy Plot tool in CLC Genomics Workbench v.3. In this particular analysis each residue is assigned a hydrophobicity score of between 4.6 (most hydrophobic) and -4.6 (least hydrophobic; Kyte and Doolittle, 1982).

Figure 3.1 Protein Sequence Comparison of Arabidopsis HSP17.6, HSP17.6a, HSP17.6b, and Wheat TaSHP16.9

The crystal structure of wheat (*Triticum aestivum*) TaHSP16.9 has been resolved previously (van Montfort et al., 2001), so protein sequences of the three Arabidopsis smHSPs of interest were aligned to this for comparison. **A** Simple sequence alignment of all four protein sequences without annotations for ease of comparison. **B** The same sequence alignment showing β-strands represented by red arrows (predicted for Arabidopsis sequences, resolved from crystal structure for the wheat sequence, and numbered according to established nomenclature (Waters et al., 1996). The disordered N-terminal region and the α-crystallin domain of the resolved wheat structure are represented by purple and green arrows, respectively.


A
B
HSP17.6
HSP17.6a
HSP17.6b
TaHSP16.9
Disordered region
a-crystallin domain

Scores for the polypeptide are averaged over windows of 11 residues and plotted.

The hydrophobicity plot in Figure 3.2 shows three features. The first is that the sequence similarity noted in the Figure 3.1 across the conserved α-crystallin domain is reflected in overlapping hydrophobicity plots, including the C-terminal shift in *HSP17.6b*, and that the disordered N-terminal regions are more variable. The second is that the *HSP17.6* and *HSP17.6a*, which are encoded at adjacent loci on chromosome 5 and thus would appear to have arisen as a result of a gene duplication event, share an almost identical hydrophobicity signature in their disordered N-terminal regions, further supporting this duplication hypothesis. The third is that a large majority of each protein sequence is hydrophilic (*i.e.* below 0 on the y-axis), suggesting a high propensity for contact with the surrounding solvent. At the time of performing this analysis the literature suggested that these smHSPs were cytosolic due to the fact that there were no known localisation signals and there had been no experimental evidence to suggest that they were located in any particular sub-cellular compartment; this hydrophobicity plot does not contradict the conclusion that they belong to a cytosolic subgroup.

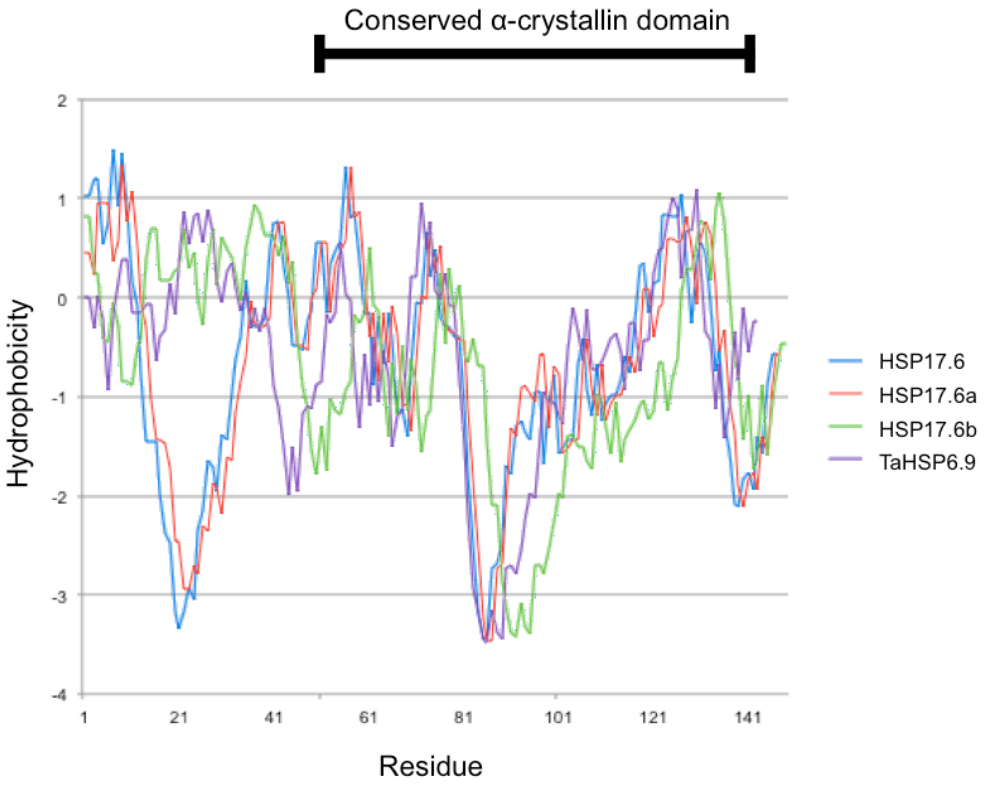


Figure 3.2 Kyte-Doolittle Hydrophobicity Plot of Arabidopsis HSP17.6, HSP17.6a, HSP17.6b, and Wheat TaHSP16.9

Hydrophobicity of each of the Arabidopsis smHSPs of interest and the 16.9kDa homolog from wheat (*Triticum aestivum*) was calculated and plotted on the same axis. All four proteins are predicted by their families to be cytosolic: HSP17.6 and HSP17.6a belong to Class II; HSP17.6b and HSP17.6b belong to Class I. Positive values indicate hydrophobic regions; negative values indicate hydrophilic regions. Windows of 11 residues were used to calculate the hydrophobicity of each region. (All *in silico* operations were carried out using the Kyte-Doolittle Hydropathy Plot tool in CLC Genomics Workbench v.3 using default parameters (moving average of 11 residues).

3.3 Expression of smHSPs of Interest

3.3.1 Promoter Analysis

The simplest explanation for the high upregulation of *HSP17.6a* and *HSP17.6b* in the microarray results would be that MYB64, being a transcription factor, directly activates their expression. To investigate this theory the promoter sequences of these two smHSP genes were interrogated for the presence of transcription factor binding sites. There was no experimental evidence in the literature to suggest a target sequence bound by MYB64, so the promoter sequences of these smHSPs were investigated for commonalities that might serve such a purpose. First, a *de novo* analysis was carried out using the Athena promoter analysis tool (O'Connor et al., 2005) to look for any other known regulatory elements. The 1kb upstream of the transcription start site (TSS) of each gene was notionally defined as the promoter for the purposes of this analysis. Had *HSP17.6* and *HSP17.6a* arisen from a duplication event intact, it is possible that the duplicated region might also have included upstream regulatory sequences.

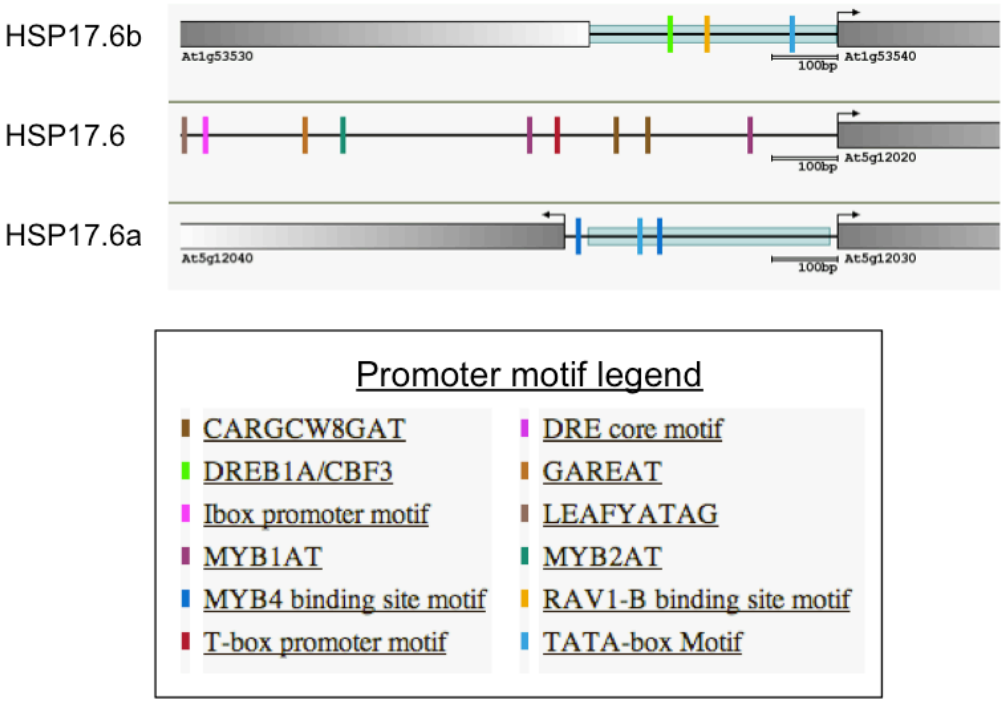
Athena is a web-based tool that holds information on the Arabidopsis genome sequence and allows the upstream region of each annotated gene to be checked for the presence of any known promoter binding sites. If several genes are examined simultaneously, the degree of enrichment (or otherwise) in the gene set of each binding site is calculated relative to the detection level across all of the promoter regions in the entire genome. The output also displays a probability value of finding such a result for each binding site identified.

Figure 3.3, **panel A**, shows such an output displaying the locations of known promoter motifs (coloured vertical bars) relative to the TSS of each gene (grey horizontal boxes at the extreme right of each row). CpG islands are also indicated (pale blue horizontal boxes). The TATA-box motif is found in the notional promoter region of both *HSP17.6a* and *HSP17.6b* and these two genes also have CpG islands just upstream of their TSSs, but there are no other common motifs across the three sequences.

Figure 3.3, **panel B**, shows the statistical significance of finding each motif in the given promoter set. The motif closest to being identified as significantly enriched is DREB1A/CBF3 (found in one of these three notional promoter regions; found in 5 % of promoter regions in the genome; p value 0.166) but with only three promoters to analyse it is perhaps difficult to draw conclusions from this. The MYB1AT motif was originally identified in the promoter of the dehydration- and salt-responsive gene *RD22*. The MYB2AT motif is bound by MYB2 which responds to water stress. The fact that these motifs are present but not significantly enriched does not preclude the possibility that they are nevertheless functional. Regions up to 3kb from each TSS were also investigated (not shown) but no enriched motifs were found.

After finding no enriched promoter motifs from a database of those already known, the existence of novel binding sites was investigated. The three notional promoter regions were aligned using the Sequence Alignment tool in CLC Genomics Workbench v3 using default parameters. Working on the hypothesis that a binding site for the MYB64 protein would be found at equivalent distances from each TSS, no such sequences were evident from the alignment (not shown). Taken together, these results indicate that regulation of these genes is either

A



B

Transcription Factor/Motif Name	prom's bound in subset		prom's bound in genome		p-value
TATA-box Motif	66%	2	75%	22707	0.849
MYB4 binding site motif	33%	1	64%	19250	0.953
T-box promoter motif	33%	1	45%	13820	0.842
Ibox promoter motif	33%	1	32%	9831	0.695
MYB1AT	33%	1	73%	22181	0.981
GAREAT	33%	1	47%	14315	0.856
LEAFYATAG	33%	1	8%	2489	0.228
RAV1-B binding site motif	33%	1	9%	2807	0.254
DREB1A/CBF3	33%	1	5%	1767	0.166
MYB2AT	33%	1	23%	7097	0.554
CARGCW8GAT	33%	1	53%	16208	0.902
DRE core motif	33%	1	18%	5550	0.457

Figure 3.3 Promoter-motif Analysis of smHSP Genes of Interest

A The notional promoter regions - *i.e.* the 1kb regions upstream of the transcription start site (TSS) - of each of At1g53540, At5g12020 and At5g12030 (numerical order by gene number, encoding HSP17.6b, HSP17.6 and HSP17.6a, respectively) were interrogated for statistical enrichment of any known promoter motifs using the online Athena tool (O'Connor et al., 2005). Motifs are represented by vertical coloured bars on the horizontal genomic DNA diagram. Several known motifs were identified, though none are present in all three of the notional promoter regions. **B** Statistical interpretation of the data on each individual motif: column 2 shows the number of promoters in the dataset containing that particular motif (n = 3); column 3 shows the number of promoters containing that motif across the complete Arabidopsis genome; column 4 shows the significance of the enrichment of each promoter within this dataset. No known motifs were present in all 3 genes of interest.

not directly mediated by binding of MYB64 to a *cis*-acting element, or that such binding occurs at cryptic sites more distal to the TSS of each gene.

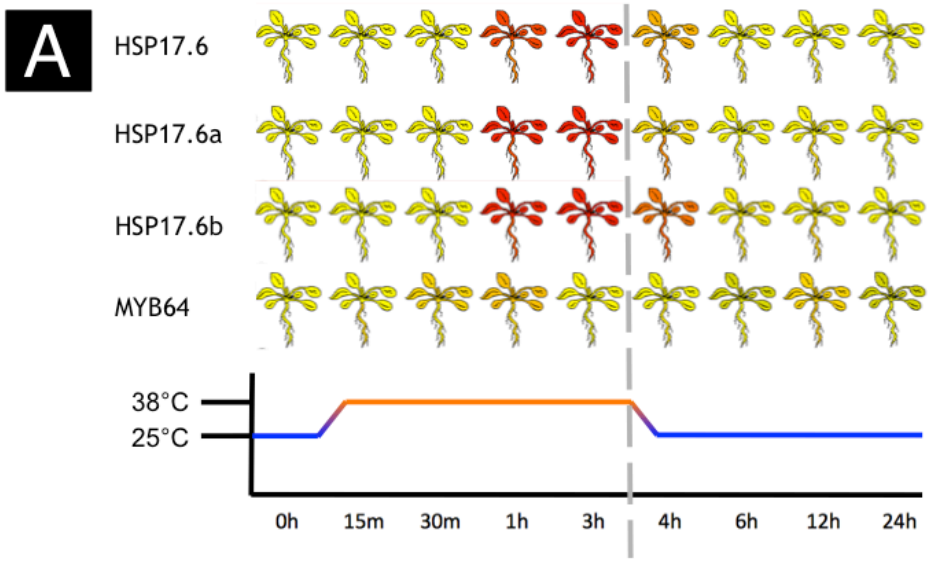
3.3.2 Data Mining of Open-Access Heat Acclimation Transcript Profiles

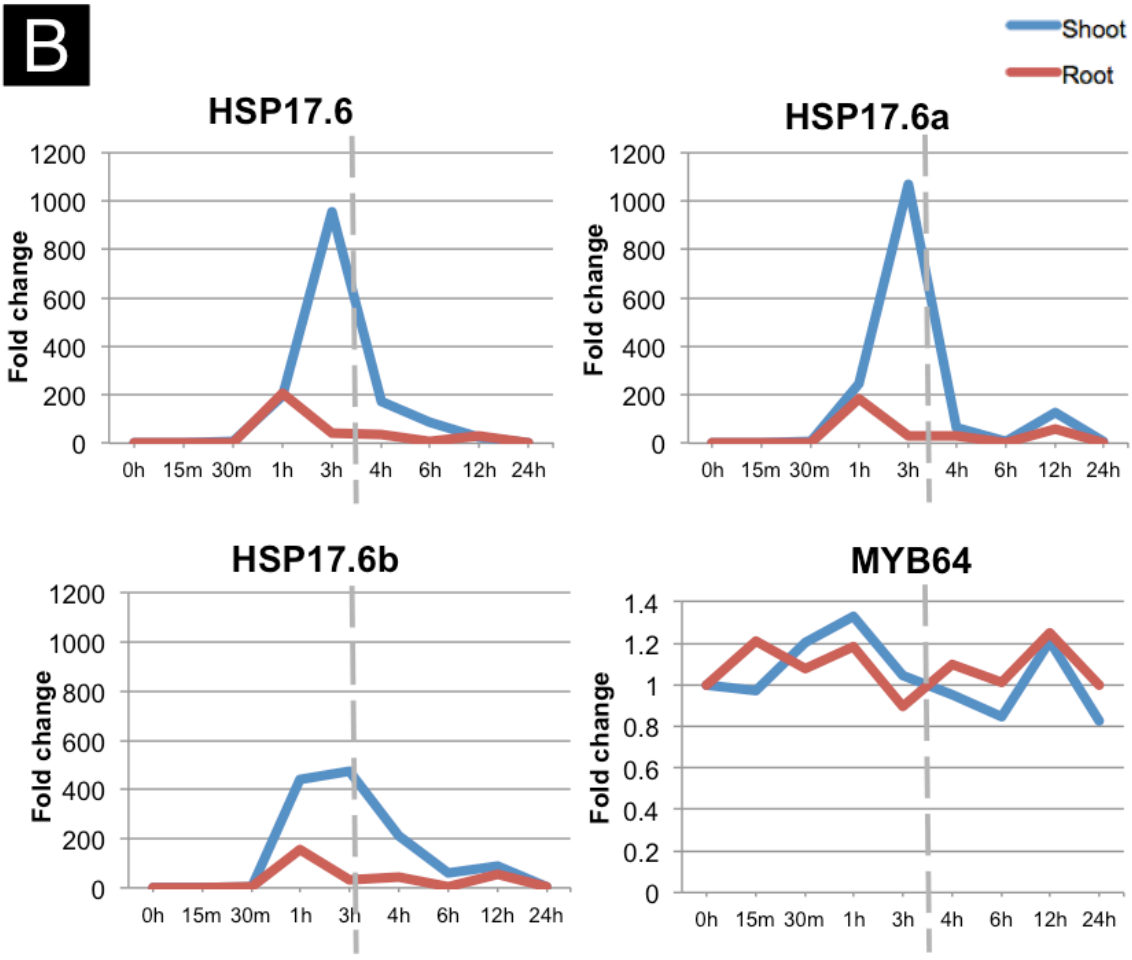
Transcript profiles from plants under a variety of experimental conditions are held at several online repositories, most commonly in the form of microarray results, in order that researchers can interrogate these databases for expression data without repeating the experiment themselves. The Arabidopsis "electronic Fluorescent Pictograph" (eFP) Browser (Winter et al., 2007) is an online tool which allows visualisation of, among others, the dataset produced by Kilian et al. (2007) as part of the AtGenExpress project. This project involved exposing seedlings to a variety of abiotic factors including heat, cold, light, UV, wounding, and chemicals. Briefly, seedlings given the heat treatment were grown in a long-day cycle at normal temperatures (24 °C) for 11 days then exposed to a high temperature (38 °C) for three hours, then returned to the growth room at 24 °C. RNA was isolated at various timepoints and hybridised to the GeneChip® Arabidopsis ATH1 Genome Array (Affymetrix, High Wycombe, UK) and expression levels were measured in duplicate. The high temperature used by Kilian et al. was very close to the temperature used in thermal acclimation experiments elsewhere in this thesis (37 °C), so it will be referred to as an acclimation treatment.

Figure 3.4, **panel A**, shows a heat-map representation of expression of the three smHSPs of interest and of *MYB64*. Figure 3.4, **panel B**, shows the fold-change in expression measured at each timepoint and in each tissue (shoots or roots)

Figure 3.4 Analysis of Open-Access Transcript Profiles of the Three smHSPs and of MYB64 Upon Thermal Acclimation

Data deposited online at the eFP browser (Winter et al., 2007) showed the temporal expression pattern of these four genes of interest before, during, and after an acclimation period. **A** False-colour representation of expression in shoots and roots. **B** Fold-change relative to non-acclimated seedlings. Seedlings were heated to the acclimation temperature (38 °C) for 3 hours then returned to normal conditions (22 °C) for a further 21 hours. Control seedlings were kept under normal conditions. RNA was extracted at the given timepoints and expression was analysed by microarray. Blue trend lines: shoots. Red trend lines: roots. Dashed grey vertical line: transition from acclimation at 38 °C back to the growth room at 25 °C.


A
HSP17.6
HSP17.6a
HSP17.6b
MYB64
38°C
25°C
0h
15m
30m
1h
3h
4h
6h
12h
24h


B
Shoot
Root
HSP17.6
HSP17.6a
HSP17.6b
MYB64
Fold change
0h 15m 30m 1h 3h 4h 6h 12h 24h

relative to the initial measurement of each transcript. *HSP17.6b*, the gene most highly upregulated in the salt-stressed activation-tagged line, HT5, underwent a 439-fold increase in shoots by the first hour and slightly increased to reach 474-fold after three hours. Three hours into the acclimation period *HSP17.6a* reached a peak of approximately 1000-fold upregulation in shoots and 200-fold in roots. *HSP17.6*, its close homolog adjacent on chromosome 5 (which was not identified in the HT5 microarray results), reached almost identical peaks. Expression of all three smHSP genes in shoots declined to approximately 200-fold or less following an hour of recovery at the control temperature of 24 °C. *HSP17.6a* and *HSP17.6b* both underwent a further transient increase in expression 12 hours after the initial exposure to the acclimation temperature (*i.e.* 9 hours after removal to the control temperature) which declined to basal levels by 24 hours.

MYB64 was investigated because of its connection with the upregulation of the smHSPs in line HT5. Expression did not change here by more than 1.4-fold and the fluctuations follow no pattern, so it can be concluded that there was no effect of heating on the level of *MYB64* transcription in this experiment.

Expression of these three smHSPs is clearly induced by the application of high temperature, and these data disagree with the hypothesis that this induction is mediated by a cascade that begins with transcription of *MYB64*.

3.3.3 Experimental Investigation of smHSP Levels in Wild Type upon Thermal Acclimation

3.3.3.1 Transcript Levels

The expression patterns described above were carried out by microarray analysis in a slightly different experimental system than the one to be used in the experiments carried out and reported in this thesis (seedlings floating on rafts rather than sown on solid agar medium; different age at sampling; different acclimation temperature) so it was of interest to confirm that these patterns would also be observed in transcript profiles performed in-house. Seedlings were grown on agar plates for two weeks (section 2.3.2) before being transferred to the incubator at 37 ± 0.5 °C for 3 hours. RNA was sampled at the intervals shown in Figure 3.5 and semi-quantitative RT-PCR was performed (sections 2.6.1 and 2.6.3) using primers to *HSP17.6* and *HSP17.6a*, and *Actin 2* as a 'housekeeping' reference.

As shown in Figure 3.5 *HSP17.6* and *HSP17.6a* were both undetectable before acclimation and were both massively induced by 3 hours at 37 ± 0.5 °C. *HSP17.6* attained a higher maximum expression level. The general trend was a decline in the abundance of both transcripts over the following 24 hours though, notably, *HSP17.6* increased slightly again at 12 hours, which was not mirrored in *HSP17.6a* levels. This is consistent with the microarray data discussed above.

3.3.3.2 Protein Levels

Up to this point, all discussion of smHSP expression has pertained to transcriptional changes. It was of interest to discern whether the acclimation

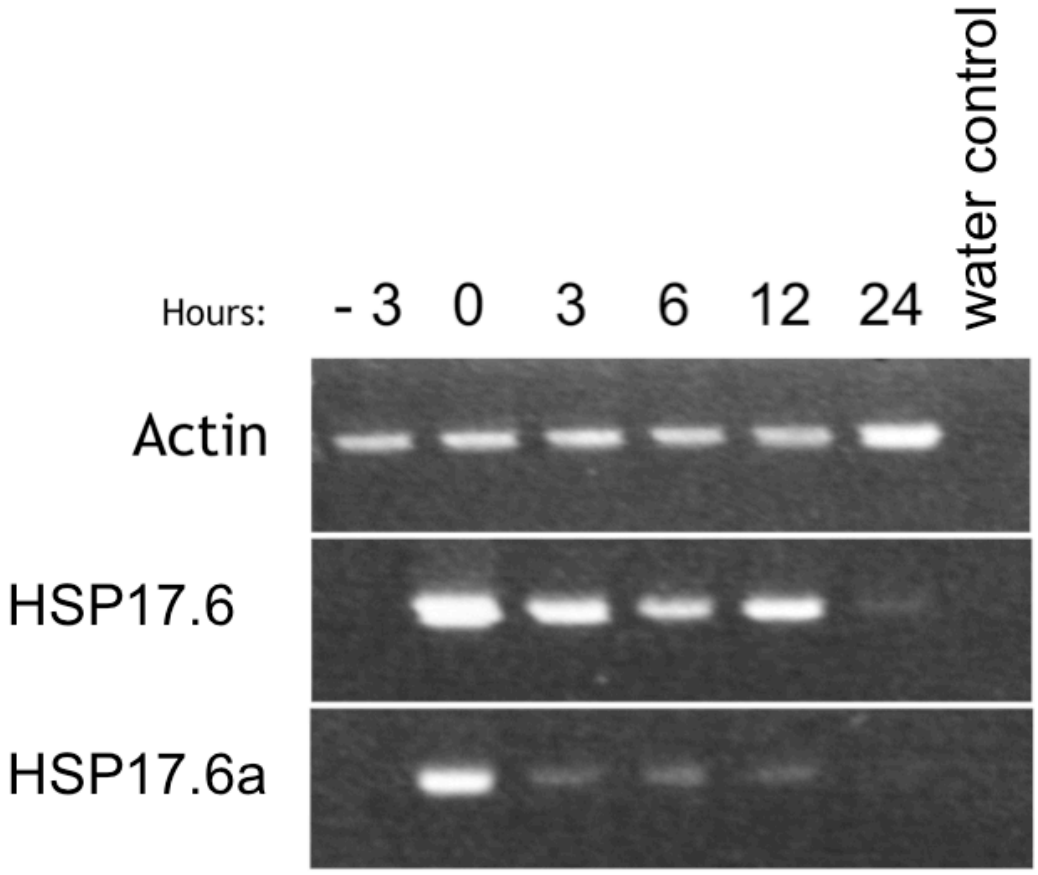


Figure 3.5 Acclimation-induced Accumulation of HSP17.6 and HSP17.6a in Wild Type Seedlings

Wild type (Col-0) plate-grown 2-week-old seedlings were given a 3-hour incubation at the acclimation temperature, 37 ± 0.5 °C, and RNA was sampled before (Pre), immediately after (0), and at 3, 6, 12 and 24 hours later. RT-PCR was carried out to measure transcript levels of two of the smHSP genes of interest. Expression of both HSP17.6 and HSP17.6a was highest immediately upon removal from the incubator and declined steadily over the subsequent 24 hours, thus confirming the open-access microarray measurements presented in Figure 3.4.

treatment described above also resulted in an increase in detectable smHSP proteins.

To monitor general smHSP protein levels upon thermal acclimation, wild type (Col 0) seedlings were sown on agar plates (section 2.3.2) for 2 weeks then transferred to the incubator at 37 ± 0.5 °C for 3 hours. This is widely used in the literature as an acclimation temperature and smHSPs are reported to reach as much as 5% of total cellular protein (Vierling, 1991), so it was reasonable to expect to be able to see a band appear at approximately 17 kDa when the smHSPs were expressed. Whole seedlings were harvested before (Pre) and immediately after acclimation (0 h), then at 3, 6, 12 and 24 h following return to the growth room at 22 °C. Samples were then applied to an SDS-PAGE gel (Figure 3.6). At 0 h following heat acclimation, *i.e.* at the timepoint when the transcript levels reached their maximum levels (see Figure 3.4 and Figure 3.5), a band was evident at around 17 kDa. This faint band persisted at a slightly lower level through the 3 and 6 hour timepoints. At 12 hours post-acclimation this band exhibited a marked increase in intensity which was also observable at 24 hours. This coincides with the transient re-elevation of transcript levels discussed above.

3.3.4 Generation of Chimeric smHSPs

In order to more accurately track expression of these smHSPs at the protein level, particularly in the absence of specific antibodies, constructs were designed to add N-terminal and C-terminal epitope and fluorescent tags. Adding epitopes would also facilitate immunoprecipitation of smHSP:substrate complexes in any future experiments that might be carried out to characterise

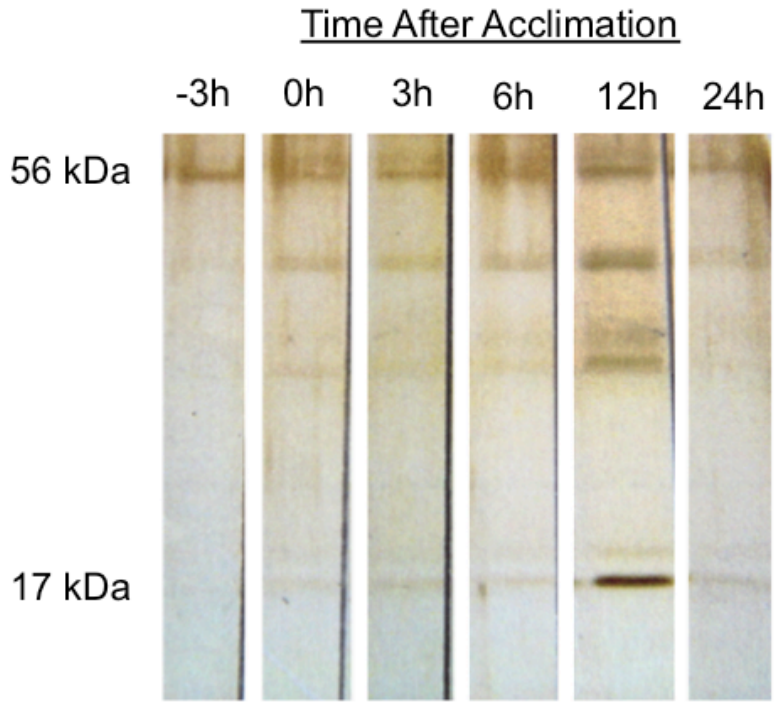


Figure 3.6 smHSP Protein Levels After Thermal Acclimation

Wild type (Col-0) plants were grown on basal agar plates in a growth room at Normal Conditions for 14 days then incubated at acclimation temperature (37 ± 0.5°C) for 3 hours before being returned to the growth room. Protein was harvested before acclimation (-3h), immediately after (0h), and 3, 6, 12 and 24 hours after the return to the growth room at 22 °C. Samples were separated by SDS-PAGE and silver stained. The large RuBisCo subunit (~ 56 kDa) is annotated as a loading control.

the differences between the range of targets of specific members of the smHSP family.

The first decision to be made was whether to add tags to either the N- or C-terminus of the proteins. Indications from the literature suggested that the N-terminus is involved in client binding and in oligomer formation (Kim et al., 1998, van Montfort et al., 2001), so it was reasoned that it was perhaps more important for function than the C-terminus. In order to increase the likelihood of preserving biological function it was decided that the fusions should therefore be made at the 3' ends of the genes.

The second decision concerned which tags to select. For experiments relating to spatial and temporal patterns of expression, it was decided that fluorescent polypeptides would provide the simplest measure of gene activation. Vectors containing GFP, RFP and YFP sequences were obtained. For potential experiments designed to elucidate client specificity, epitope tags for which there are commercially-available antibodies would allow smHSP:substrate complexes to be immunoprecipitated. Vectors encoding haemagglutinin (HA), FLAG and MYC tags were also obtained.

The final decision regarded selection of appropriate promoters. The 1kb region upstream of each gene had already been investigated and no common activation sequences were found to be shared by all three smHSPs (see section 3.3.1), but this did not preclude the possibility that some cryptic *cis*-acting elements within these regions were still responsible for driving expression of the genes in response to acclimation temperatures. Constructs under the control of these 1kb 'own-promoter' regions might be useful in assessments of endogenous temporal patterns of expression. The constitutive 35S promoter, on the other hand, would

be useful in sub-cellular localisation experiments where it was not important to find physiologically representative conditions to stimulate expression, but simply to determine which compartment each smHSP would occupy when expressed. Both own-promoter and 35S promoter constructs were therefore planned.

Thus, each smHSP was cloned into a variety of plasmid vectors to create fusion constructs (NB: 'x' in the following nomenclature is variable and represents slightly different vectors and gene combinations, each with slightly different properties. For a full list see Table 2.2): pGWBxx :: HSPxx : xFP constructs to constitutively express fluorescently-tagged smHSPs under the control of the 35S promoter; pGWBxx :: HSPxxpro : HSPxx : xFP constructs to express the same chimeric proteins under the control of 'own-promoters'; and px7xWG2 :: HSPxx : epitope constructs to constitutively express various epitope-tagged versions of the smHSPs (Figure 3.7 **panels A**, **B**, and **C**, respectively). Experiments performed with the fluorescently tagged proteins are described below. The epitope-tagged constructs were generated but immunoprecipitation and substrate identification experiments were not completed due to time constraints.

With so many combinations of vectors and genes it seemed unreasonably time-consuming to transform wild type Arabidopsis, select, screen, genotype, and finally perform the relevant experiments, especially without first knowing how the biological function of such a relatively small protein might be affected by a comparatively large tag. Rather than creating transgenic Arabidopsis lines, these constructs were used in transient expression experiments in *N. benthamiana* leaves.

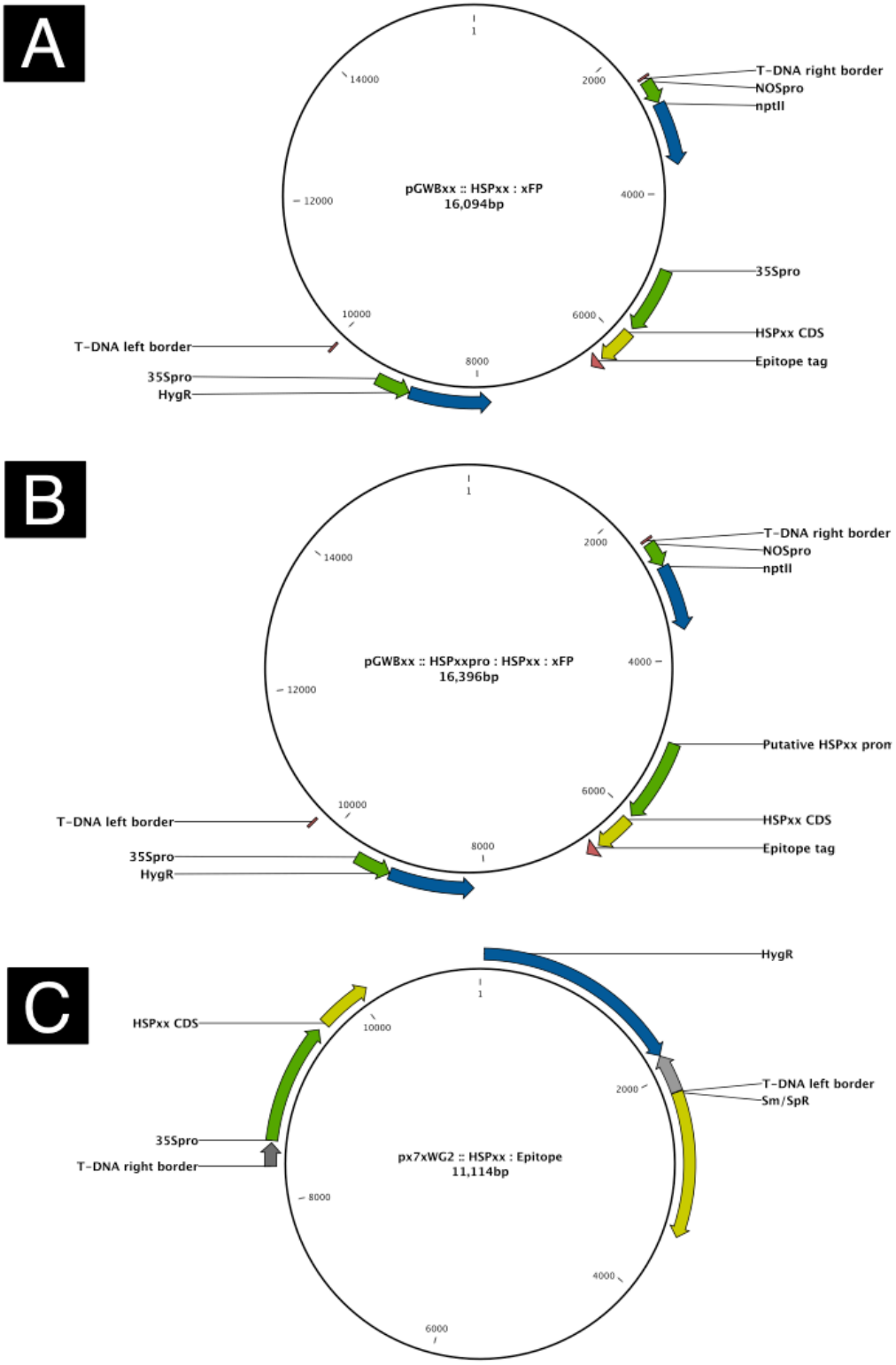


Figure 3.7 Generic Plasmid Maps of Chimeric smHSP Constructs

smHSPs were cloned into a series of binary vectors with a combination of promoters and C-terminal tags. The pGWBxx series of vectors add a variety of epitope tags for use in immunological assays. A suite of 35Spro (**A**) and putative smHSP-promoter region (**B**) constructs were made. The px7xWG2 series of vectors already carry the 35S promoter and allow the addition a variety of fluorescent tags for use in imaging assays (**C**). For a full description of the vectors used, see Table 2.2.

3.3.5 Temporal Expression of smHSP Proteins

N. benthamiana leaves were infiltrated with Agrobacterium carrying each of the fluorescently tagged 'own-promoter' constructs: HSP17.6pro:HSP17.6:XFP; HSP17.6apro:HSP17.6a:XFP; and HSP17.6bpro:HSP17.6b:XFP; where 'XFP' represents either RFP, YFP, or GFP. All three combinations of fluorescent tag were tested with all three genes. Infiltrations were performed separately. Plants were allowed 36 hours at Normal Conditions for transient transfection to take place. Leaves were then held at the acclimation temperature of 37 ± 0.5 °C for 3 hours controlled by a thermal PCR block. A block of metal (approximately 0.5 cm thick) was laid over the heat block normally used to incubate PCR tubes, leaves were wrapped in cling film to prevent transpirational cooling and laid on top of this metal block, then a piece of thermally insulating rubber foam was placed on top to hold the leaf in contact with the metal. A thermocouple probe was inserted between the metal and the foam to accurately measure leaf temperature (T_{LEAF}). During this acclimation treatment 0.5 cm x 1 cm sections were removed with a scalpel at 30 minute intervals, as well as at 1 hour, 3 hours, 6 hours and 12 hours post-acclimation. These sections were immediately prepared and examined by confocal microscopy but, despite exhaustive attempts, no fluorescence signal was detected from any of the transfections (not shown).

3.3.6 Sub-Cellular Localisation of Constitutively Expressed smHSPs

N. benthamiana leaves were again infiltrated with Agrobacterium, this time carrying constitutive 35S promoter constructs: 35Spro:HSP17.6:GFP;

35Spro:HSP17.6a:YFP; and 35Spro:HSP17.b:RFP. Plants were left for 36 hours at Normal Conditions for transient transfection to take place. Sections 0.5 cm x 1 cm were removed with a scalpel, prepared, and immediately examined under a confocal microscope.

3.3.6.1 Localisation of HSP17.6

Of all the fluorescent HSP17.6 constructs, the version tagged with RFP exhibited the clearest fluorescent signal (35Spro:HSP176.6:RFP) (Figure 3.8). Morphologically, the fluorescence signal did not appear to coincide with any obvious sub-cellular compartments. A diffuse cytosolic signal can be seen in the lobes of these epidermal cells (**arrow A**); cytoplasmic strands across vacuolar areas are visible (**arrow B**); some areas are mixtures of diffuse cytosolic signal and sharper punctate patches (**arrow C**); a perinuclear pattern is obvious (**arrow D**); and uniform globular patches are evident near the periphery of some cells (**arrows E**). For a video reconstruction, see Video Files 1 and 2 (for legends, see page 105).

3.3.6.2 Localisation of HSP17.6a

None of the fluorescently tagged versions of HSP17.6a gave rise to any detectable signal under confocal microscopy (not shown).

3.3.6.3 Localisation of HSP17.6b

As with HSP17.6, the clearest signal from HSP17.6b was obtained from the RFP construct (35Spro:HSP17.6b:RFP - Figure 3.9). Similar patterns of sub-cellular localisation were also observed, though the cytosolic network pattern was much

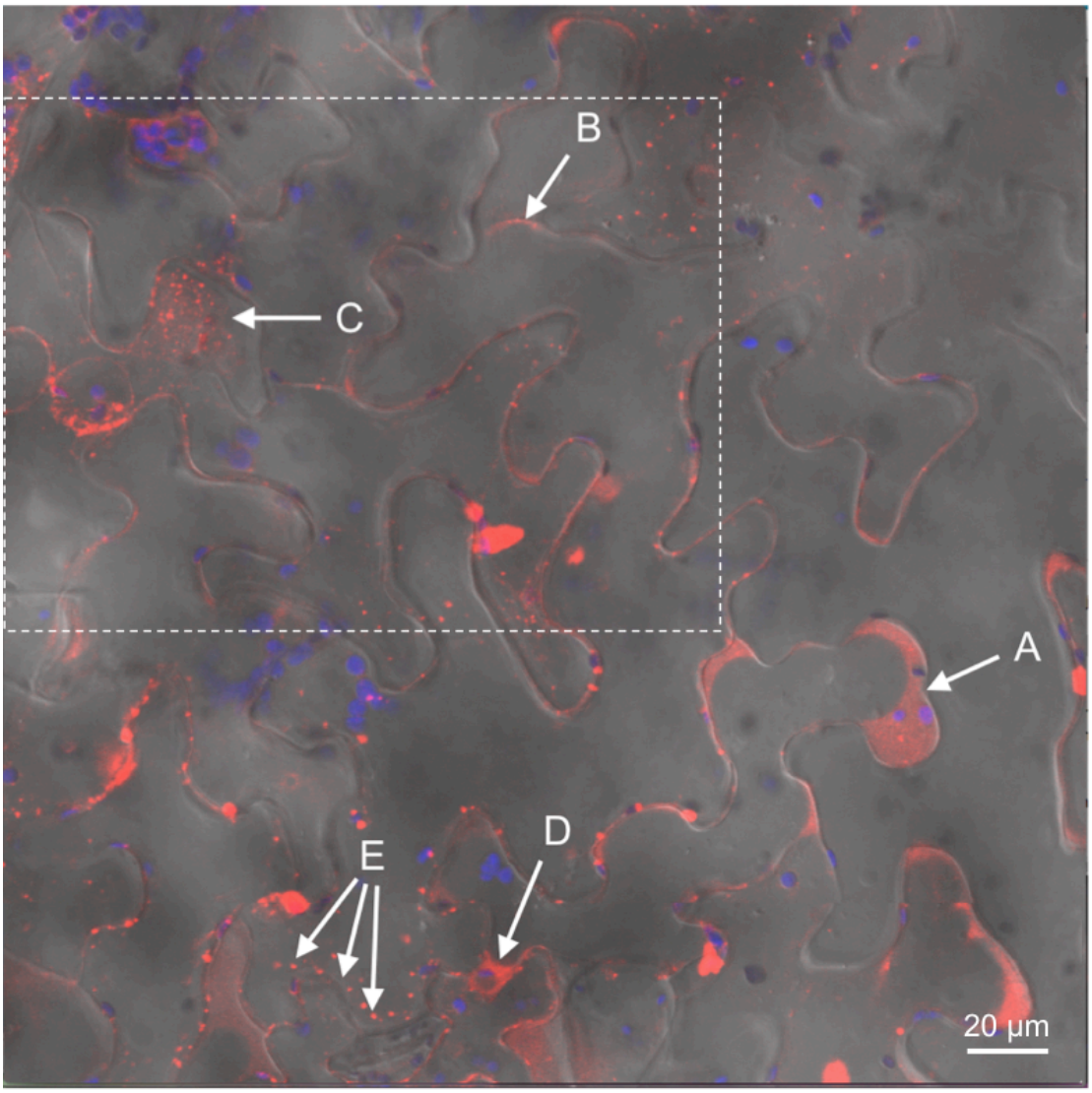


Figure 3.8 Sub-cellular Localisation of HSP17.6

Nicotiana benthamiana leaves were imaged by confocal microscope 36 hours post-transfection with 35Spro:HSP17.6:RFP. (**A**) diffuse cytosolic signal; (**B**) cytoplasmic strand; (**C**) complex mixture of smooth and punctate signal; (**D**) perinuclear localization; (**E**) punctate pattern at cell periphery. RFP fluorescence is in red; chloroplast autofluorescence is in blue. For 3D reconstructions from a Z-stack series of images, see Video File 1. Dashed line represents the area imaged in the timecourse in Video File 2.

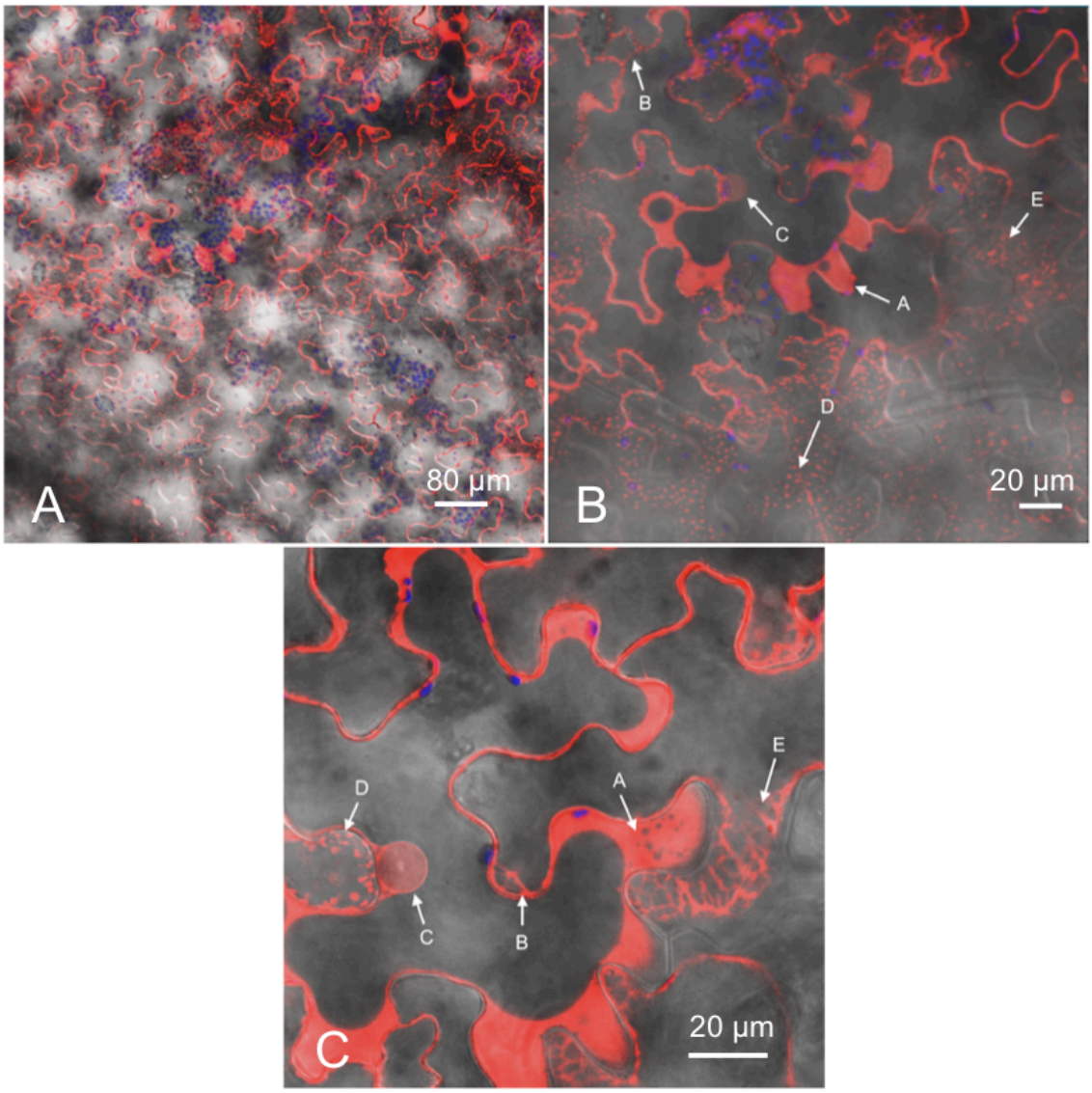


Figure 3.9 Sub-cellular Localisation of HSP17.6b

Nicotiana benthamiana leaves were imaged by confocal microscope 36 hours post-transfection with 35Spro:HSP17.6b:RFP. RFP fluorescence is in red; chloroplast autofluorescence is in blue.

A Overview of localisation patterns across a large area of lower leaf epidermis.

B Enlargement of panel A. Arrows A: diffuse cytosolic signal with negative signal representing membrane-bound organelles. Arrows B: cytoplasmic strand. Arrows C: perinuclear localization. Arrows D: punctate pattern exhibited in some cells and not in others. Arrows E: network pattern exhibited in some cells and not in others.

C Enlargement of panel B. Arrows as above.

more pronounced (**arrows B** and **E**). For a video reconstruction, see Video Files 3 and 4 (for legends, see page 105).

3.3.6.4 Co-localisation of HSP17.6 and HSP17.6b

Both HSP17.6 and HSP17.6b appeared to be localised in the cytoplasm. To confirm that they were in the same compartment as each other *N. benthamiana* leaves were co-infiltrated with two Agrobacterium cultures, one transformed with 35Spro:HSP17.6:GFP and the other transformed with 35Spro:HSP17.6b:RFP. 36 hours after infiltration a 0.5 x 1 cm tissue section was examined under a confocal microscope. While the transfection efficiency was reduced in this double-transfection, it was possible to find cells expressing both constructs (Figure 3.10). One such cell was imaged and analysed with the Intensity Correlation Analysis plugin for ImageJ (Schneider et al., 2012) to determine the extent of co-localisation. Two separate cross-sections were used (**lines 1** and **2**). The RFP and GFP signals were found to be strongly correlated, indicating that both of these chimeric proteins were located in the same compartment(s), which appear to be parts of the cytoplasm.

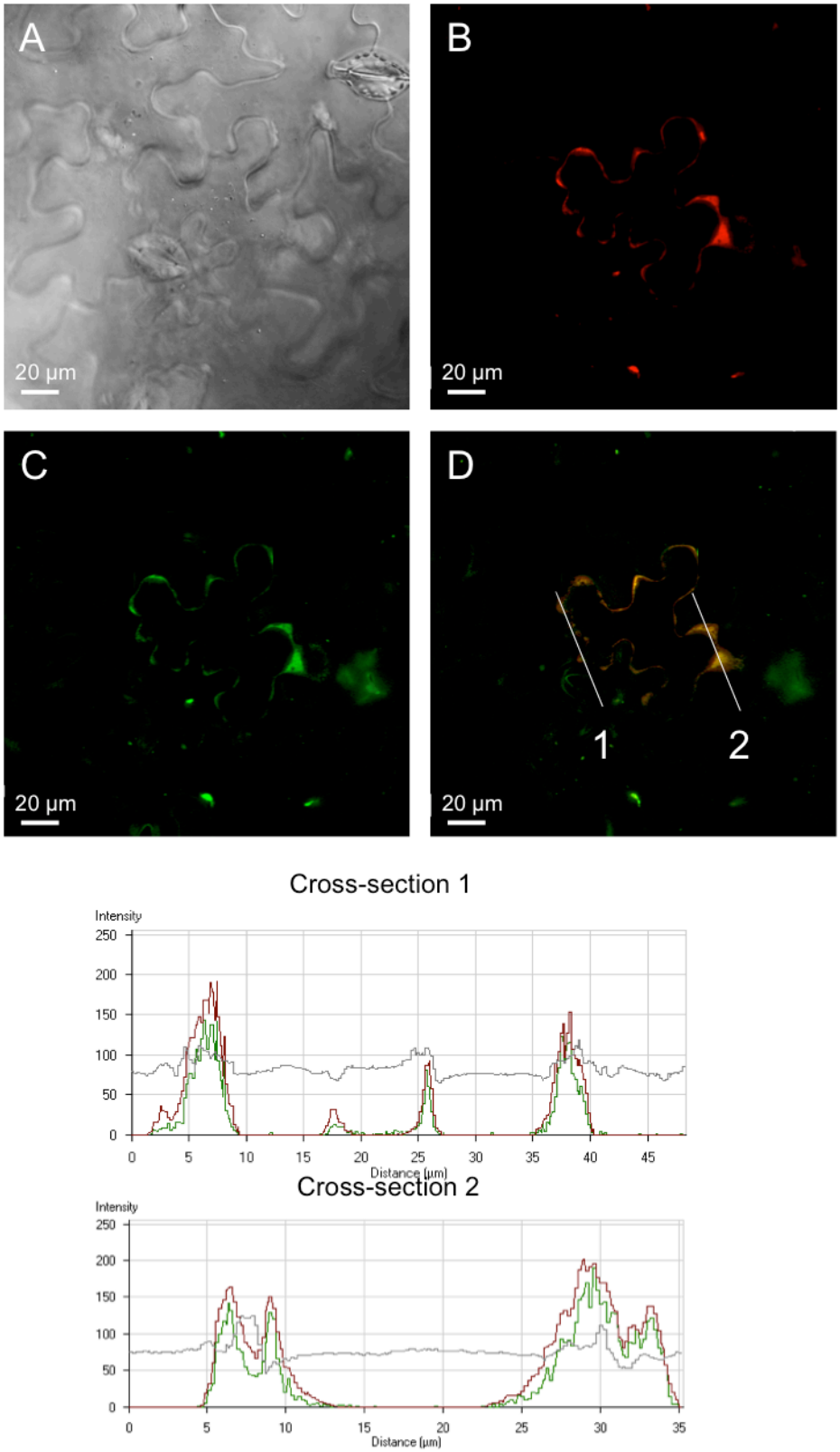


Figure 3.10 Co-localisation of HSP17.6:GFP and HSP17.6b:RFP

N. benthamiana leaves (bright field; **A**) were co-transfected with 35Spro:HSP17.6b:RFP (**B**) and 35Spro:HSP17.6:GFP (**C**) and imaged by confocal microscopy 36 hours later. Signal intensity from each colour channel was compared across two linear sections (**D**, lines 1 and 2) using the Image Correlation Analysis plugin in ImageJ (Schneider et al., 2012). The two fluorescent signals were strongly correlated. Green plots: HSP17.6:GFP intensity. Red plots: HSP17.6b:RFP intensity. Grey plots: bright field image intensity.

3.3.6.5 Figure Legends for Accompanying Video Files

Video File 1 Sub-cellular Localisation of HSP17.6

(*This video file complements* Figure 3.8) *N. benthamiana* leaves were imaged by confocal microscopy 36 hours post-transfection with 35Spro:HSP17.6:RFP. A Z-stack of images was captured through the lower epidermal layer of these leaves. **Clip A**: Images from the Z-stack are played in series from the surface of the tissue to the deepest point, then in reverse. 9 loops, 26 seconds total. **Clip B**: The Z-stack was used to produce a 3D reconstruction of the tissue which rotates around the Y axis. 8 loops, 50 seconds total. **Arrow A**: diffuse cytosolic signal; **Arrows B**: cytoplasmic strand; **Arrows C**: complex mixture of smooth and punctate RFP signal; **Arrows D** perinuclear localisation; **Arrows E**: punctate pattern at cell periphery. RFP fluorescence is in red; chloroplast autofluorescence is in blue.

Video File 2 Sub-cellular Dynamics of HSP17.6

(*This video file complements* Figure 3.8) *N. benthamiana* leaves were imaged by confocal microscopy 36 hours post-transfection with 35Spro:HSP17.6:RFP. A timeseries of images was captured over a period of 4 minutes 15 seconds with a capture framerate of approximately 1 image every 6 seconds. This video loops 3 times. **Arrows** as above.

Video File 3 Sub-cellular Localisation of HSP17.6b

(*This video file complements* Figure 3.9) *N. benthamiana* leaves were imaged by confocal microscope 36 hours post-transfection with 35Spro:HSP17.6b:RFP. A Z-stack of images was captured through the lower epidermal layer of these leaves. RFP fluorescence is in red; chloroplast autofluorescence is in blue.

Video File 4 Sub-cellular Dynamics of HSP17.6b

(*This video file complements* Figure 3.9) *N. benthamiana* leaves were imaged by confocal microscope 36 hours post-transfection with 35Spro:HSP17.6b:RFP. **Clip A**: A timeseries of images was captured over a period of 5 minutes with a capture framerate of approximately 1 image every minute. **Clip B**: A timeseries of images was captured over a period of 4 minutes with a capture framerate of approximately 6 images per minute. RFP fluorescence is in red; chloroplast autofluorescence is in blue.

3.3.6.6 Heat-induced Alteration in Localisation of smHSPs

After establishing that these smHSPs were cytosolic, it was decided to investigate whether their localisation altered upon activation at the thermal acclimation temperature (37 °C). The same procedures described above were used to transfect leaves with 35Spro:HSP17.6:RFP and 35Spro:HSP17.6b:RFP, and leaf sections were heated while on the microscope stage by constantly flushing with a stream of warmed water in a specially designed chamber. Temperature was monitored using a thermocouple held inside the same chamber. RFP was chosen as it provided the strongest fluorescence signal.

The temperature was steadily increased from room temperature (22 °C) to the target temperature (37 °C) over a period of ten minutes. The subcellular dynamics of the tagged proteins did not alter during the initial stages of this treatment, but the fluorescence was gradually quenched as the temperature exceeded 30 - 32 °C. This meant that it was not possible to obtain images at the temperature necessary to induce thermal acclimation (images not shown). As the temperature was gradually lowered to room temperature, the fluorescence signal returned.

This finding demonstrates the unsuitability of RFP for use in heat-stress investigations *in planta* in real-time.

3.4 smHSP Knockout Lines

As part of an effort to characterise the function of *HSP17.6* and its chromosomal neighbour *HSP17.6a*, knockout lines were obtained and investigated.

3.4.1 Selection and Acquisition of Knockout Lines

The Nottingham Arabidopsis Stock Centre (NASC, Scholl et al., 2000) carries seed stock generated from various large-scale mutagenesis events. Their database was interrogated for any lines annotated as carrying mutations within or around the coding regions of *HSP17.6* and *HSP17.6a* (Figure 3.11). An insertional mutant (stock #N507510) was available for *HSP17.6* carrying a T-DNA insertion of approximately 10 kb within the last 200 bp of the gene's single exon. For *HSP17.6a* there were no insertional mutants in the coding region, though there was a line available with an insertion within the 200 bp upstream of its TSS (stock #N672448), which was generated by the same large-scale insertional mutagenesis and thus carried a copy of same approximately 10 kb T-DNA. This was within the notional promoter region, so while it was not as likely to result in a complete loss of function, it was obtained on the basis that this upstream lesion might still attenuate expression of *HSP17.6a*.

Each of these putative knockout lines, notionally designated *hsp17.6* and *hsp17.6a*, were supplied labelled as homozygous seed stock. To confirm the genotypes, PCR primers were designed to the sequences just 5' and just 3' of each insertion site (Figure 3.11, red flags; see Table 2.12 for primer sequences). Seeds were sown on soil, leaves were sampled from 3 separate seedlings of each line, genomic DNA was isolated, and PCR was carried out using these primers

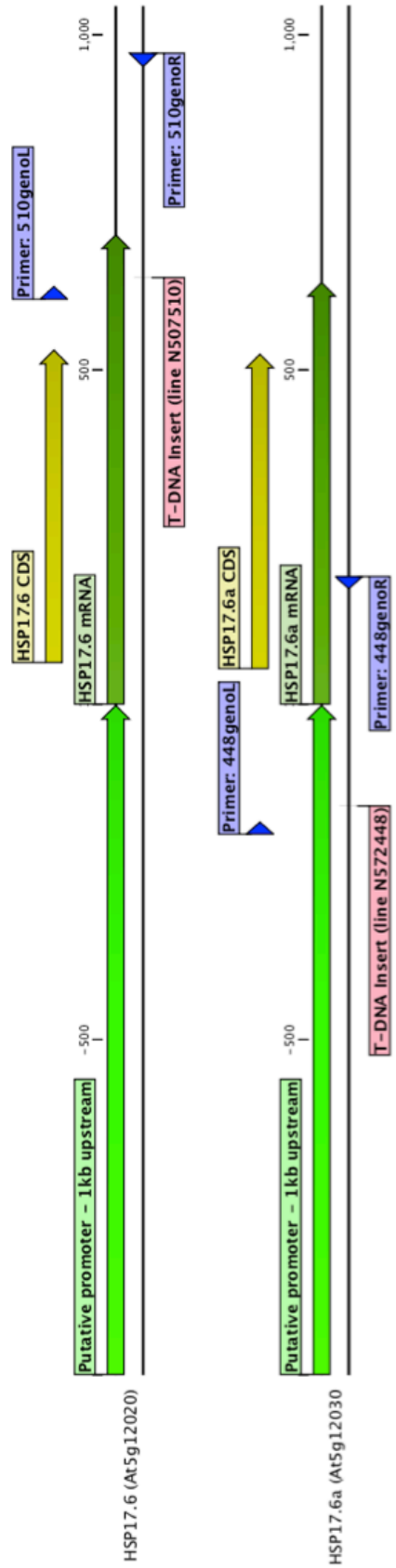


Figure 3.11 Locations of T-DNA Insertions and Associated Genotyping Primers

Two independent lines carrying T-DNA insertions were identified as potential loss-of-function mutants and seed was acquired from the Nottingham Arabidopsis Stock Centre (NASC, Scholl et al., 2000). Line N507510 was annotated in the catalogue as carrying T-DNA in the 3' UTR of *HSP17.6* and therefore the insertion might disrupt proper transcription and mRNA processing. Line N572448 was annotated as carrying T-DNA within the 200 bp 5' of the transcriptional start site of *HSP17.6a* and was therefore considered to be a candidate promoter-knock-down line.

Genotyping primers (blue flags) were designed to anneal to the genomic regions 5' and 3' of the insertion sites (pink flags). If a T-DNA insertion was present, its size (~10 kb) would prevent a PCR product from being made within the extension time any typical PCR cycle, thus confirming the presence of a potentially mutant allele. If a T-DNA insertion was *not* present, a product of ~400 bp would be created from the wild type allele.

to detect the presence of any wild type alleles (Figure 3.12). Since the expected result indicating a knockout would in this case take the form of a negative PCR result, positive control PCRs were performed comprising the reciprocal primer & sample pairs.

These primers to the wild type *HSP17.6* allele failed to produce a product from all three *hsp17.6* gDNA samples (left part of Figure 3.12) indicating that all three of those plants were homozygous for the insertion allele, while the same primers all successfully produced a wild type product from all three gDNA samples from the other mutant, *hsp17.6a*, thus proving that the primer pair would have worked had a wild type allele template been present. The same was also found for primers to the wild type *HSP17.6a* allele (right part of Figure 3.12) in the reciprocal combination. Thus, it was confirmed that each plant was indeed homozygous for its respective T-DNA insertion. These seedlings were grown to maturity, seed was harvested, and their progeny were used for knockout investigations.

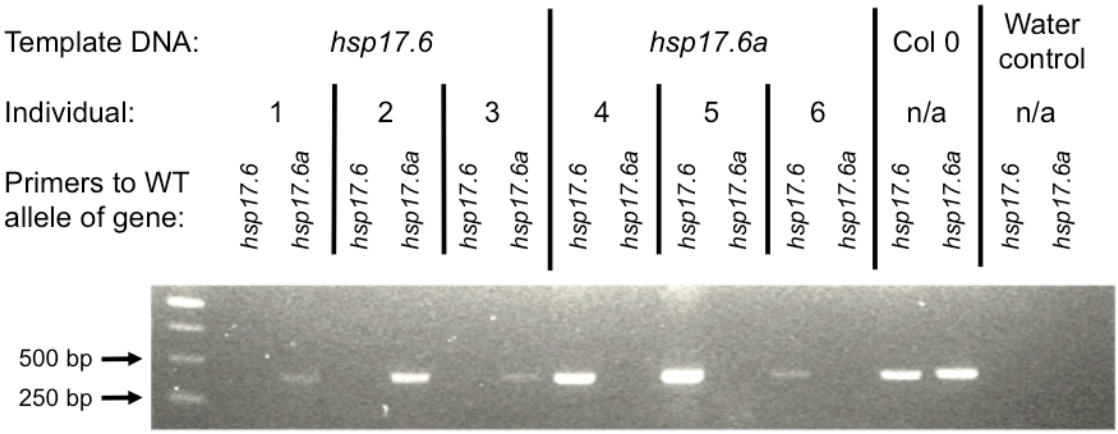


Figure 3.12 PCR Confirmation that Putative Knockout Lines *hsp17.6* and *hsp17.6a* are Homozygous

Genotyping primers (Table 2.12) were used on gDNA extracted from 3 individual seedlings of each putative knockout line. Primers were designed to detect the presence of a wild type allele; absence of a band therefore represents a line homozygous for the ~10 kb T-DNA insert between the primer binding sites. Reciprocal PCR reactions were performed as controls: presence of a band when *hsp17.6* genotyping primers were used on *hsp17.6a* gDNA, for example, provided evidence that the template was intact.

3.5 Thermotolerance of smHSP Knockout Lines

These knockout lines were then investigated for thermosensitivity phenotypes at various stages of development.

3.5.1 Thermotolerance of Wild Type Arabidopsis

Thermotolerance can be classified as one of two types: basal or acquired. Basal thermotolerance involves protection that does not require a lag time in order to be effective. The precise nature of these basal mechanisms is poorly defined in the literature, and is mentioned mainly in relation to innate differences in heat stress responses between species, ecotypes and genotypes (Barnabas et al., 2008, Clarke et al., 2004). Acquired thermotolerance, on the other hand, is the phenomenon by which an organism can survive a lethal temperature if it is allowed to reach that temperature gradually (Song et al., 2012). In plants, this type of response can involve signalling mediated by plant hormones such as ABA, with these signals leading to transcriptional cascades, and ultimately to the expression and/or post-translational activation of individual effector proteins (Nakashima et al., 2009, Pieterse et al., 2012), the most notable of which are members of the HSP families. The lag time necessary to achieve a state of acquired thermotolerance can be attributed to these stepwise processes.

Two parameters had to be defined before any thermotolerance experiments could be carried out on the smHSP knockout lines: the temperature above which non-acclimated wild type Arabidopsis is unable to survive (the 'killing temperature' - T_K), and the optimum duration of the lag phase required to allow

wild type plants to reach a state of acquired thermotolerance (the 'acclimation period').

3.5.1.1 Killing Temperature (T_K)

Wild type seedlings are widely reported in the literature to be able to survive at temperatures up to approximately 43 or 44 °C. A typical killing protocol used for screening involves an exposure of 3 to 4 hours (Nurcahyanti, 2009). A preliminary experiment was carried out to confirm that these temperatures used for these durations would be sufficient for seedlings sown using our protocols and heat-treated using our incubators.

Wild type seedlings were sown on agar plates (section 2.3.2) and germinated horizontally and allowed to grow for 2 weeks before being transferred to the incubator for 3 hours or 4 hours, and this was repeated at temperatures of either 42, 44 or 46 °C. Experimental measurements made with a thermocouple probe held at various positions around the interior of the incubator showed that the actual air temperature varied from the thermostatically set temperature by ± 0.5 °C. It was determined that 44 ± 0.5°C for 3 hours was sufficient to kill these 2-week old wild type seedlings (data not shown, and Nurcahyanti, 2009)).

3.5.1.2 Acclimation Period

There were two components of the acclimation period to be defined: the length of time required to initially *induce* the process of thermotolerance and the length of time required to *express or activate* the necessary components. It was reasoned that it would be best to first give a conservatively long induction period and determine the length of time required to activate the protective

mechanisms, than to experiment with reducing the induction period to find its lower limit. For both these trials 37 °C was selected as the acclimation temperature as this (± 1 °C) has consistently been reported in the literature to be sufficient (e.g. Vierling et al., 1986, Chen et al., 1990, Vierling, 1991, Larkindale and Vierling, 2008, Pavlova et al., 2009, Zupanska et al., 2013).

To investigate the optimal length of the activation period, plants were given a long induction at 37 ± 0.5 °C for 3 hours followed by an expression/activation period at 22 °C of either 0, 3, 6, or 12 hours, then a stress at the previously determined T_K (see section 3.5.1.1) of 44 ± 0.5 °C for 3 hours (Figure 3.13, **panel A**). Plants were qualitatively assessed for viability 36 hours after the stress, and again 5 days later. Representative images are presented in (Figure 3.13, **panel B**).

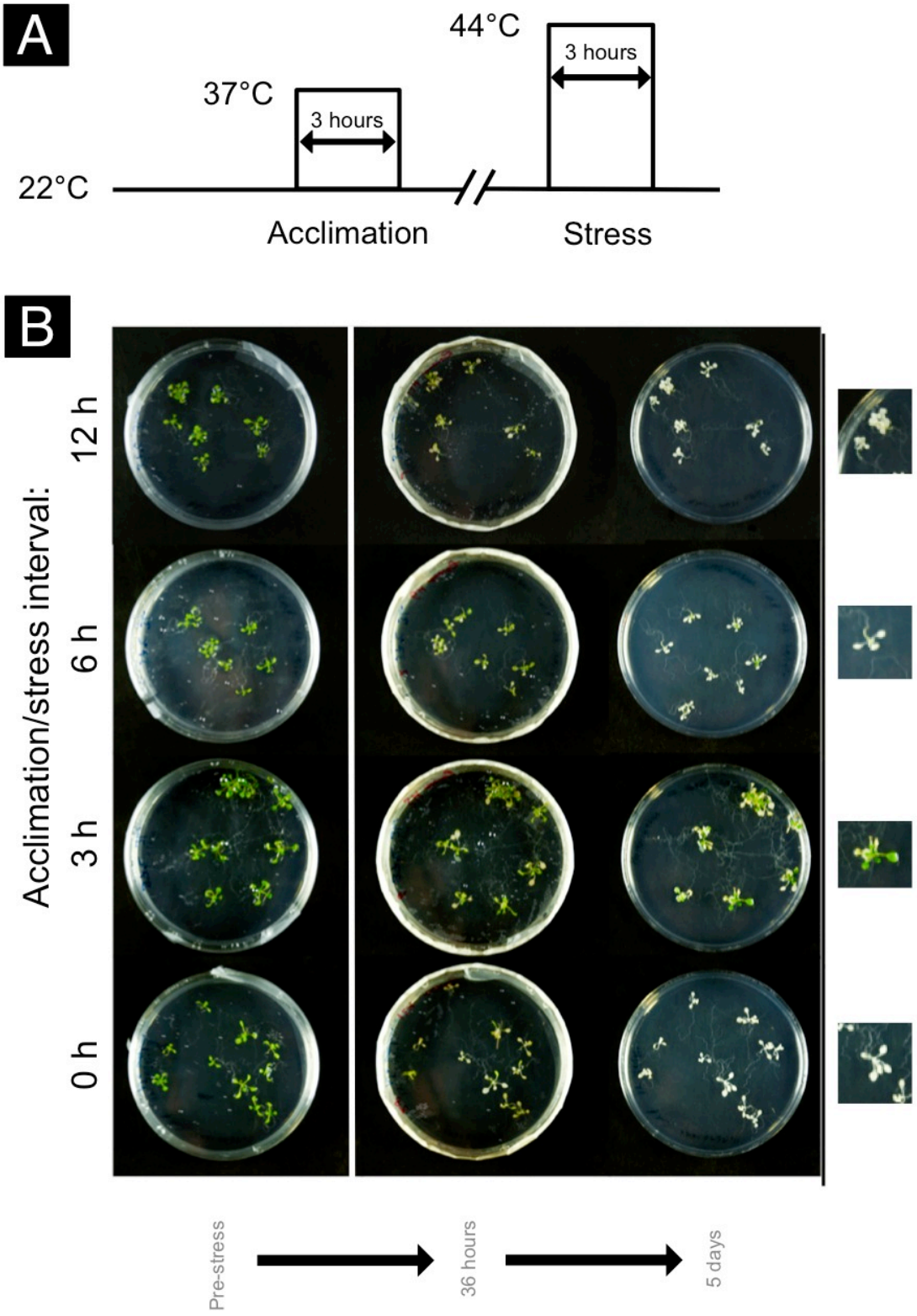


Figure 3.13 Empirical Determination of the Most Protective Time Interval Between Acclimation and High Temperature Stress

While it was not possibly to gradually increase the temperature of the incubation chamber from 37 °C to 44 °C in a smooth and reproducible manner, it was known that an acclimation period at 37 °C followed by a 'rest' period improved survival of seedlings subsequently exposed to the killing temperature, 44 °C (Larkindale and Vierling, 2008). **A** Plants were acclimated at 37 ± 0.5 °C for 3 hours, returned to Normal Conditions at 22 °C for 3, 6, or 12 hours, then exposed to the killing conditions of 44 ± 0.5 °C for 3 hours. **B** Seedlings were photographed before, 36 hours after, and 5 days after exposure to the killing temperature. Magnifications or representative plants are presented to the right.

It was observed that a direct transfer from acclimation temperature to the T_K did not induce acquired thermotolerance – all plants were completely white / brown by day 5. Plants given a 3 hour activation period exhibited the greatest survival – more than half of the leaves on each seedling remained green (particularly the younger leaves, perhaps indicating that those which are actively growing and are therefore more metabolically active are most capable of mounting the appropriate defence response). Among the seedlings given a 6 hour activation period a small number of leaves remained green but the vast majority of the aerial tissue appeared to have died by day 5, and plates given 12 hours responded similarly to those directly transferred to the T_K.

To investigate the minimum length of time required at the acclimation temperature in order to induce acquired thermotolerance, seedlings were induced at 37 °C for either 1, 2, or 3 hours, then given an expression period at 22 °C for 3 hours as recommended above, and then a stress at 44 ± 0.5 °C for 3 hours (Figure 3.14, **panel A**). Seedlings were qualitatively assessed for survival at 36 hours and 5 days post-treatment (Figure 3.14, **panel B**). In all seedlings a number of older leaves died but the majority remained green and there was no observable difference between survival after an induction of 1, 2, or 3 hours.

Based on these observations future thermotolerance experiments incorporated an acclimation protocol consisting of 1 hour at 37 ± 0.5 °C to induce acquired thermotolerance followed by 3 hours at 22 °C to allow activation of the appropriate defence mechanisms, then exposure to the T_K of (unless otherwise specified) 44 ± 0.5 °C.

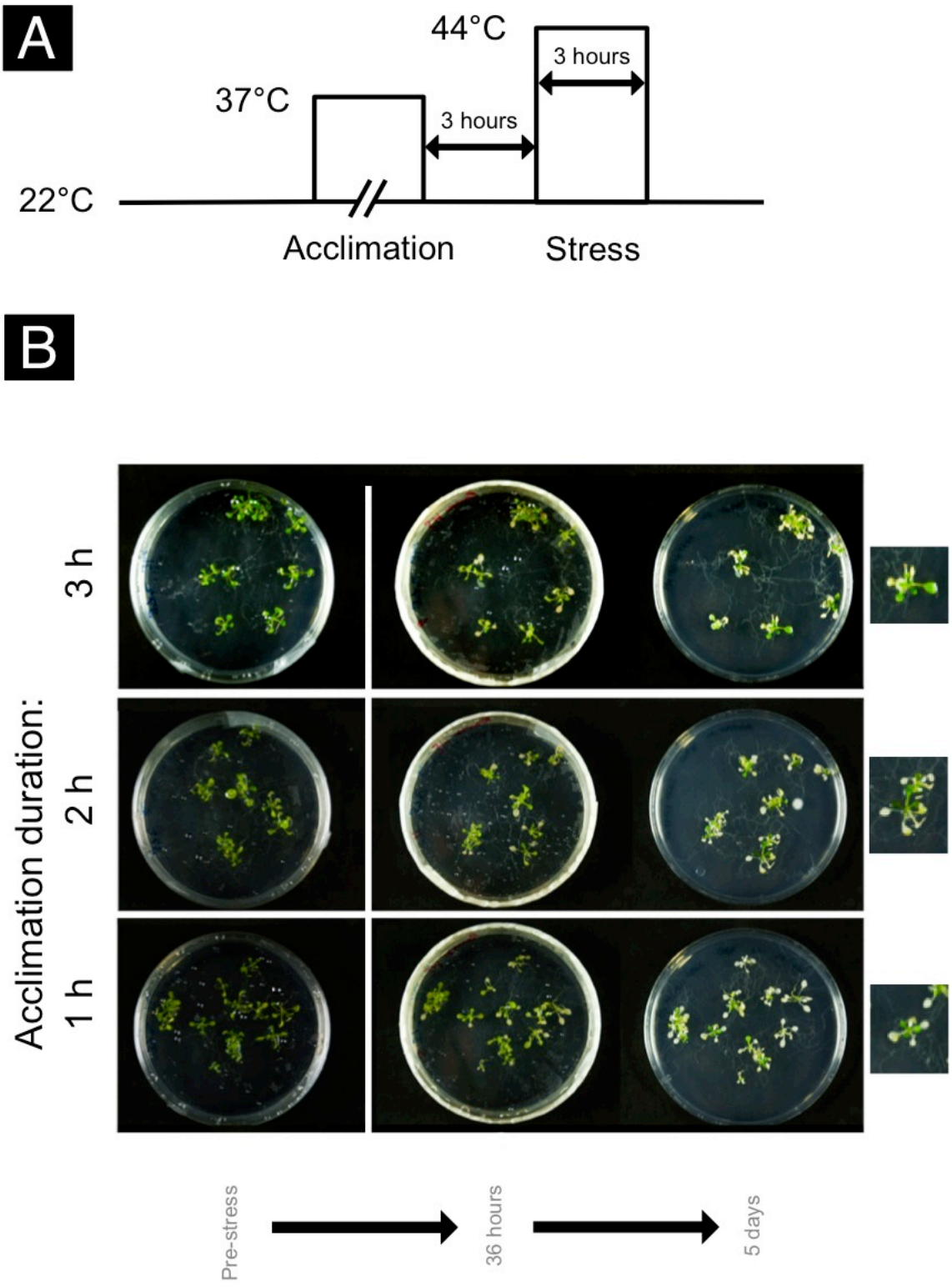


Figure 3.14 Empirical Determination of the Most Effective Thermal Acclimation Period

After demonstrating that an acclimation prior to exposure to the killing temperature provided protection to seedlings (Figure 3.13), the optimum duration of acclimation period was investigated. **A** Plants were be acclimated at 37 ± 0.5 °C for either 1, 2, or 3 hours, returned to Normal Conditions at 22 ± 0.5 °C for 3 hours, then exposed to the killing conditions of 44 ± 0.5 °C for 3 hours. **B** Seedlings were photographed before, 36 hours after, and 5 days after exposure to the killing temperature. Magnifications of representative plants are presented to the right.

3.5.2 Thermotolerance of smHSP Knockouts: Overall Morphology

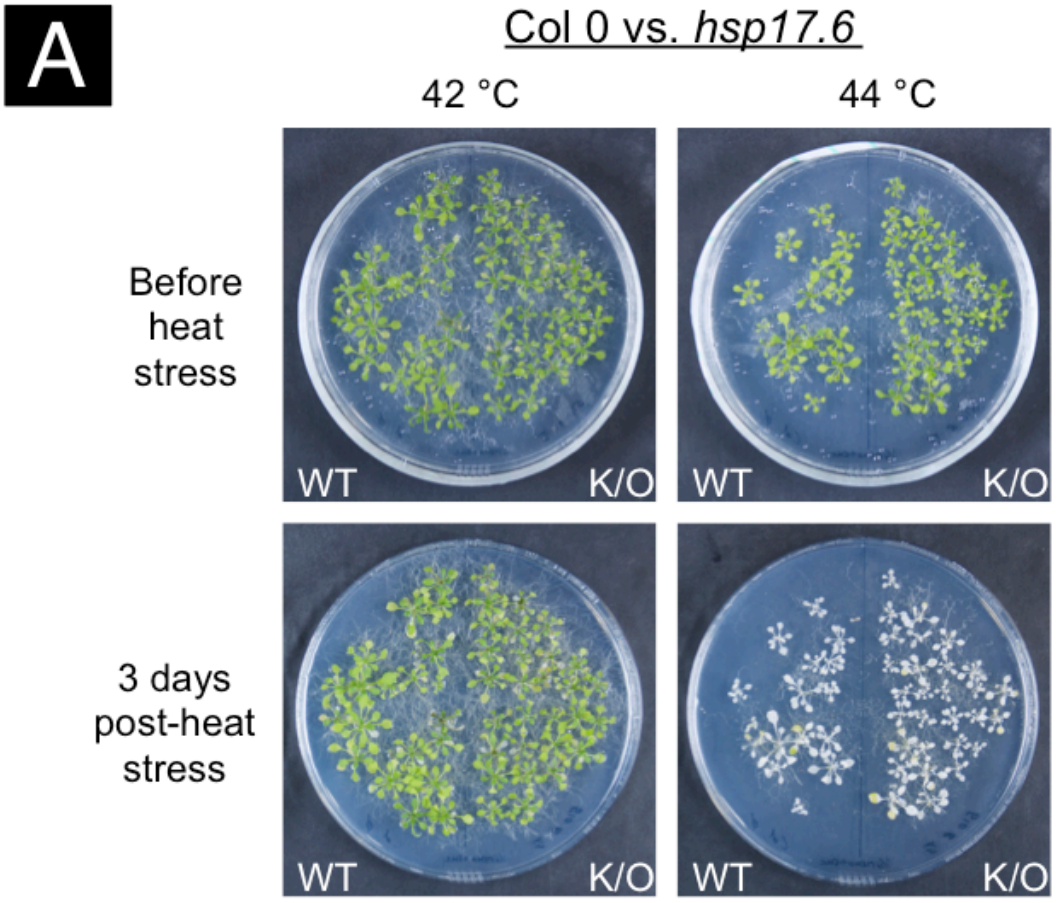
Seedlings were assessed at 2 weeks after germination to determine whether the knockout lines were more sensitive to high temperatures than wild type.

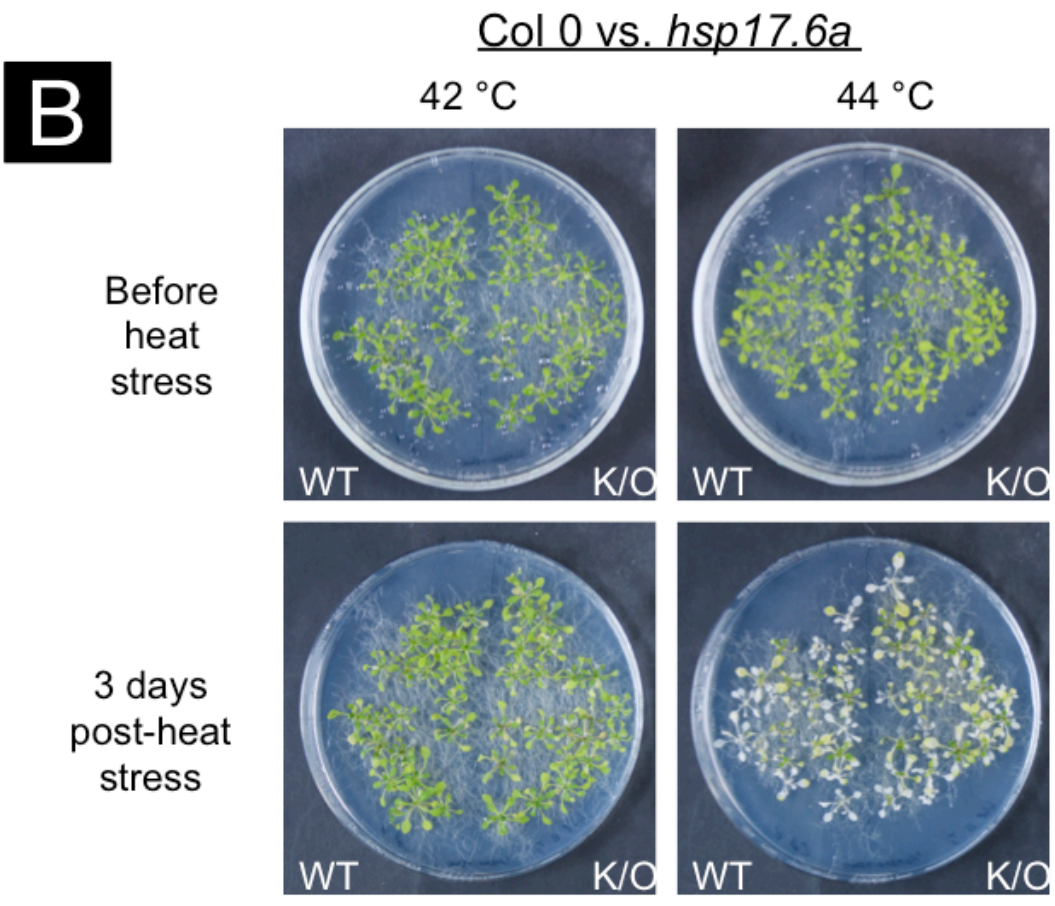
Approximately 30 seeds of wild type Col 0 were sown on bipartite agar plates (section 2.3.2) alongside 30 seeds of either line *hsp17.6* or *hsp17.6a* and grown under Normal Conditions. Two weeks after germination, seedlings were acclimated for 1 hour at 37 °C (see section 2.5) and then heat stressed for 3 hours at either 42 or 44 ± 0.5 °C. After a recovery period of 3 days at Normal Conditions plates were assessed visually for survival (Figure 3.15).

Plants exposed to the wild type killing temperature (44 ± 0.5 °C; right-hand-side of figure) all suffered a degree of bleaching and death. There was some between-plate variation which can be seen from the way that all plants on the plate containing line *hsp17.6* (**panel A**) turned completely white and all plants on the plate containing *hsp17.6a* (**panel B**) had some remaining green leaves, but within plates there were no clear differences between knockout and wild type lines. Plants exposed to a temperature slightly below the killing temperature (42 ± 0.5 °C; left-hand-side of figure) all retained their green pigmentation, indicating that the knockouts are not thermosensitive by this particular measure of survival.

Figure 3.15 Survival of Heat-stressed 2 Week Old Seedlings of *hsp17.6* and *hsp17.6a*

Seedlings of Col 0 and either line *hsp17.6* (**A**) or line *hsp17.6a* (**B**) were sown on split plates, grown until 2 weeks old, then given a moderate heat acclimation at 37 ± 0.5 °C for 1 hour to induce smHSP expression. Following a 3 hour rest under Normal Conditions, the seedlings were incubated either at the wild type killing temperature (44 ± 0.5 °C) or a sub-lethal temperature to test for hypersensitivity (42 ± 0.5 °C) for 3 hours, then returned to the growth room at 22 °C and visually assessed 3 days later. At the wild-type killing temperature, all plants died and there was no difference between knockouts and Col 0. At 42 ± 0.5 °C the knockout lines showed no sign of being compromised – there was no sign of bleaching and none of the leaves had become flattened onto the medium.


A
Col 0 vs. hsp17.6
42 °C
44 °C
Before heat stress
3 days post-heat stress
WT
K/O


B
Col 0 vs. hsp17.6a
42 °C
44 °C
Before heat stress
3 days post-heat stress
WT
K/O

3.5.3 Thermotolerance of smHSP Knockouts: Germination

As discussed previously, members of the smHSP family are known to be expressed during seed maturation / germination (section 1.2.1). If *HSP17.6* and *HSP17.6a* were important for thermotolerance at this stage then it is reasonable to hypothesise that applying a heat stress during the early stages of germination would reveal a sensitive phenotype. This hypothesis was tested.

Seeds of all three genotypes (Col 0, *hsp17.6* and *hsp17.*6a) were sown on tripartite agar plates (section 2.3.2) so that each line under each condition was exposed to an identical microclimate within the incubator. Twenty seeds of each line were used and plates were set up in duplicate. Immediately following stratification at 4 °C, plates were incubated at 22, 42, 44, or 46 ± 0.5 °C for 3 hours. Plates were then moved to the growth room at Normal Conditions to recover. Germination was scored 1 week later (Figure 3.16).

Wild type line Col 0 was able to survive the 42 ± 0.5 °C stress with a germination rate of 100 % which dropped slightly to 87 % when stressed at 44 ± 0.5 °C. Knockout line *hsp17.6* achieved 93 % germination at 42 and 44 ± 0.5 °C. Line *hsp17.6a* germinated less well than either of the other two lines under control conditions (70 %) and this was not further reduced at either 42 or 44 ± 0.5 °C. All lines completely failed to germinate after an exposure to 46 ± 0.5 °C.

Analysis of variance tests on these data indicated no significant differences ($p>0.05$; see Appendix ii-a) in the main effects of line and temperature, or their interaction. This indicated that knocking out either of these two smHSP genes did not cause a thermosensitive phenotype at germination.

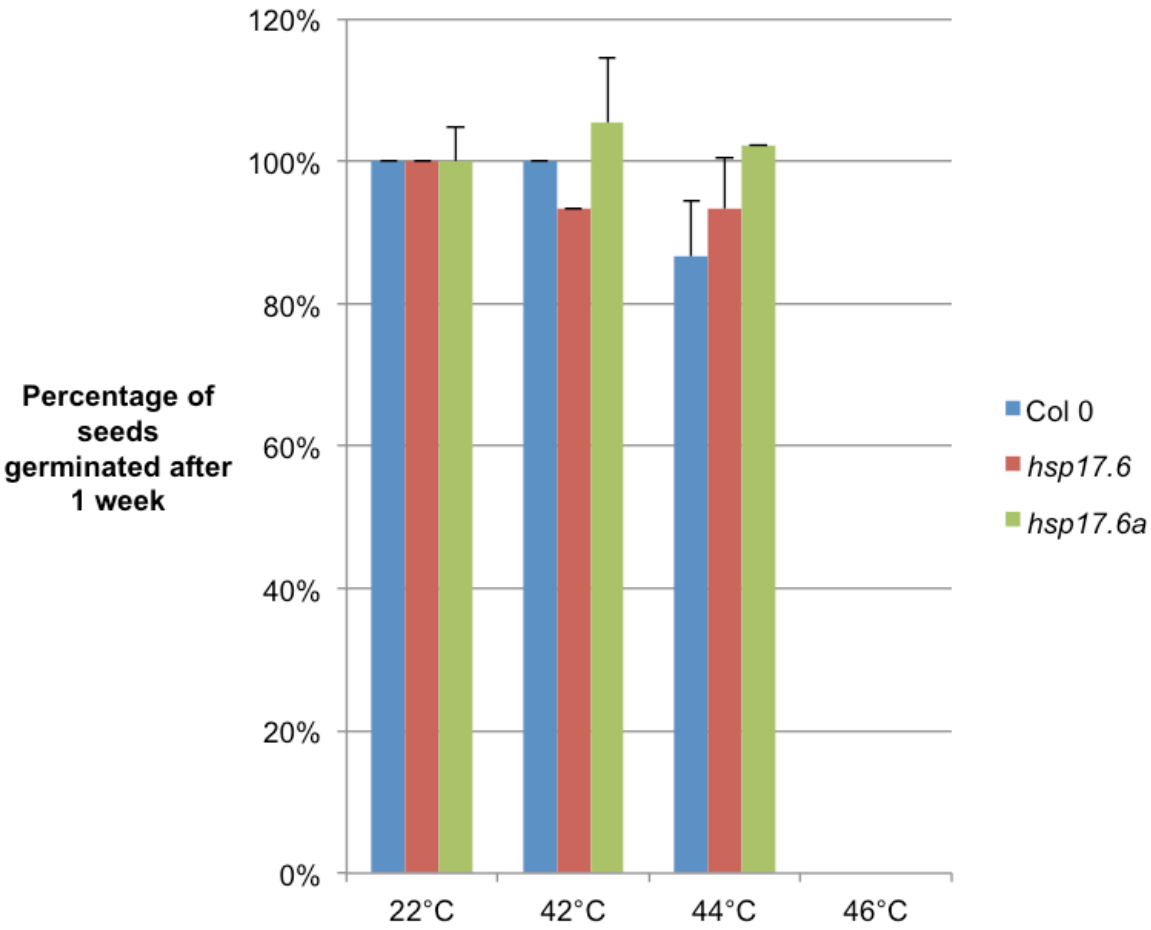


Figure 3.16 Germination of *hsp17.6*, *hsp17.6a*, and Wild Type Seeds Following Heat Stress

Seeds of lines Col 0, *hsp17.6*, and *hsp17.6a* were sown on split tripartite plates, in duplicate, and stratified at 4 °C in the dark for 3 nights. Seedlings were incubated immediately afterwards at 22, 42, 44 or 46 ± 0.5°C for 3 hours, then returned to the growth room at 22 °C. Germination was scored one week later. Each line's measurements are normalised against that line's germination rate at 22 °C. Error bars represent means ± SE; *n* = 2 replicates of 15 plants each.

3.5.4 Thermotolerance of smHSP Knockouts: Root Development

While smHSPs were shown to be expressed at a lower level in the roots than in shoots (Figure 3.4), they were still upregulated 200-fold in response to a thermal acclimation. To test the hypothesis that these smHSPs play a role in thermotolerance in roots, root lengths were measured before and after a heat stress.

8 seeds each of Col 0 and either of the two knockout lines were sown on bipartite agar plates (section 2.3.2), in duplicate, and germinated vertically. Plates were incubated at stress temperatures 5 days later (see section 2.5). The extent of root growth in the 3 days following the stress was recorded and plotted in Figure 3.17 (**panel A**: *hsp17.6*, **panel B**: *hsp17.6a*).

When transferred directly from 22 °C to either 42 or 44 ± 0.5 °C (non-acclimated; left side), all root growth in all lines was effectively stopped. When given an acclimation at 37 ± 0.5 °C (right side) root growth continued after a stress at 42 ± 0.5 °C and it did so slightly more in wild type (approx. 90 % of non-stressed control, **panels A** and **B**) than in *hsp17.6* (approx. 70 % of non-stressed control, **panel A**) or in *hsp17.6a* (approx. 50 % of non-stressed control, **panel B**) though neither of these reductions in knockout lines were statistically significant ($p>0.05$). Root growth was effectively stopped in all lines when acclimated and then transferred to the wild type killing temperature (44 ± 0.5 °C).

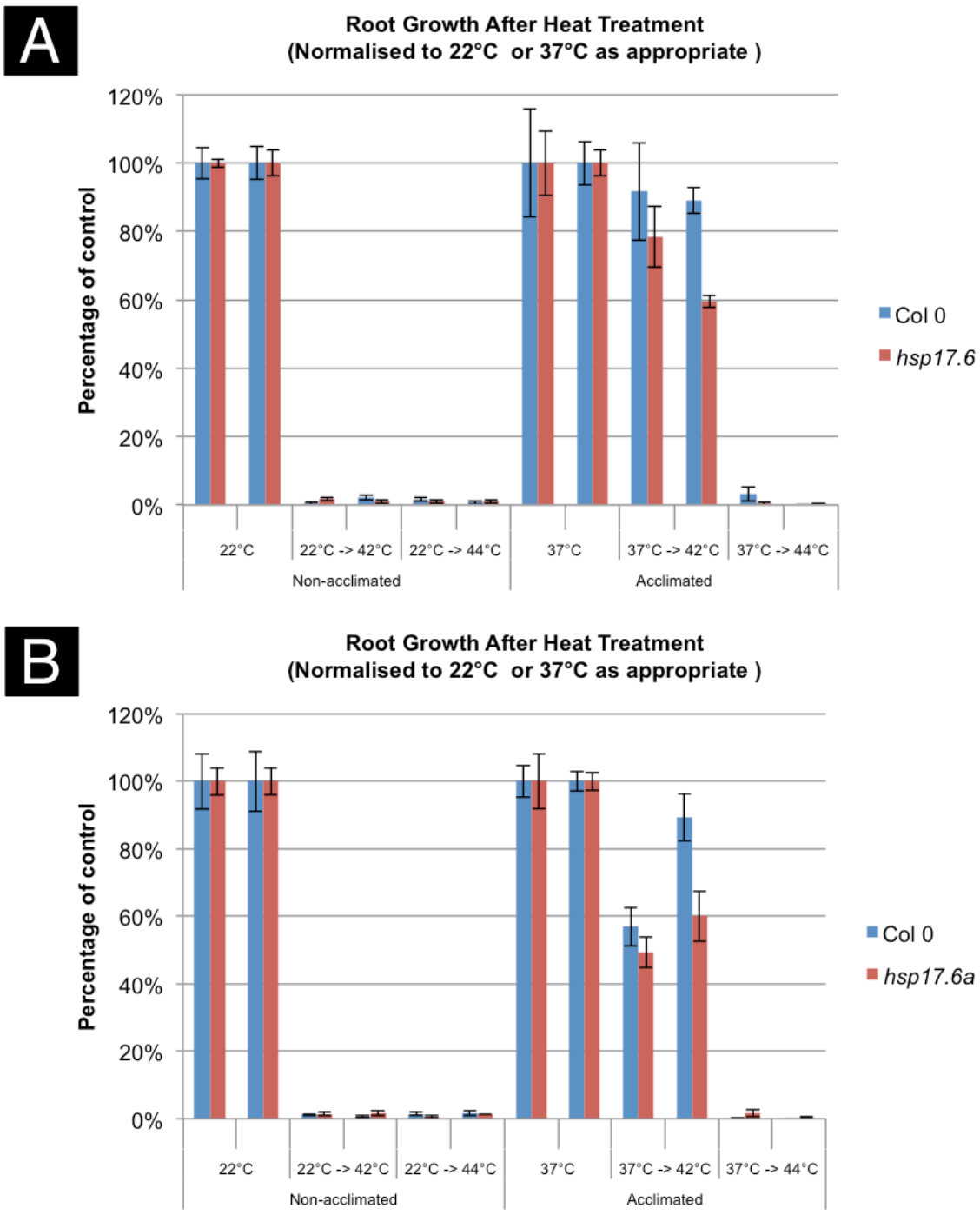


Figure 3.17 Effect of Heat Stress on Root Extension in Knockout Lines *hsp17.6* and *hsp17.6a* and in Wild Type

Seeds of wild type (Col 0), *hsp17.6* and *hsp17.6a* were sown on split bipartite agar plates and germinated vertically. Heat stress was applied 5 days later. Root extension was measured a further 3 days later. Data from two individual replicate plates are presented for each treatment as there may be some variation attributable to plates / position of plates within the incubator. **A** Post-stress root growth exhibited by Col 0 and *hsp17.6*. **B** Post-stress root growth exhibited by Col 0 and *hsp17.6a*. Non-acclimated samples are normalised against 22 °C; acclimated samples are normalised against 37 ± 0.5°C. Error bars represent means ±SE; n = 15.

3.5.5 Thermotolerance of smHSP Knockouts: Hypocotyl Extension

Seeking to discern whether thermotolerance was conferred upon the aerial tissues shortly after germination, an equivalent experiment was performed to measure the length of dark-grown hypocotyls before and after a heat stress. Germinating seeds in the dark causes an etiolated phenotype characterised by exaggerated height and thin hypocotyls. Evolutionary adaptations drive the seedlings to gain height until they break through the soil layer into the light where photosynthesis can proceed, rather than expending energy on developing laterally-oriented leaves in the dark. This exaggerated extension in only one direction provides a convenient method for measuring growth rate.

Fifteen seeds each of wild type Col 0 and of the two knockout lines were sown on tripartite agar plates (section 2.3.2). Plates were wrapped in aluminium foil immediately after sowing seeds so that the plants were entirely dark-grown and placed in the growth room. After 3 days the plates were briefly unwrapped in low light so that the height of the hypocotyls could be marked on the walls of the plate, then were re-wrapped and exposed to a heat regime as follows: 22 ˚C (non-acclimated control); 22 ˚C -> 44 or 46 ± 0.5 ˚C (non-acclimated heat-stressed); 37 ± 0.5 ˚C (acclimation control); and 37 ± 0.5 ˚C -> 44 or 46 ± 0.5 ˚C (acclimated heat-stressed). Hypocotyls were measured before and 2 days after heat stress, and the difference represented post-stress growth.

Figure 3.18 shows hypocotyl extension measurements expressed as a percentage of the appropriate control (either 22 °C for non-acclimated samples or 37 ± 0.5 °C for acclimated samples). Lines *hsp17.6* and *hsp17.6a* were significantly more

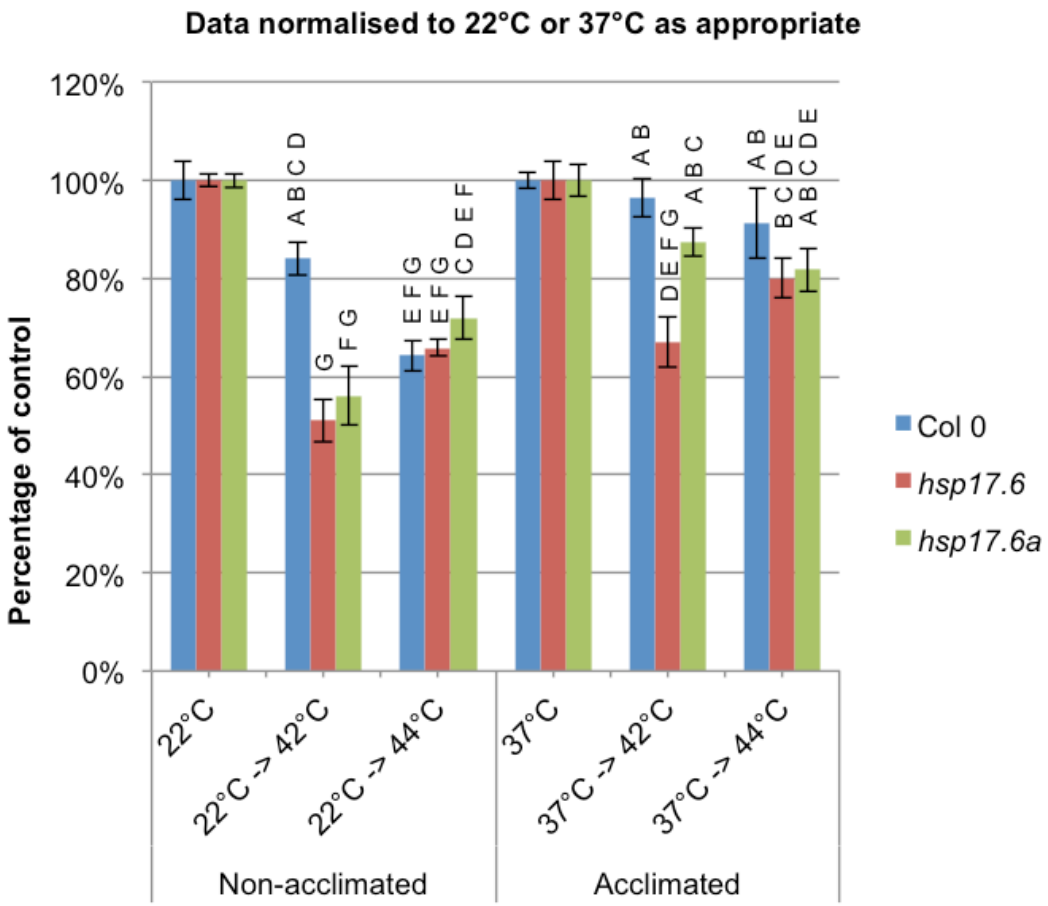


Figure 3.18 Effect of Heat Stress on Hypocotyl Extension in Knockout Lines *hsp17.6* and *hsp17.6a* and in Wild Type

Seeds of wild type (Col 0), *hsp17.6* and *hsp17.6a* were sown on split tripartite agar plates, wrapped in tinfoil to maintain a dark environment and germinated vertically. Heat stress was applied 5 days later. Hypocotyl extension was measured a further 3 days later. Non-acclimated samples are normalised against 22 °C; acclimated samples are normalised against 37 ± 0.5°C. Error bars represent means ±SE; *n* = 15. Means that do not share a letter are significantly different (see Appendix ii-b).

susceptible than wild type ($p<0.05$; see Appendix ii-b) to a non-acclimated stress at 42 ± 0.5 °C (they achieved hypocotyl extensions of 51 % and 56 % of those exhibited by non-stressed controls while wild type achieved 84 %), but all three lines were inhibited roughly equivalently when the non-acclimated stress temperature was 44 ± 0.5 °C (wild type: 64 %, *hsp17.6*: 66 %, *hsp17.6a*: 72 % of control extension). With an acclimation before stress, line *hsp17.6a* performed significantly better than without ($p<0.05$; see Appendix ii-b), and again the two knockout lines exhibited less hypocotyl extension than wild type when stressed at the lower temperature of 42 ± 0.5 °C though this was only significant for line *hsp17.6* ($p<0.01$; see Appendix ii-b) (wild type: 97 %, *hsp17.6*: 87 %, *hsp17.6a*: 67 %) but there was no significant difference between lines after stress at 44 ± 0.5 °C (wild type: 91 %, *hsp17.6*: 80 %, *hsp17.6a*: 82 %).

This suggests that there is a very mild thermosensitivity phenotype among these knockouts shortly after the germination stage if plants are exposed to a stress temperature slightly below that of the normal wild type killing temperature (in this case 42 °C rather than 44 °C).

3.7 Halotolerance of smHSP Knockout Lines

There are two reasons for investigating the halotolerance of knockout lines for these two genes. The first is that the genes were originally identified as being upregulated in an experiment assessing survival at a high concentration of salt. The second is based on an extrapolation of what was already known and hypothesised about their function. If they operate as molecular chaperones to prevent denaturation, then perhaps they have a role in other stress conditions that would also lead to a change in protein conformation. Increased salt

concentration causes a change in the ionic strength of a solution, thus altering the balance of the electrostatic forces between a folded polypeptide and its environment that contribute to determining which parts of the protein are folded internally and which are exposed on the surface. Perturbations in salt concentration therefore have the potential to denature proteins and it would be of evolutionary benefit to an organism to upregulate the expression of chaperones such as the smHSPs. Conversely, it is hypothesised that loss of these proteins might confer a halosensitive phenotype.

Eight seeds of each knockout line were sown alongside wild type on bipartite agar plates, in triplicate, and supplemented with a final concentration of 0, 40, or 60 mM NaCl. Seedlings were grown vertically for 2 weeks. Root length was taken as the measure of growth since this was the most obvious difference between lines (Figure 3.19).

A salt concentration of 40 mM was sufficient to cause a reduction to 36 % of control root length in wild type, while roots of *hsp17.6* were only reduced to 82 % and those of *hsp17.6a* to 53 %. Root extension of both these knockouts was significantly different from each other, and from that of wild type (for a full statistical analysis see Appendix ii-c). At 60 mM NaCl the wild type was reduced to 32 % of control and hsp17.6a was reduced to 38 % - these values were not significantly different from each other – while hsp17.6 was only reduced to 59 % of control. Therefore it was concluded that the two knockout lines are less sensitive to 40 mM salt than the wild type line, though only the hsp17.6 knockout was more tolerant at 60 mM.

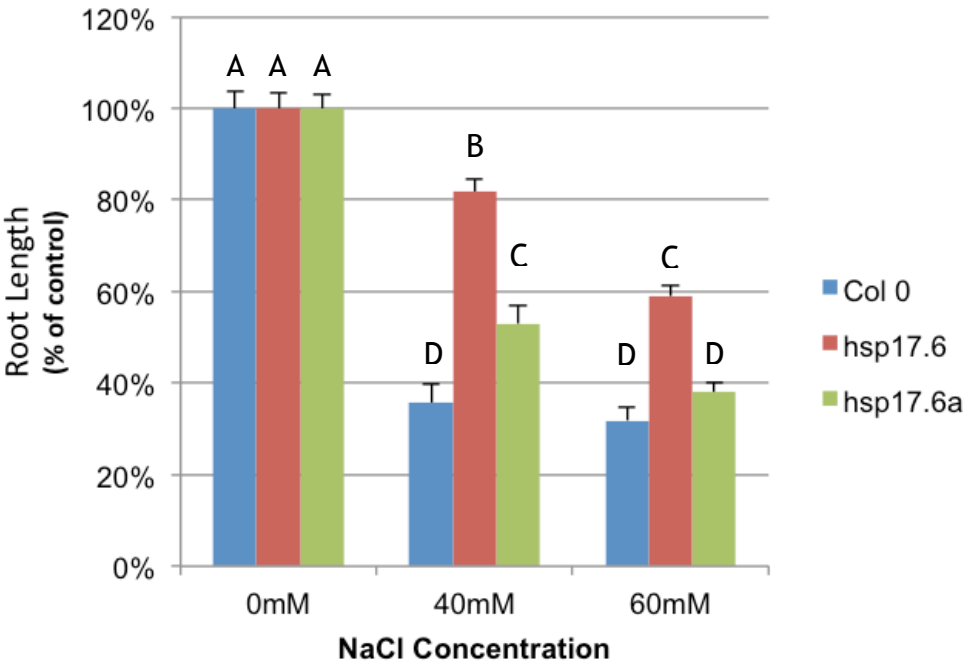


Figure 3.19 Halotolerance of *hsp17.6* and *hsp17.6a* Knockout Lines

Col 0 and smHSP knockouts were germinated together on split plates containing basal medium supplemented with 0, 40, or 60 mM NaCl and grown vertically. Roots were measured 14 days later. Error bars represent averages ± SE; *n* = 3 replicates of 15 plants each. Means that share a letter are not significantly different (see Appendix ii-c).

3.8 Discussion

The smHSPs discussed here have high similarity in their primary sequence and therefore in their hydrophobicity profile and predicted secondary structures, which strongly implies that they share a common function. If this is the case, it raises the question of why there are so many similar members in the Arabidopsis superfamily, and more generally in plants when compared with the other kingdoms. The simplest hypothesis would be that these chaperones are each responsible for the protection of a different subset of clients, whether those subsets be defined by groups of related proteins, proteins localised in different subcellular compartments, or simply different structural features or binding sites at different points on the surface of any given client. The initial plan was to attempt to answer this question regarding specificity by performing protein isolations on extracts sampled at appropriate points during heat stress. It is for this reason that the epitope-tagged smHSP constructs were created. In the intervening time, however, developments in the literature began to suggest that smHSP thermal activation and client-interaction relied more on quaternary protein dynamics (whereby smHSP multimers dissociate into dimers, revealing surfaces that are able to interact promiscuously with polypeptides at risk of aggregation) than on sequence specificity (Baldwin et al., 2011a, Benesch et al., 2010, Stengel et al., 2010), so this line of investigation was not actively pursued.

Recent publications have reiterated that the N-terminal region has proven difficult to resolve crystallographically (for more detail see Section 1.5.1 - smHSPs in Plants and Chapter 6 (“General Discussion”). The implication would be that these N-terminal arms are free to move in solution and this would provide a degree of flexibility in client-binding. When multimerised, some of

these arms appear to be located centrally to the oligomer and others externally, providing for a variety of three-dimensional orientations in which protein-protein interactions might take place. The hydrophobic tract in the N-terminal arms of HSP17.6 and HSP17.6a would add diversity to the arsenal of potential chaperone:client interactions possible in one of the activated heteromultimers.

The attempts to identify promoter binding sequences in the regions upstream of the smHSPs described here did not lead to any novel discoveries. The failure to express from own-promoter constructs in *N. benthamiana* leaves could be either because the regulatory sequences are more distal than 1kb upstream, or because the *N. benthamiana* transcription machinery is sufficiently different that it does not bind to those sequences.

The knockout lines produce some evidence of mild thermosensitivity by the measures employed here. At 2 weeks old, and at germination, the delivery of a heat stress did not hamper growth of the knockouts significantly more than wild type (Figure 3.15 and Figure 3.16). Root growth measurements include between-plate discrepancies which are yet to be satisfactorily resolved but are probably due to localised heating differences: on some plates the wild type and the knockout were suppressed equally; on other plates the knockout performed significantly worse (Figure 3.17). Dark-grown hypocotyl extension measurements incorporating an acclimation demonstrated a reduction in knockout growth only in line *hsp17.6* relative to wild type, and only when the stress temperature was 42 ± 0.5 °C but not at the standard T_K 44 ± 0.5 °C (Figure 3.18). Without an acclimation, however, hypocotyl extension of both knockouts at this lower stress temperature was significantly less than wild type.

Measurements of seedlings grown on high NaCl, however, revealed a striking difference in the response of *hsp17.6* from that of wild type and the same is true of *hsp17.6a*, though to a lesser extent.

The increase in halotolerance of the knockouts was unexpected and perhaps suggests that without so much of these proteins there is a lack of input into a negative feedback loop that acts to moderate the expression of the smHSP family, such as has been observed by Rhoads with HSP26 (2005) and by Montero-Barrientos with HSP70 (2010). This would result in the overcompensation observed here (Figure 3.19) and warrants further investigation.

These smHSPs, upregulated in line HT5 which overexpressed MYB64, represented an interesting line of research, and exhaustive investigations showed that they may act redundantly and so more work must be done to tease apart the contributions that each member of the large smHSP family provides. The weak phenotypes described by the smHSP knockout lines and the dimer-based polydispersity paradigm developing in the literature were all motivating factors to shift focus back to MYB64 and the genes it might be responsible for activating under stress.

3.8.1 Technical Considerations

3.8.1.1 Historical Microarray Identification of smHSPs in Line HT5

The HT5 transcriptome profile was revealed using microarray technology (specifically, the GeneChip® Arabidopsis ATH1 Genome Array from Affymetrix (High Wycombe, UK)) and this raises a concern. These smHSPs are highly conserved. *HSP17.6* and *HSP17.6a* are 86% identical at a nucleotide level and the

other members of the smHSP family all share the conserved α-crystallin domain, so it is possible that there was cross-hybridisation of the cDNA to the probes. This would have implications for the expression levels reported in Table 1.1. If further work were to be carried out on these smHSPs, it would be desirable to confirm by qPCR that they are indeed upregulated under salt stress conditions in line HT5 and perhaps investigate the levels of the other cytosolic members of the family. Unfortunately this was not possible as the phenotype of the HT5 line had been lost by the start of the investigations documented in this thesis.

3.8.1.2 Use of Segregated Wild Type Lines

Several early experiments were carried out using Col 0, the wild type supplied by the Nottingham Arabidopsis Stock Centre (NASC), as a control. This was done for historical reasons and because Col 0 was the background in which the mutants had been made. It became apparent, however, that Col 0 from different stocks did not all behave as one. Pools of seed which could be traced back to the same original stock were no more reliable than pools obtained from other researchers who received their Col 0 stock at different times. This could have been due to a number of reasons (epigenetic effects, novel mutations, cross-pollination, historical mislabelling of stocks) so it was reasoned that the most appropriate wild type control to use for any experiment on a mutant or transgenic line was one which had segregated from the original mutagenesis event. These lines were confirmed as such first by screening on the appropriate selective agent to confirm sensitivity, then by PCR to confirm the absence of the relevant insertion (data not shown).

These early findings strongly support the recommendation that the wild type used in any experiment should be as closely related as possible, by segregation

and in terms of the number of intervening generations, to the experimental line(s) in question.

4 35Spro:MYB64 Transcript Profiling

4.1 Introduction

As explained in detail in section 1.4, the microarray results from previous research by Price (2005) were obtained from an activation-tagged line carrying an enhancer element inserted 1370 bp upstream of the transcription start site of *MYB64*. A moderate, long-term (3-4 week) salt stress was used in this screen to drive expression of salt-responsive genes, so that the inserted tag simply enhanced endogenous salt-responsive *MYB64* transcription levels. After a 35Spro:MYB64 construct was made and stably transformed Arabidopsis lines were generated, it was of interest to find out whether the transcript profile of this new, constitutively overexpressing line was comparable with that of the activation-tagged one, particularly in the absence of the moderate long-term salt stress. By this time new techniques were available which provided more sensitivity and a higher resolution of expression differences than microarrays. Next generation sequencing (NGS), for example, involves the parallel sequencing of thousands of template DNA molecules from a biological sample and the real-time integration of that data into a list of sequenced reads. This method was chosen for analysis of the 35S:MYB64 transcriptome and the results gave fresh insights that led to investigation of other potential downstream effectors of *MYB64*.

4.2 Next-Generation Sequencing (NGS)

Modern nucleic acid sequencing has developed from the original process of DNA replication with chain termination, known as ‘Sanger sequencing’ (Sanger et al.,

1977). It is helpful to first understand the basis of this technique. Briefly, a template DNA molecule is replicated many times using a pool of free nucleotides spiked with one of four analogs carrying a hydrogen at the 3' position instead of an OH group, leading to a loss of the ability to further extend the chain. This chain-terminating analog is radiolabelled to facilitate detection of DNA strands into which they are incorporated. By saturating the reaction so that the chain is replicated enough times that it will be terminated at each possible position, and by comparing the lengths of each molecule from all four reactions at single-base resolution in tandem, it is possible to establish the sequence of a complete stretch of DNA. This technique was superseded by the use of fluorescently-labelled terminators such that all four analogs could be incorporated in one single reaction, and the identity of the base at each position could be simply determined by the colour at each position in the sequence (Smith et al., 1986).

Illumina sequencing technology (Illumina United Kingdom, Essex, UK) uses a variation on this process known as 'sequencing by synthesis'. Briefly, the chain-terminating analogs are labelled with a fluorescent moiety which is removable, and sequencing is performed on each template molecule progressively from one end to the other.

This technique was chosen to compare the transcript profiles of 35Spro:MYB64 line 141 and the segregated wild type, Col 0. Sixty-four siblings of each line were grown to maturity on soil in the greenhouse for approximately four weeks. Trays were randomly relocated in the greenhouse every few days to minimise the impact of environmental fluctuations on plant development. Leaves were harvested from each of the 64 individuals and grouped into pools of 8, then total RNA was extracted, yielding a total of 8 pools of RNA per line. These samples

were then pooled again, yielding one sample per line containing RNA representing all 64 siblings. This reduced any potential noise in the data that would have arisen from variation in growth or RNA extraction, and meant that a reasonably sized population could be sequenced economically. RNA was then passed to the technical staff at the Sir Henry Wellcome Functional Genomics Facility (College of Medical, Veterinary & Life Sciences, University of Glasgow) for preparation for sequencing on the Illumina platform.

Illumina sequencing takes place inside a flow cell to which the DNA (in this case cDNA) molecules are physically attached (Illumina, 2010). In the preparation stage of the process, total RNA was extracted from a tissue sample and messenger RNA (mRNA) was isolated by hybridisation to oligodT magnetic beads. This mRNA was washed, eluted and randomly fragmented into short sequences by incubation with divalent cations at high temperature, and then reverse transcription was carried out to create cDNA. Molecules were size-selected (in this case to ~35 bp) and short adapters ligated to each end (Figure 4.1, **panel A**). The products were then denatured into single strands and attached to the surface of a flow cell channel pre-coated with a dense lawn of primers complementary to the adapter sequences (**panel B)**. Free ends of the cDNAs were allowed to anneal to unbound primers and bridge-amplification followed (**panel C**), creating double-stranded copies of the original sequences (**panel D**) which were once again denatured into surface-bound single strands (**panel E**). This process was repeated multiple times to exponentially increase the number of copies into a 'spot' at the location of each template molecule's original position (**panel F**). The samples were then ready for sequencing by synthesis.

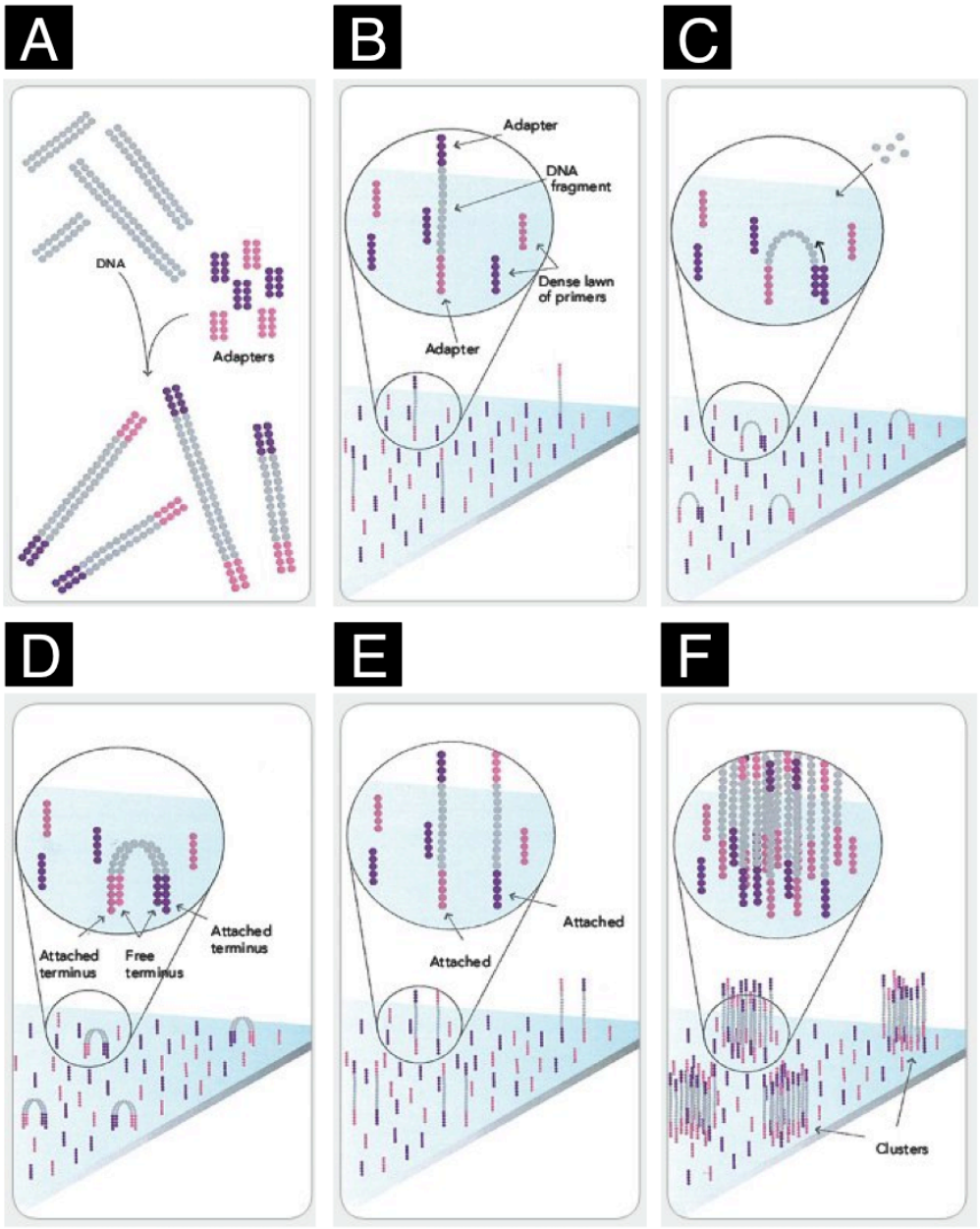


Figure 4.1 Schematic Diagram of the Illumina Preparation Process

A DNA is fragmented and adapters are ligated to each end **B** These modified sequences are denatured into single strands and bound to the surface of a flow cell, pre-coated with primers complementary to the adapters **C** Free adapter ends anneal to surface-bound primers and bridge amplification begins **D** A new surface-bound copy of the original molecule is produced **E** Products are denatured to single strands **F** The process is repeated until dense clusters, representing the original sequences, are created across the surface of the flow cell. This preparation increases the copy number of template molecules, and sequencing-by-synthesis using a reversible dye-labelled chain-termination technique follows.

(Figure adapted from Illumina, 2010)

The first base in each ~35 bp sequence was identified by adding a primer complementary to the adapter and each of the four fluorescently labelled deoxyribonucleotide triphosphates (dNTPs) along with the appropriate enzymes and buffers, and then scanning the flow cell with a laser to build a record of the colour at each spot. The process is known as *reversible* chain termination as the fluorescent labels were then cleaved from the nascent DNA strands, allowing another labelled dNTP to bind at the next position during the next round of synthesis. This process of single-base extension, laser scanning, and label cleavage was repeated to the end of the template and the order of the colours at each spot was used to reconstruct the order of the bases at each position. The output is a data file containing a list of the sequences ('reads') of each original cDNA fragment. Typically up to 5 million reads per channel and with an 8-channel flow cell up to 40 million molecules can be sequenced (> 1 gigabase).

The output file of sequenced reads was obtained from the Functional Genomics Facility and analysed using the RNA-Seq Analysis tool within CLC Genomics Workbench v3 using default parameters. This tool aligned each read to the annotated Arabidopsis reference genome (downloadable from The Arabdiopsis Information Resource (TAIR); available at www.arabidopsis.org (Swarbreck et al., 2008)) and the number of reads that aligned with each individual gene model was counted to build a global expression profile. The alignment output included a visual representation of each gene model together with each short cDNA read aligned to it (Figure 4.2). This method of expression analysis had an advantage over microarray data in that it allowed inferences to be made about splice variants. Where individual exons were matched with fewer reads it is likely that an exon had been spliced out in a proportional fraction of transcripts, and conversely where reads were matched to *introns* it is possible that in some

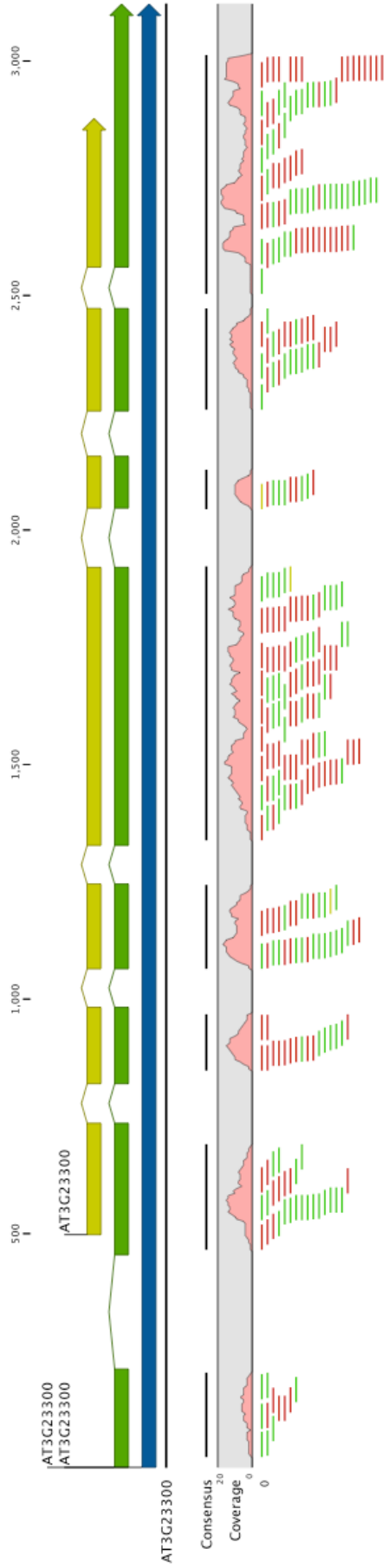


Figure 4.2 Example of Next-Generation Sequencing Output

Alignments were made between the mRNA sequence data from each of the experimental samples and the Arabidopsis reference genome (available online at www.arabidopsis.org) using CLC Genomics Workbench v3 using default parameters. In this example of an alignment output, the genomic DNA is represented by a narrow, solid, black, horizontal line near the top with block arrows above representing known gene features. Aligned sequencing reads (~35bp each) are represented as green, red, and yellow bars at the bottom. Green bars indicate fragments sequenced in the forward orientation, red bars in the reverse orientation. Yellow bars represent ambiguous reads (i.e. those that match >1 locus in the genome) which had been assigned proportionally to one of the matching loci. Sequencing coverage of the genomic region is indicated by the pink graph across the middle. Levels of coverage can be used to make inference about splicing; exons with only 50% of the coverage of the rest of the transcript are likely to be spliced out in 50% of mRNA molecules; sequencing coverage across introns may indicate retention in alternative splice variants, or alternatively that some pre-mRNA was captured in the extraction process before mRNA processing was completed.

Gene shown: At3g23300 (a methyltransferase of unknown function; for illustrative purposes only).

cases that intron had been retained. The technique did not rely on hybridisation, so false positives that can arise in microarrays from the annealing of cDNAs to close (but not perfect) matches were eliminated.

Since the exact sequence of each read was known, it also had the advantage of being completely unambiguous. Some reads matched more than one genomic locus, however. The software could be configured to handle these in one of two ways; either to take the ambiguous reads and assign them to one of the possible matching genes in a manner proportional to the number of flanking unambiguous reads at each of those genes, or to disregard them. In this instance the former setup was used.

The total number of reads aligned with each gene needed to be corrected for two factors before a between-samples comparison could be made: the length of each gene (since longer genes would of course be expected to be matched with a greater number of reads) and the total number of reads per sample (since variation in input concentration would lead to an artificially high level of expression reported for every gene in that sample). Thus, transcript abundance for each gene was adjusted and expressed as the number of Reads Per Kilobase, per Million mapped reads (RPKM). The RPKM values of each gene in each sample were then used to calculate a ratio of expression in the 35Spro:MYB64 line vs. Col 0. This ratio was used as the final measure of increased or decreased abundance of each transcript.

4.3 Transcript Profiling of 35Spro:MYB64 Line 141

The aim of this profiling experiment was to identify those transcripts that were most differentially abundant in the non-stressed line overexpressing *MYB64*. This

includes sequences that were more, as well as those that were less, abundant than in the wild type. For the sake of brevity, and since the changes ultimately stem from altered levels of a transcription factor, these sets of genes will herein be referred to as 'upregulated' or 'downregulated' (though it is not to be assumed that *MYB64* is modulating the transcription of all of these genes directly, nor that changes in abundance can be purely assigned to changes in transcriptional activity). For a list of the top 50 most highly 'upregulated' transcripts (plus one particularly interesting gene from the top 100) see Table 4.1 and a complete list, with a lower cutoff of 8-fold 'upregulation', is presented in Table A1.1 (in Appendix i). No attempt was made to filter the lists on the basis of low detection in the wild type sample. While this would remove those genes detected in wild type by only one or two reads and thus improve the validity of the ratio calculations, this would artificially exclude those genes which were nevertheless hugely 'upregulated' in the transgenic line. It was more important at this stage to determine which functional groupings were enriched than to precisely measure the extent of that enrichment. Another set of 'upregulated' transcripts were completely undetectable in the wild type sample at the depth of sequencing used (5 million reads) making it impossible to calculate a ratio at all. These were classified by the analysis software as 'infinitely upregulated' and include *MYB64*. A list of these can be found in Supplementary File 1a (see List of Accompanying Material, page xiii).

The list of the top 50 'upregulated' transcripts in Table 4.1 represents a multitude of abiotic stress responses, particularly those that one might expect to function relatively high up in a stress-response signalling cascade (e.g. MAP kinases and transcription factors). The 100 most highly 'upregulated' genes were grouped by common function, which is traditionally done with analogous

Table 4.1 Names and Functions of the Top 50 'Upregulated' Transcripts in 35Spro:MYB64 Transgenic Line (Plus One Extra Gene of Interest)

Transcript levels (as measured by Illumina Sequencing) were expressed as reads per kilobase per million mapped reads (RPKM) for each gene, then an RPKM ratio of transgenic : WT was calculated in order to show the most highly 'upregulated' transcripts. From left, this table shows: gene numbers, symbols, and functions as annotated in the database at TAIR; absolute expression differences (expressed in RPKM) and ratios between transgenic:wild type lines; and RPKM values along with the raw data used in their calculation. The top 50 genes in the list are presented here, plus one extra gene from within the top 100 (*DREB2A*) notable for its prevalence in the literature regarding abiotic stress responses including water balance and heat. Entries are colour coded based on functional annotations at TAIR. A subset of genes (gold) were selected for further study (see section 4.4 "Identification of Genes of Interest").

[1] RPKM is calculated using the number of unique exon reads. The number of unique gene reads is presented for information only.

						35Spro:MYB64				Col 0			
Rank	AGI Number	Gene Symbol	Function	Expression Difference (RPKM)	Ratio	Expression values (RPKM)[1]	Exon length	Unique gene reads[1]	Unique exon reads	Expression values (RPKM)[1]	Exon length	Unique gene reads[1]	Unique exon reads
1	AT4G23680	F9D16.150	Biotic stress response; Major latex protein-related.	172.26	466.15	172.63	751	845	845	0.37	751	1	1
2	AT1G01560	**MPK11**	ABA responsive; unknown role.	22.96	142.16	23.12	1710	333	257	0.16	1710	31	1
3	AT3G01830	F28J7.16	Calmodulin-related protein; unknown role.	10.01	46.84	10.23	1273	83	83	0.22	1273	1	1
4	AT4G25000	AMY1	Alpha-amylase-like; starch mobilisation.	16.95	44.63	17.34	1432	162	162	0.39	1432	2	2
5	AT1G58420	F9K23.5	Unknown.	12.67	34.71	13.05	740	63	63	0.38	740	3	1
6	AT1G10540	NAT8	Transmembrane transporter; unknown role.	4.57	33.06	4.71	1951	61	60	0.14	1951	1	1
7	AT5G66650	MSN2.3	Unknown.	25.35	30.31	26.22	1286	220	220	0.87	1286	4	4
8	AT3G53980	F5K20.280	Lipid transport; unknown role.	36.12	26.63	37.53	592	144	144	1.41	592	3	3
=9	AT3G57240	BG3	Glycosyl hyrdolase	11.26	25.90	11.71	1230	94	94	0.45	1230	2	2
=9	AT4G39670	T19P19.60	Glycolipid binding; unknown role.	8.35	25.90	8.69	829	47	47	0.34	829	1	1
11	AT2G22880	T20K9.9	Unknown.	35.25	22.87	36.87	345	83	83	1.61	345	2	2
12	AT1G80540	T21F11.13	Unknown.	3.49	22.04	3.65	1678	40	40	0.17	1678	1	1
13	AT1G28370	**ERF11**	DNA binding (repression?); ethylene and ABA response.	55.16	20.28	58.02	972	629	368	2.86	972	176	10
14	AT1G61340	T1F9.17	Unknown.	27.06	19.84	28.50	968	210	179	1.44	968	15	5
15	AT5G48657	(None)	Defence-related protein.	3.17	18.73	3.35	1557	34	34	0.18	1557	1	1
16	AT1G05575	(None)	Unknown.	19.54	18.46	20.66	497	98	67	1.12	497	15	2
17	AT3G10930	F9F8.25	Unknown.	47.37	18.10	50.14	703	230	230	2.77	703	7	7
18	AT3G61190	BAP1	Abiotic stress response.	25.72	18.07	27.23	923	164	164	1.51	923	5	5
=19	AT2G04495	(None)	Unknown.	4.92	17.63	5.21	941	30	30	0.30	941	1	1
=19	AT5G05300	K18I23.10	Unknown.	14.97	17.63	15.87	309	32	32	0.90	309	1	1
21	AT5G04340	**CZF2**	Cold-induced zinc finger protein. TF.	112.02	17.38	118.86	976	757	757	6.84	976	24	24
22	AT1G60190	T13D8.8	E3-ligase, unknown role.	17.95	17.33	19.05	2276	289	283	1.10	2276	12	9
23	AT1G12940	NRT2.5	Nitrate transporter; unknown role.	2.54	17.08	2.70	1762	36	31	0.16	1762	3	1
24	AT5G06510	NF-YA10	Nuclear factor; development?	14.48	16.99	15.39	1842	184	184	0.91	1842	6	6
25	AT4G25490	**CBF1**	TF; activates COR gene expression.	42.22	16.96	44.87	946	273	273	2.65	946	8	8
26	AT3G48640	T8P19.150	Unknown.	6.40	16.53	6.81	675	30	30	0.41	675	1	1
27	AT5G54490	PBP1	Pinoid binding protein; auxin response.	48.79	16.44	51.95	528	179	179	3.16	528	6	6
28	AT5G59220	**PP2C**	Protein phosphatase; drought & ABA response.	2.55	15.98	2.72	1632	29	29	0.17	1632	1	1
29	AT3G29000	K5K13.13	EF-hand family protein; unknown role.	32.51	15.87	34.70	636	144	144	2.19	636	5	5
30	AT3G02840	F13E17.22	Fungal elicitor protein; biotic stress response.	22.10	15.51	23.62	1278	197	197	1.52	1278	7	7
31	AT5G25240	F21J6.107	Unknown.	6.77	15.43	7.24	593	28	28	0.47	593	1	1
32	AT1G18570	MYB51	Response to salt, ABA, pathogens, ethylene.	28.83	15.29	30.85	1654	1253	329	2.02	1654	687	12
33	AT5G47230	ERF5	TF; response to cold, ethylene & chitin.	59.23	15.09	63.43	1191	493	493	4.20	1191	18	18
=34	AT1G56240	PP2-B13	Carbohydrate binding; unknown role.	3.38	14.88	3.62	1142	5	5	0.24	1142	1	1
=34	AT3G21780	UDPGT71B6	ABA glycosyl-transferase	5.18	14.88	5.55	1491	41	41	0.37	1491	2	2
36	AT4G27652	(None)	Unknown.	41.85	14.14	45.04	524	154	154	3.18	524	6	6
37	AT4G17490	ERF-6-6	TF; ethylene & chitin response.	149.71	14.07	161.17	1044	1098	1098	11.45	1044	43	43
38	AT2G22500	DIC1	Mitochondrial membrane transport; unknown role.	115.08	13.40	124.36	1528	1240	1240	9.28	1528	51	51
=39	AT3G06890	F17A9.4	Unknown.	4.65	13.22	5.03	731	24	24	0.38	731	1	1
=39	AT4G01360	F2N1.26	Unknown.	4.91	13.22	5.31	1386	48	48	0.40	1386	2	2
=39	AT4G28085	(None)	Unknown.	6.28	13.22	6.80	541	24	24	0.51	541	1	1
=39	AT5G45630	MRA19.3	Unknown.	6.60	13.22	7.14	515	24	24	0.54	515	1	1
43	AT2G32030	F22D22.22	Acetyl transferase, GCN5-related; unknown role.	7.66	12.95	8.30	868	47	47	0.64	868	2	2
44	AT2G41730	T11A7.17	Unknown.	6.83	12.67	7.42	475	23	23	0.59	475	1	1
45	AT1G76650	CML38	Calmodulin-like; response to wounding.	92.60	12.62	100.56	768	504	504	7.97	768	22	22
46	AT5G51190	ERF MWD22.13	TF; ethylene & chitin response.	96.51	12.56	104.86	833	570	570	8.35	833	25	25
47	AT1G17380	JAZ5	Jasmonate response.	8.46	12.49	9.20	1133	70	67	0.74	1133	3	3
48	AT4G34410	RRTF1	Redox-responsive TF; ethylene & chitin response.	25.01	12.32	27.22	1385	225	225	2.21	1385	10	10
=49	AT1G30190	(None)	Unknown.	3.70	12.12	4.03	837	180	22	0.33	837	81	1
=49	AT3G09020	T16O11.2	Glycosyltransferase; unknown role.	7.06	12.12	7.69	1315	66	66	0.63	1315	3	3
82	AT5G05410	DREB2A	TF; response to drought, heat and H2O2.	5.18	10.10	5.75	1465	55	55	0.57	1465	3	3

Abiotic stress responses

Biotic stress responses

Both abiotic and biotic stress responses

No functional annotation listed

Gene of particular interest; selected for further study

microarray data, using the Gene Ontology tool available online at the DAVID Bioinformatics Resources (Huang et al., 2008). This tool has a database with entries for all of the genes in the Arabidopsis genome and a list of functional description terms associated with those genes and their gene products (*e.g.* 'transcriptional activator' or 'calcium-binding EF hand'). The tool takes a user-defined input list of genes, fetches the functional terms from the database associated with each of them, and clusters the genes based on common functions. Terms that relate to similar functions are grouped into clusters, the clusters are given an 'enrichment score', and it is those clusters of functional terms that form the basis of a gene ontology readout.

The results for the most highly enriched clusters of annotations are presented in Table 4.2 (a full list of GO clusters can be found in Supplementary File 2 - see List of Accompanying Material, page xiii). The cluster with the highest enrichment score ('Annotation Cluster 1') contains annotations related to transcriptional modulation and responses to hormone signalling. The next three clusters are associated with abiotic stress responses (particularly those related to water balance and temperature regulation). Annotations are ranked within each cluster in descending order of fold-enrichment. One drawback of such gene ontology analysis is that the same transcript can be associated with more than one functional annotation within a cluster as well as across clusters, so the number of terms listed in a cluster does not directly correspond to number of unique genes, however, the large fraction associated with gene expression in cluster 1 shows qualitatively that altering *MYB64* levels causes many other downstream transcriptional pathways to be differentially regulated. Also of note is the fact that the most highly enriched annotations relate to ethylene stimulus (approx. 12-fold), cold stimulus (approx. 8-fold), abscissic acid stimulus (approx.

Annotation Cluster 1							
Term	Count	%	P Value	Fold Enrichment	Bonferroni	Benjamini	FDR
DNA-binding region:AP2/ERF	9	9	1.59E-08	18.22	1.92E-06	1.92E-06	1.82E-05
Response to chitin	8	8	3.56E-07	16.96	8.26E-05	2.07E-05	4.57E-04
Pathogenesis-related transcriptional factor and ERF, DNA-binding	9	9	5.63E-08	16.82	6.59E-06	6.59E-06	6.44E-05
AP2	9	9	2.24E-07	13.08	5.60E-06	5.60E-06	1.82E-04
Response to carbohydrate stimulus	9	9	5.47E-07	12.17	1.27E-04	2.54E-05	7.02E-04
Ethylene signaling pathway	7	7	2.44E-05	12.11	1.61E-03	1.61E-03	2.49E-02
Ethylene mediated signaling pathway	7	7	3.36E-05	11.08	7.77E-03	1.11E-03	4.32E-02
Two-component signal transduction system (phosphorelay)	7	7	1.34E-04	8.64	3.05E-02	3.44E-03	1.71E-01
Response to ethylene stimulus	8	8	3.67E-05	8.41	8.48E-03	1.06E-03	4.71E-02
Activator	9	9	2.69E-05	7.47	1.77E-03	8.87E-04	2.75E-02
Response to hormone stimulus	17	17	7.28E-08	5.03	1.69E-05	5.63E-06	9.35E-05
Response to endogenous stimulus	18	18	2.87E-08	4.97	6.65E-06	3.33E-06	3.68E-05
Response to organic substance	21	21	1.62E-09	4.81	3.76E-07	3.76E-07	2.08E-06
Hormone-mediated signaling	7	7	4.64E-03	4.36	6.60E-01	5.52E-02	5.80E+00
Cellular response to hormone stimulus	7	7	4.64E-03	4.36	6.60E-01	5.52E-02	5.80E+00
Regulation of transcription, DNA-dependent	13	13	4.27E-04	3.21	9.43E-02	8.22E-03	5.46E-01
Regulation of RNA metabolic process	13	13	4.49E-04	3.20	9.89E-02	7.97E-03	5.74E-01
Transcription regulation	14	14	4.13E-04	3.17	2.69E-02	9.04E-03	4.21E-01
Transcription	14	14	5.03E-04	3.10	3.26E-02	8.26E-03	5.13E-01
Transcription	14	14	6.47E-04	2.87	1.39E-01	1.07E-02	8.28E-01
DNA-binding	13	13	5.37E-03	2.49	2.99E-01	6.86E-02	5.36E+00
Transcription factor activity	15	15	1.83E-03	2.45	1.62E-01	1.62E-01	2.00E+00
Intracellular signaling cascade	7	7	6.61E-02	2.39	1.00E+00	4.70E-01	5.84E+01
Regulation of transcription	17	17	1.63E-03	2.26	3.15E-01	2.49E-02	2.07E+00
Transcription regulator activity	15	15	6.15E-03	2.15	4.47E-01	1.79E-01	6.57E+00
Nucleus	16	16	8.18E-03	2.09	4.18E-01	7.45E-02	8.06E+00
DNA binding	16	16	2.10E-02	1.80	8.69E-01	3.99E-01	2.08E+01
Annotation Cluster 2							
Term	Count	%	PValue	Fold Enrichment	Bonferroni	Benjamini	FDR
Response to cold	7	7	1.92E-04	8.09	4.36E-02	4.45E-03	2.46E-01
Response to temperature stimulus	8	8	2.77E-04	6.08	6.22E-02	5.83E-03	3.55E-01
Response to abiotic stimulus	12	12	3.34E-03	2.70	5.40E-01	4.74E-02	4.21E+00
Annotation Cluster 3							
Term	Count	%	PValue	Fold Enrichment	Bonferroni	Benjamini	FDR
Response to abscisic acid stimulus	9	9	7.97E-06	8.50	1.85E-03	3.08E-04	1.02E-02
Response to salt stress	5	5	4.50E-02	3.65	1.00E+00	3.72E-01	4.46E+01
Response to osmotic stress	5	5	5.69E-02	3.37	1.00E+00	4.33E-01	5.29E+01
Response to abiotic stimulus	12	12	3.34E-03	2.70	5.40E-01	4.74E-02	4.21E+00
Annotation Cluster 4							
Term	Count	%	PValue	Fold Enrichment	Bonferroni	Benjamini	FDR
Response to water deprivation	5	5	3.81E-03	7.65	5.87E-01	5.08E-02	4.78E+00
Response to water	5	5	4.55E-03	7.28	6.53E-01	5.71E-02	5.68E+00
Response to abiotic stimulus	12	12	3.34E-03	2.70	5.40E-01	4.74E-02	4.21E+00
Annotation Cluster 5							
Term	Count	%	PValue	Fold Enrichment	Bonferroni	Benjamini	FDR
EF-Hand type	4	4	1.68E-02	7.33	8.63E-01	6.29E-01	1.76E+01
EF-hand 2	3	3	6.40E-02	7.02	1.00E+00	9.82E-01	5.33E+01
EF-hand 1	3	3	6.50E-02	6.96	1.00E+00	9.34E-01	5.38E+01
EF-HAND 2	4	4	1.99E-02	6.87	9.05E-01	5.44E-01	2.06E+01
Calcium-binding EF-hand	3	3	7.95E-02	6.34	1.00E+00	8.56E-01	6.12E+01
EF hand	3	3	8.48E-02	6.11	1.00E+00	8.23E-01	6.37E+01
EF-HAND 1	4	4	4.28E-02	5.09	9.94E-01	7.22E-01	3.93E+01
EFh	3	3	1.18E-01	4.93	9.56E-01	7.91E-01	6.38E+01
Calcium ion binding	7	7	3.13E-03	4.74	2.60E-01	1.40E-01	3.39E+00
Calcium	5	5	3.38E-02	4.07	8.97E-01	2.23E-01	2.96E+01

Table 4.2 Gene Ontology Term Enrichment Among Genes Measurably 'Upregulated' in the 35Spro:MYB64 Line

The 100 most highly 'upregulated' transcripts in the MYB64 overexpressing line are significantly enriched for functions associated with gene expression and DNA binding, hormone response (especially ethylene and ABA), and abiotic stimulus response (especially related to water and temperature stress). For a full list see Supplementary File 2.

8-fold), water stimulus (approx. 7-fold), Ca^{2+} signalling (approx. 4 to 7-fold), and temperature stimulus (approx. 6-fold). This represents a greater diversity of functional relationships than had been expected for a single transcription factor originally identified from a salt tolerance screen .

4.4 Identification of Genes of Interest

This gene ontology analysis revealed that *MYB64* upregulation leads to changes in a variety of distinct stress response pathways. If we were to investigate exactly how *MYB64* is involved, arguably the simplest and most logical way to begin would be to look for connections to the 'master regulators' of such networks. The list of the 50 + 1 most highly 'upregulated' transcripts in 35Spro:MYB64 line 141 (Table 4.1) was used as a shortlist, then the Arabidopsis gene annotation database (TAIR) was interrogated manually one at a time to find those with specific functional annotations that suggested they might be such master regulators, and finally the literature was consulted in more detail in order to finalise shortlist of those genes that warranted further study based on their importance in stress responses. Below are the results of this investigation with a brief summary of the reasons each gene was selected (also marked in yellow in Table 4.1).

CBF1 (C-repeat Binding Factor 1; At4g25490) encodes a classic and well-characterised cold-responsive transcription factor (Seki et al., 2001, Fowler and Thomashow, 2002, Maruyama et al., 2004, Canella et al., 2010). Transcription factors in this family bind the C-repeat (CRT) / Drought Responsive Element (DRE) in promoters of genes that typically enable plants to tolerate chilling, freezing, and other abiotic stresses.

CZF2 (Cold-induced Zinc Finger 2; At5g04340), also known as *ZAT6* (Zinc finger of Arabidopsis Thaliana 6), is a member of a family known to respond to drought, low temperature, and salinity (Sakamoto et al., 2004, Mittler et al., 2006). It is not inconceivable that some of the stress responses exhibited by plants to seemingly quite different stresses may, in fact, be designed to implement similar changes to cellular architecture. It has been shown to affect root development (Devaiah et al., 2007). Members of the CZF/ZAT family have recently been shown to inhibit expression of sequences associated with auxin-mediated growth and also to inhibit negative regulators of ABA-directed growth suppression (Kodaira et al., 2011), and also to positively regulate the germination of seeds at high salinity as a result of MAP Kinase phosphorylation (Liu et al., 2013). It was decided that *CZF2*, therefore, represented a further interesting extension to the salt / heat interaction first uncovered by the HT5 / *MYB64* experiment.

PP2C (Protein Phosphatase 2C; At5g59220) belongs to a large family of protein phosphatases, of which there are at least 76 of subtype 2C in Arabidopsis, and this one clusters on the same phylogenetic branch as *ABI1* and *ABI2* (Schweighofer et al., 2004, Zhang et al., 2013). ABA is one of the best-known stress response phytohormones, so it is perhaps not surprising that a transcription factor identified from a stress tolerance screen has here been shown to go on to activate a component of this pathway. If this interaction is proven to be genuine – that is, that *MYB64* activates *PP2C* transcription when not artificially overexpressed – it would imply that *MYB64* is under the control of ABA. These three phosphatases comprise a branch of the phylogenetic tree that is distinct from the rest of the PP2Cs, so although there is no direct evidence

that this PP2C is responsible for the transduction of ABA signals it is possible that it acts redundantly with one of the other two ABI proteins in sensing ABA levels.

ERF11 (Ethylene Response Factor 11; At1g28370) encodes a transcriptional repressor from a family involved in the transduction of stress signals in Arabidopsis. Ethylene is another phytohormone produced when plants become stressed (Kendrick and Chang, 2008), and it has a well-established link to biotic stresses (Adie et al., 2007, Zhao and Qi, 2008). This raises the suggestion that *MYB64* might act as a node between abiotic and biotic stress responses. This hypothesis has been supported recently by the discovery that *ERF11* has been shown to play a key role in the suppression of ethylene signalling by binding to the promoter regions of two genes in the ethylene biosynthesis pathway (*ACS2* and *ACS5*) (Li et al., 2011)

MPK11 (MAP Kinase 11) encodes an uncharacterised kinase related to the better-studied *MPK4* which has been shown to play a role in various abiotic stress responses, including the ABA-induced closure of stomata (Hettenhausen et al., 2012) as well as biotic stresses such as challenge with elicitor peptides (Bethke et al., 2012). In its capacity as a kinase, MPK11 may be responsible for activating a suite of proteins that go on to prime the plant to become stress tolerant. As another potential 'high-level' component in the signalling pathway, *MPK11* was chosen for further study for its likelihood of representing a more appealing biotechnology target than effectors such as heat shock proteins.

4.5 qPCR Validation of NGS Transcript Profiles of Wild Type vs. 35Spro:MYB64

While the next-generation sequencing experiment was designed to be robust in that it was performed on a large pool of 64 individual plants to reduce noise, and that it also provided large amounts of high-resolution quantitative data, it was important to confirm the most significant findings by quantitative RT-PCR (qPCR). Primers were designed to the set of transcripts discussed in section 4.4, which encode well-characterised stress response genes (for primer sequences, see Table 2.13). Before carrying out any measurements it was important to consider which reference gene(s) to use to normalise between samples.

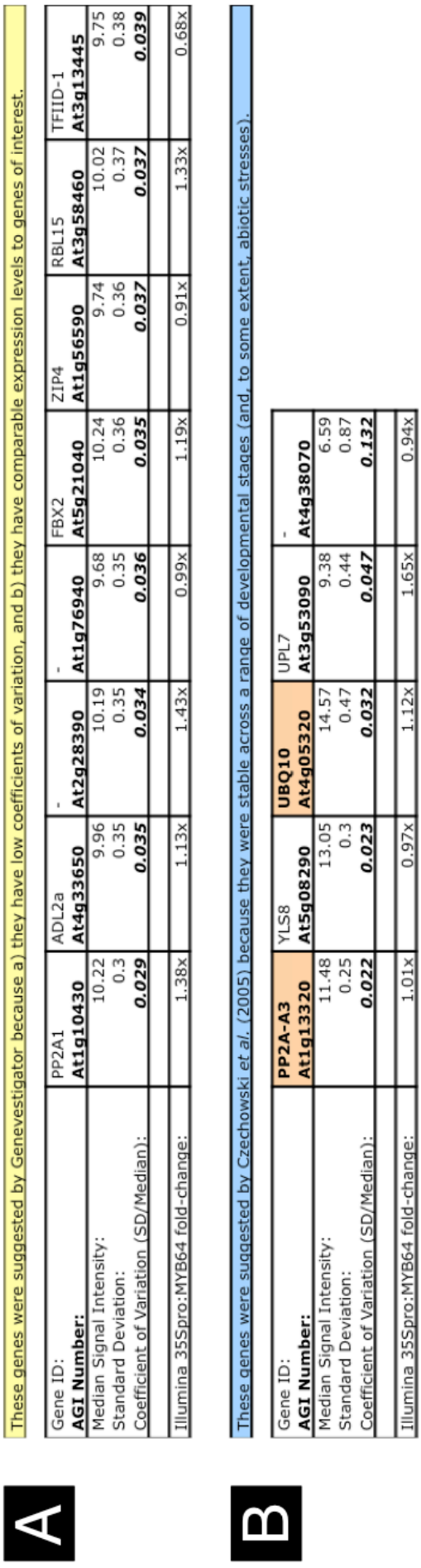
4.5.1 Reference Gene Selection

Actin 2 (At3g18780) was initially selected as an internal reference gene candidate because of its widely assumed stable expression pattern which has led to its use as a suitable reference in the field of plant molecular biology, but this was proven not to be the case when *MYB64* was constitutively overexpressed (NGS results revealed that levels of Actin 2 transcript in transgenic 35Spro:MYB64 line 141 were 8 % lower than in wild type, after normalisation for RPKM, and the work of Nurcahyanti (2009) proved that its expression was affected by heat stress. An effort was made, therefore, to identify a gene (or genes) that would vary less in expression between lines or with application of abiotic stresses.

The Biomarker Selection tool available online at Genevestigator (available at www.genevestigator.net, Hruz et al., 2008) provided a list of transcripts with

low variability based on 4728 individual microarray experiments (performed on the ATH1 22k Affymetrix Genome Array chip using seedlings subjected to a variety of abiotic stresses). Selection criteria were that a) absolute expression levels should be similar to that of transcripts being measured, and b) the coefficient of variation (*i.e.* standard deviation / mean expression level) should be as low as possible. The list of suggested genes were ranked from lowest to highest coefficient of variation, and then those at the low end of the list were cross-checked for minimal fold changes in the 35Spro:MYB64 NGS data. The top 8 candidate genes were found to differ from wild type by up to a third (from 0.68x to 1.43x, Table 4.3, panel **A**), *i.e.* more than Actin 2.

A literature search was also undertaken to find out whether theoretical results like these had already been validated experimentally. Czechowski et al. (2005) provide a set of reference gene candidates whose expression was found by microarray to be stable across the full range of developmental stages and, to some extent, abiotic stresses (Table 4.3, **panel B)**. Cross-referencing these data with the somewhat larger abiotic stress microarray dataset generated by the community and available at Genevestigator, as well as with our NGS data, allowed some of those candidates to be discounted and the list was narrowed to two on the basis of being more stably expressed than others: a protein phosphatase subunit (PP2A-A3; At1g13320) involved in hypocotyl and root development (Blakeslee et al., 2008), and polyubiquitin 10 (UBQ10; At4g05320), one of five genes in the Arabidopsis genome encoding ubiquitin, a protein which is added to substrates targeted for degradation (Bachmair et al., 2001) (Table 4.3, **panel B**, cells highlighted in orange). As primers had already been designed, tested and published (Czechowski et al., 2005), it was concluded that these two



A

These genes were suggested by Genevestigator because a) they have low coefficients of variation, and b) they have comparable expression levels to genes of interest.

Gene ID:	PP2A1	ADL2a	-	-	FBX2	ZIP4	RBL15	TFIID-1
AGI Number:	**At1g10430**	**At4g33650**	**At2g28390**	**At1g76940**	**At5g21040**	**At1g56590**	**At3g58460**	**At3g13445**
Median Signal Intensity:	10.22	9.96	10.19	9.68	10.24	9.74	10.02	9.75
Standard Deviation:	0.3	0.35	0.35	0.35	0.36	0.36	0.37	0.38
Coefficient of Variation (SD/Median):	***0.029***	***0.035***	***0.034***	***0.036***	***0.035***	***0.037***	***0.037***	***0.039***
Illumina 35Spro:MYB64 fold-change:	1.38x	1.13x	1.43x	0.99x	1.19x	0.91x	1.33x	0.68x

B

These genes were suggested by Czechowski *et al.* (2005) because they were stable across a range of developmental stages (and, to some extent, abiotic stresses).

Gene ID:	**PP2A-A3**	YLS8	**UBQ10**	UPL7	-
AGI Number:	**At1g13320**	**At5g08290**	**At4g05320**	**At3g53090**	**At4g38070**
Median Signal Intensity:	11.48	13.05	14.57	9.38	6.59
Standard Deviation:	0.25	0.3	0.47	0.44	0.87
Coefficient of Variation (SD/Median):	***0.022***	***0.023***	***0.032***	***0.047***	***0.132***
Illumina 35Spro:MYB64 fold-change:	1.01x	0.97x	1.12x	1.65x	0.94x

Table 4.3 Selection of Candidate Reference Genes for qPCR

Actin 2 had historically been used as a reference gene for semi-quantitative RT-PCR earlier in this body of work, but it was decided that for highly-sensitive quantitative RT-PCR experiments a reference gene should be empirically chosen for its stability of expression levels throughout the course of a heat treatment. A shortlist of candidate reference genes was compiled based on the suggestions of the Biomarker Selection Tool at Genevestigator (Hruz et al., 2008) (**panel A**, yellow) and also based on the experimental observations of Czechowski et al. (2005) (**panel B**, blue). Median signal intensity, standard deviation, and coefficient of variation values presented in these tables were gathered from publically-available expression datasets deposited at Genevestigator. The final *in silico* selection step was to compare the data from our Next-Generation Sequencing data on the 35Spro:MYB64 transgenic and wild-type lines; the final row in each panel shows which of the candidate genes varied least.

Genes PP2A-A3 (protein phosphatase 2A, subunit A3) and UBQ10 (ubiquitin 10) were chosen for experimental validation as reference genes by qPCR (see Figure 4.3) and primers were designed and ordered (for sequences, see Table 2.13).

might be the best candidate reference genes against which to normalise during further qPCR investigations.

Stability of their expression levels following thermal acclimation was tested experimentally. Segregated wild type seeds were sown on agar plates (section 2.3.2) and grown under Normal Conditions for 2 weeks. Plates were then incubated at the acclimation temperature of 37 ± 0.5 °C for 3 hours. Seedlings were harvested before, immediately after, 25 hours after, and 48 hours after this acclimation step. RNA was extracted and cDNA prepared (sections 2.6.1 and 2.6.3) and qPCR was carried out using primers to Actin 2, UBQ10, and PP2A-A3 (Figure 4.3). UBQ10 level varied most: immediately after acclimation the transcript abundance had almost tripled (2.77-fold) then, following a dip to 1.97-fold at 25 hours, the level once again increased to 10.1-fold higher than it had been initially. Actin 2 levels were the next most variable: a small elevation of 1.57-fold directly followed the acclimation period, rising to 4.70-fold at 25 hours and 4.51-fold at 48 hours. PP2A-A3 varied least, exhibiting 0.87-fold, 1.90-fold and 3.05-fold changes at each of the three timepoints. Based on these data, PP2A-A3 was chosen as the reference gene for further qPCR studies.

4.5.2 qPCR Comparison of Non-Stressed Lines

In order to validate the NGS results regarding the genes downstream of *MYB64*, qPCR was performed on mature, non-stressed, wild type and 35Spro:MYB64 plants. Two independent transformant lines were sampled to find out whether the use of different vectors led to any difference in molecular phenotype. Seeds were sown on soil (section 2.3.1) and grown under Normal Conditions until 4 weeks old. Rosette leaves were harvested (~15 plants per sample) and RNA

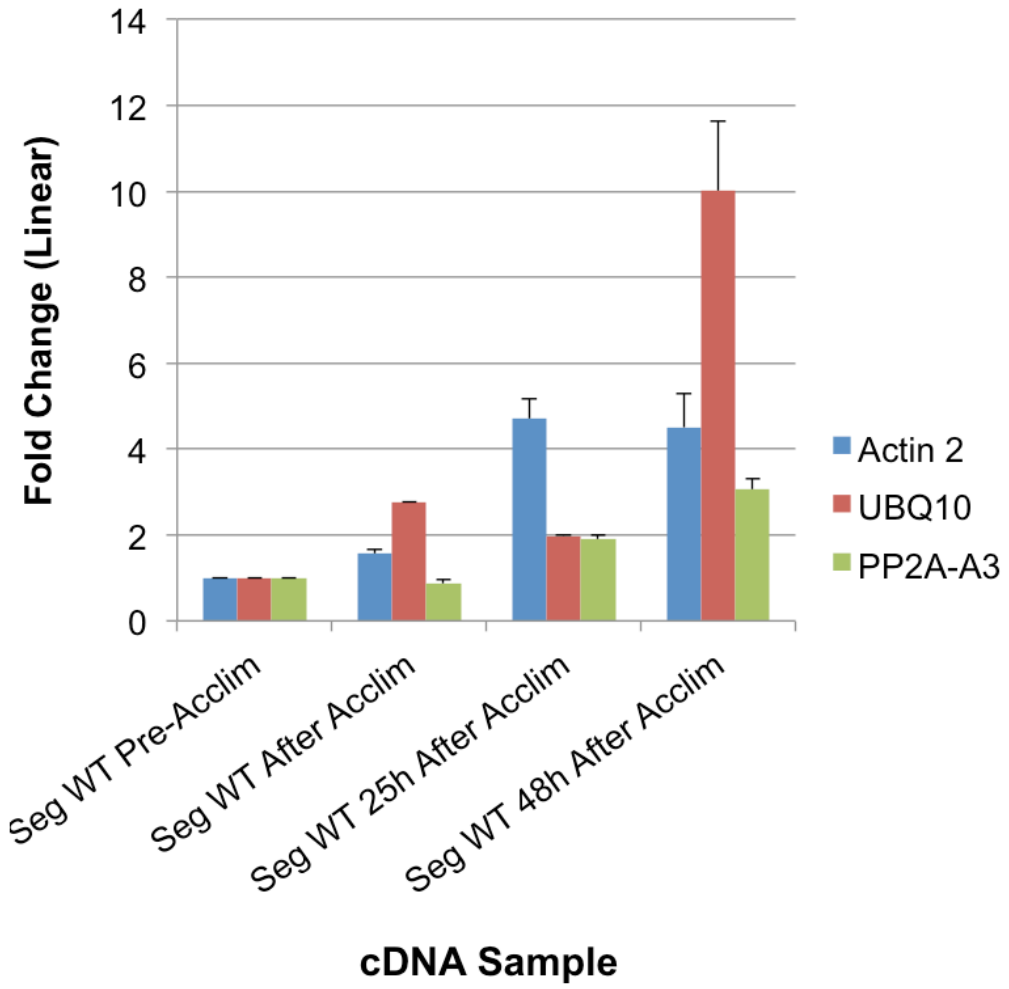


Figure 4.3 qPCR Validation of Reference Gene Candidates

The two candidate reference genes identified *in silico*, PP2A-A3 and UBQ10 (section 4.5.1, "Reference Gene Selection"), were experimentally compared with Actin 2, the control used previously in semi-quantitative RT-PCR experiments, to validate their use as reference genes in subsequent qPCR experiments involving heat treatments. Expression levels of the 3 transcripts in segregated wild type were followed for 48 hours after a heat acclimation at 37 °C (± 0.5 °C) for 3 hours. Error bars represent means ± SE; *n*=3

extracted to prepare cDNA (sections 2.6.1 and 2.6.3). The results of this no-stress qPCR experiment are presented in Figure 4.4. In line 35Spro:MYB64-141, the two *MPK11* splice variants – *MPK11.1* and *MPK11.2* – were upregulated 12 and 23-fold, *CZF2* 45-fold, *ERF11* 145-fold, and *MYB64* 198-fold. This is in strong agreement, at least qualitatively, with the NGS transcript profiling results (Table 4.1). The expression levels from line 35Spro:MYB64-127 are somewhat different from these data; the downstream transcripts are still more abundant than in the segregated wild type but to a lesser extent. This does not directly contradict the NGS results as the data came from two different lines. The variation in *MYB64* upregulation between the two lines (252-fold in line 127 vs. 198-fold in line 141) suggests the possibility of a dose-dependent interaction of *MYB64* on downstream genes, or a more complex mode of action for *MYB64* than simply activating transcription; perhaps post-transcriptional or post-translational modification, which is carried out differently in each of the two lines.

4.5.3 qPCR Comparison of Heat-Acclimated Wild Type vs. 35Spro:MYB64 Line 127

If *MYB64* was not consistently driving the activation transcription of these 'downstream' genes, it might be that its role is instead in the potentiation of stress-related gene expression. If this were the case, the hypothesis follows that expression levels detected in the 35Spro:MYB64 lines should be *further* enhanced upon delivery of an appropriate abiotic stress. Since parallel experiments were under way to characterise the heat stress response of these plants (see Chapter 4) and to establish whether there was a connection between *MYB64* and the small heat shock proteins (see Chapter 3), high temperature was chosen as the abiotic stress to use here also.

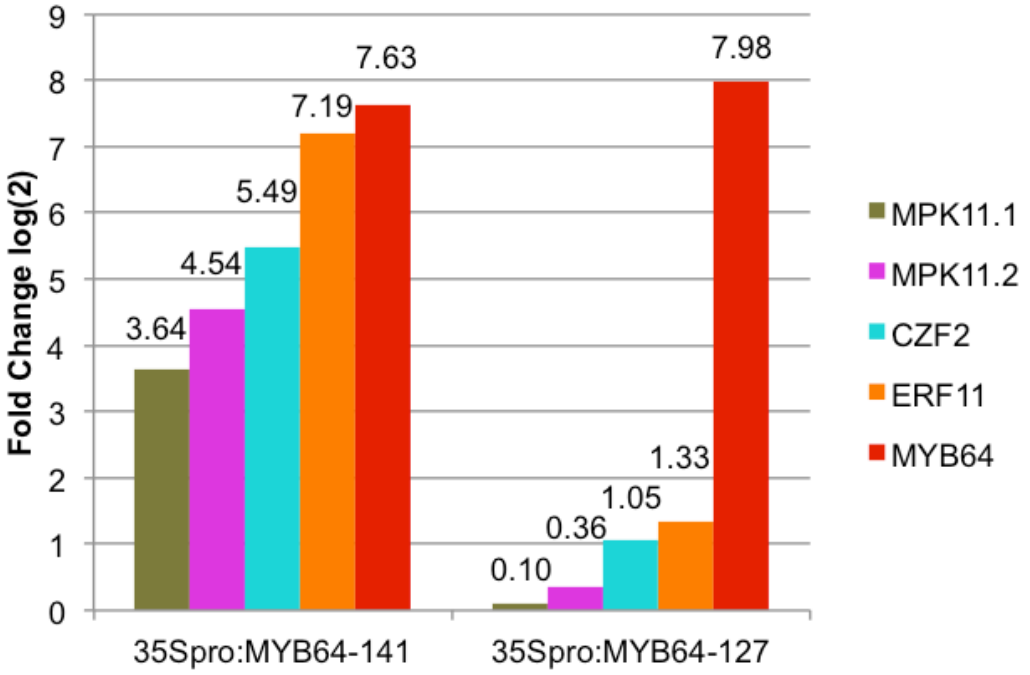


Figure 4.4 qPCR Analysis of Selected Over-abundant Transcripts Relative to Segregated Wild Type

Several genes of interest identified from the NGS results (see Table 4.1) were examined (*MPK11* has two splice variants) to confirm the degree of 'upregulation' in two independent 35Spro:MYB64 transformants generated with different vectors. Biological replicates: n = 15. Technical replicates: n = 1.

Seeds of 35Spro:MYB64 line 127 and segregated wild type were sown on agar plates (section 2.3.2) and grown under Normal Conditions for 10 days. Plates were acclimated at 37 ± 0.5 °C for 3 hours and samples were harvested for RNA extraction and cDNA preparation (sections 2.6.1 and 2.6.3) either before or 24 hours after this treatment. qPCR was performed using primers to *HSP17.6*, *HSP17.6a*, *ERF11*, *CZF2*, and both *MPK11* splice variants. The expression level of each transcript in the 35Spro:MYB64 line was normalised to its expression in the segregated wild type given the same treatment and sampled at the same time, and values were plotted on a $\log_2$ scale (Figure 4.5).

Before acclimation, each transcript (with the obvious exception of *MYB64*) was expressed at roughly equivalent levels in each of the two lines (**panel A**). This was roughly in agreement with the qPCR results for this line (127) presented in Figure 4.4. Twenty-four hours after acclimation, however, there were substantial differences (**panel B**). Expression of the two smHSPs was *lower* in the transgenic line than in the segregated wild type, in contrast to the findings of the HT5 microarray (though that microarray represented a response to salt stress). The *ERF11* level in the transgenic line was higher by just over two logs. *CZF2* expression in the transgenic line was just slightly more than one log higher than that measured in the wild type. Both splice variants of *MPK11* were more highly increased: approx. 3.5 and 4 logs for *MPK11.1* and *MPK11.2*, respectively.

These findings are consistent with the qPCR results presented in Figure 4.4 in that line 127 overexpresses *MYB64* transcript several hundred-fold under Normal Conditions. The expression of the downstream genes is not significantly altered until an external stimulus is applied – in this case, heat acclimation.

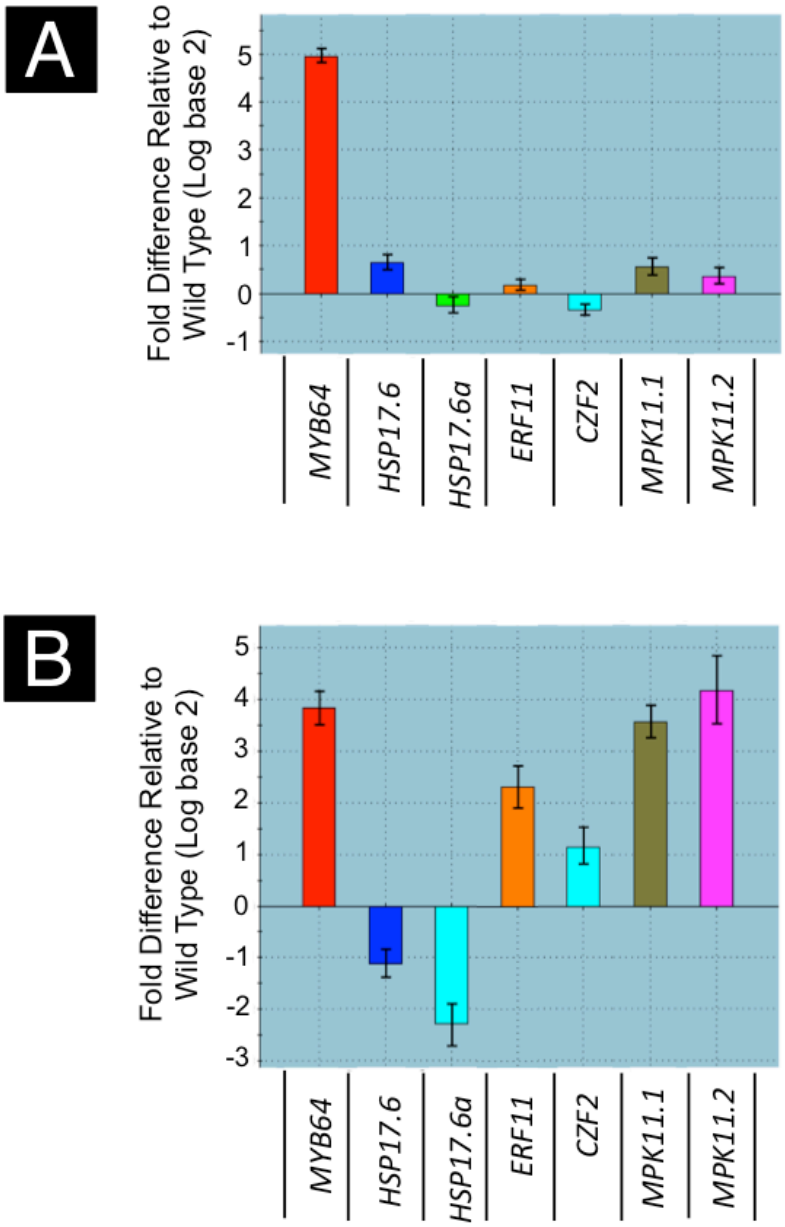


Figure 4.5 Effects of Heat Acclimation on Transcript Levels of Genes of Interest in 35Spro:MYB64 Line 127 Relative to Segregated Wild Type

A selection of upregulated genes of interest were investigated to determine whether their expression would increase even further upon heat acclimation. Seedlings of 35Spro:MYB64 line 127 and segregated wild type were grown on plates for 10 days. **A** Plates designated as controls were sacrificed before acclimation and the seedlings harvested for RNA extraction. **B** Seedlings were acclimated at 37 ± 0.5 °C for 3 hours then returned to the growth room. RNA samples were harvested 24 hours later for qPCR. Values represent expression in the transgenic line normalised against expression in segregated wild type. Error bars represent the mean ± SE of *n* = 2 plates, each with 15 seedlings.

4.6 qPCR Timecourse of Heat-Acclimated Transcript Profile of Segregated Wild Type

There was clearly a connection between these 'downstream' effectors and *MYB64* overexpression, but it was not clear whether *MYB64* directly initiated their expression in wild type plants or whether it acted as a potentiator, maintaining a high level of expression after transcription of those genes had already begun in response to some uncharacterised cue. It was important to determine whether levels of *MYB64* began to rise sufficiently far enough in advance of levels of the 'downstream' genes in order to be able to place them in sequence with each other.

Seeds of segregated wild type were sown on basal agar plates and grown under normal conditions for 10 days, then acclimated for 90 minutes. Samples were taken for qPCR throughout, and up to 15 hours following, acclimation (Figure 4.6).

MYB64 levels increased to approximately 17-fold higher than initial levels during the acclimation period, then returned to baseline levels after 1 hour at 22 °C. They gradually rose to approximately 3-fold higher again 15 hours later. *CZF2* and *PP2C* (panel **B)** both underwent a relatively large increase in expression level before that of *MYB64*. *ERF11*, *MPK11.1* and *MPK11.2* (**panel C**) all underwent minor changes in expression, both positive and negative, during the acclimation period, and gradually rose to between 4 and 6-fold higher 15 hours after the return to 22 °C. These data suggest that *MYB64* is probably not directly activating the expression of *CZF2* or *PP2C* in segregated wild type, but may play a role in activation of *ERF11* and *MPK11*.

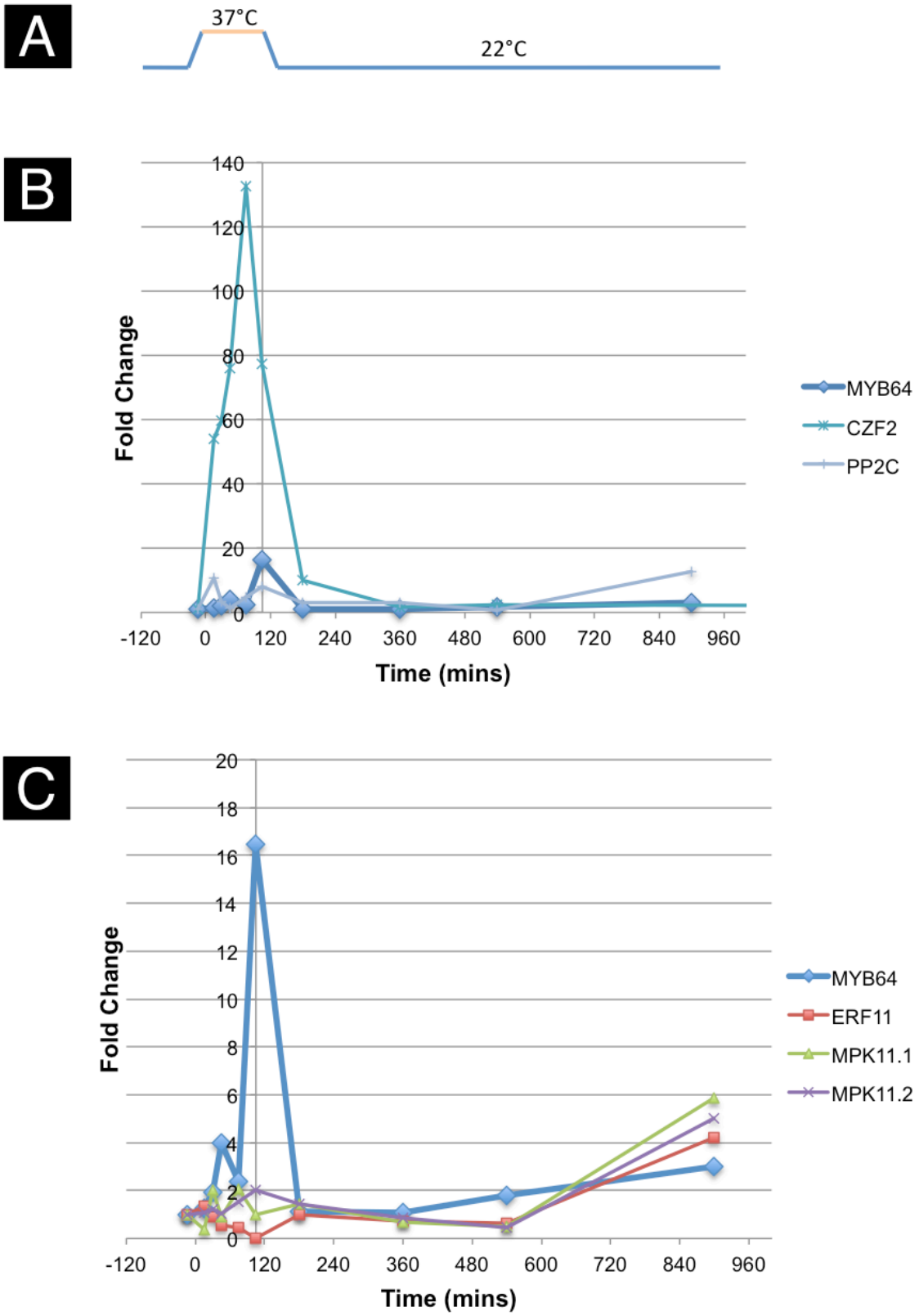


Figure 4.6 qPCR Timecourse of Segregated Wild Type MYB64 and 'Downstream' Gene Expression Upon Thermal Acclimation

To investigate whether *MYB64* expression preceded that of the suspected downstream targets, segregated wild type seedlings were given an acclimation at 37 ± 0.5 °C for 90 mins (**A**) and expression was measured by qPCR throughout, and for the following 15 hours. **B** MYB64, CZF1 and PP2C levels plotted on a 0 – 140 fold scale. **C** MYB64, ERF11, MPK11.1 and MPK11.2 plotted on a 0 – 20 fold scale.

4.7 *MYB64* Transcription Response to ABA

Since overexpression of *MYB64* led to the 'upregulation' of transcripts associated with an ABA response (Table 4.1), and since 35Spro:MYB64 seedlings had an ABA-tolerant phenotype (see Chapter 4), it was hypothesised that *MYB64* might act as one of the transducers of this hormone signal. The transcript levels of *MYB64* in the wild type line were measured before and after a treatment with exogenous ABA.

Segregated wild type plants were grown on soil until four weeks old then sprayed either with 0 or 25 μM ABA dissolved in methanol. RNA samples were harvested before and 3 hours after treatment. cDNA was prepared and qPCR carried out using primers to *MYB64* (Table 2.13). Figure 4.7 shows a significant upregulation of *MYB64* in response to exogenous ABA at a concentration of 25 μM, supporting the hypothesis that it lies upstream of the genes of interest in the Arabidopsis response to ABA.

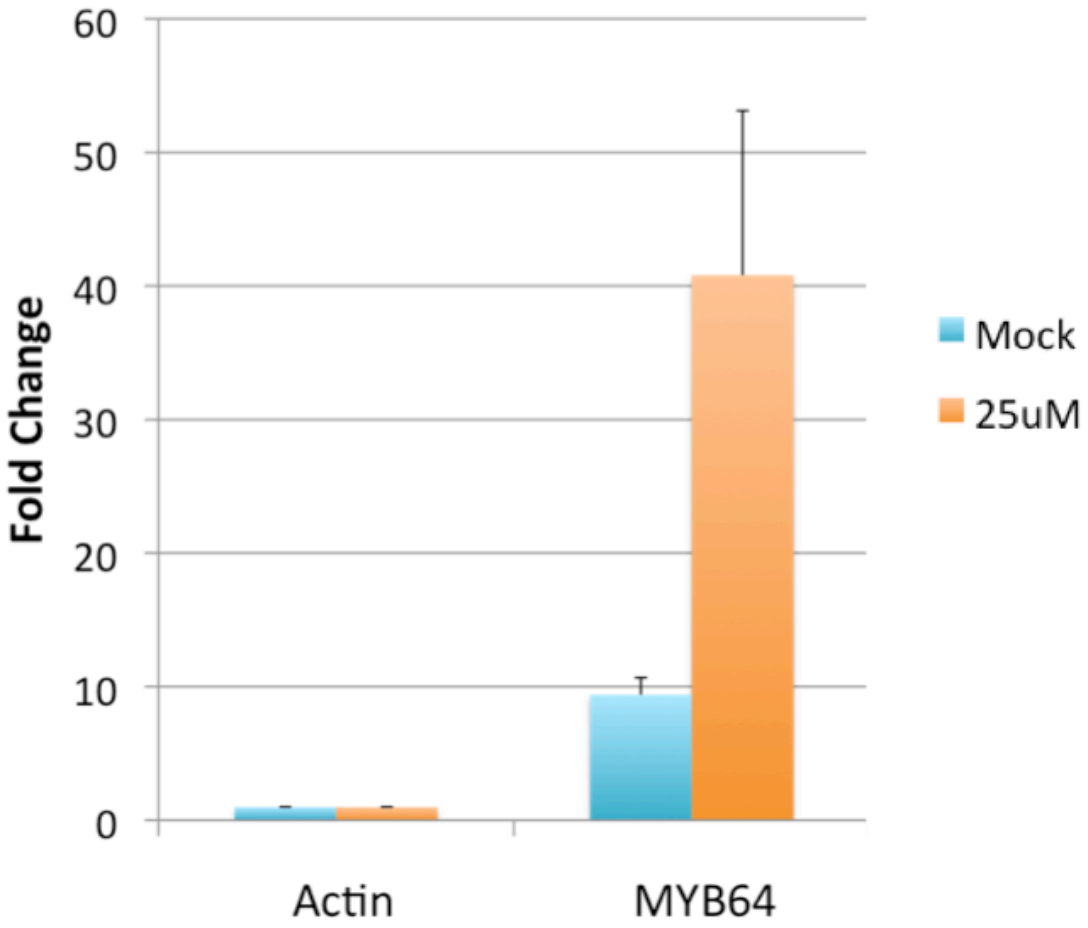


Figure 4.7 ABA Responsiveness of *MYB64* Levels in Segregated Wild Type

Segregated wild type seeds were sown on soil and allowed to grow under normal conditions for 4 weeks. Initial RNA samples were harvested from rosette leaves, then plants were either sprayed with 25 μM ABA (orange bars) or mock treated with a spray of the ABA solvent (methanol; blue bars) and a second RNA sample was harvested 3 hours later. cDNA was prepared and qPCR performed using primers to Actin or to *MYB64*. Fold changes were calculated relative to the initial sample. Error bars represent means ± SE; $n = 2$

4.9 Similar Transcript Profiles Identified in Independent Activation-Tagged Thermotolerant Lines

Following the reverse-genetics approach of the research that led to the identification of the halotolerant line HT5 (Price, 2005), a screen was devised to identify novel mutants with increased *thermotolerance*. The same pool of approximately 22,000 activation-tagged seed was once again acquired from the Nottingham Arabidopsis Stock Centre (NASC, available at www.arabidopsis.info, Scholl et al., 2000). These seeds (T_1) were sown on agar plates supplemented with BASTA, the selective agent for the insertion, grown under normal conditions until 7 days old, then incubated at the T_K for wild type Arabidopsis (44 ± 0.5 °C) for 3 hours. Plates were observed for the following 5 days and seedlings that survived were rescued and placed in soil and allowed to set seed. This T_2 generation was given a secondary screen to confirm the phenotype, identical to the primary screen except that there was no BASTA in the medium. Wild type (Col 0) seeds were included as an internal control. Plants that survived 5 days later were deemed to be authentic positives and named 'Ac-Tag A', 'Ac-Tag B' *etc*. To date, five such lines have been identified. For an example of the primary screen phenotypes see Figure 4.8.

The locus and the genetic nature of each of these novel insertions have not yet been characterised, but following the hypothesis that the surprisingly wide-ranging network of responsive genes triggered by *MYB64* overexpression is a result of extensive stress response cross-talk, it might also be the case that these novel insertions triggered the same transcript profile changes in these independent thermotolerant lines. If such similar transcript profiles were found

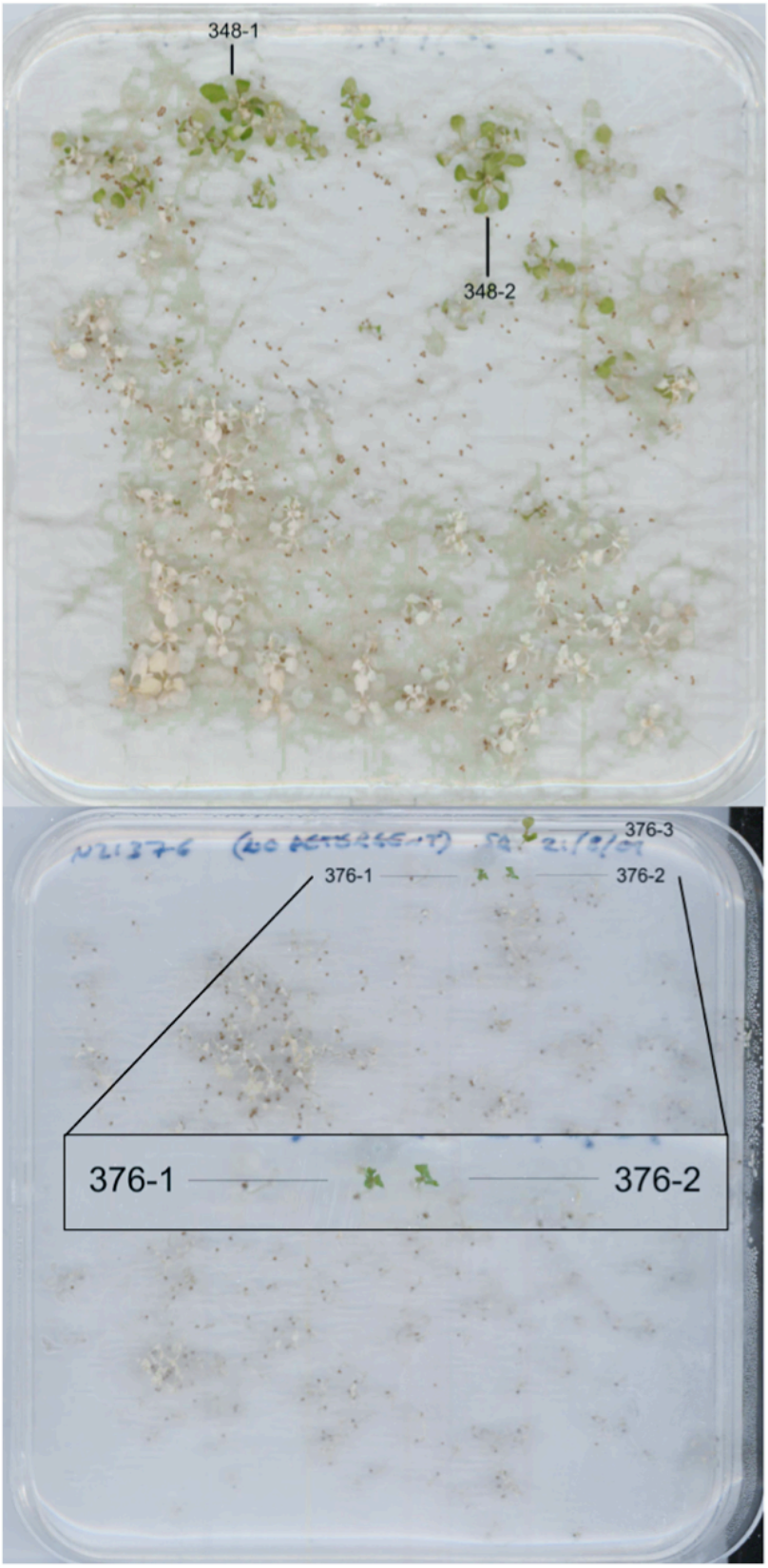


Figure 4.8 Example of Activation-Tagged Mutants Rescued from Thermotolerance Screen

The Weigel set of Arabidopsis seed carrying randomly positioned T-DNA insertions was obtained and screened for thermotolerance by incubating at the killing temperature (44 ± 0.5 °C) for 3 hours. Seedlings that were still alive several days later were rescued for further analysis and numbered according to the pool from which they came.

it would validate the theory that stress response networks are more complex than initially thought. Expression of some of the same set of genes of interest discussed above was therefore investigated in lines Ac-Tag A and Ac-Tag B under non-stress conditions (Figure 4.9).

All 5 genes tested (including two transcripts for *MPK11*) were significantly more abundant in the Ac-Tag lines than in the Col 0 wild type supplied by NASC, *i.e.* the background in which the Ac-Tag mutants were made. (Measurements are not available for *HSP17.6* in 35Spro:MYB64 line 127 due to technical failure; this should be rectified in future work.)

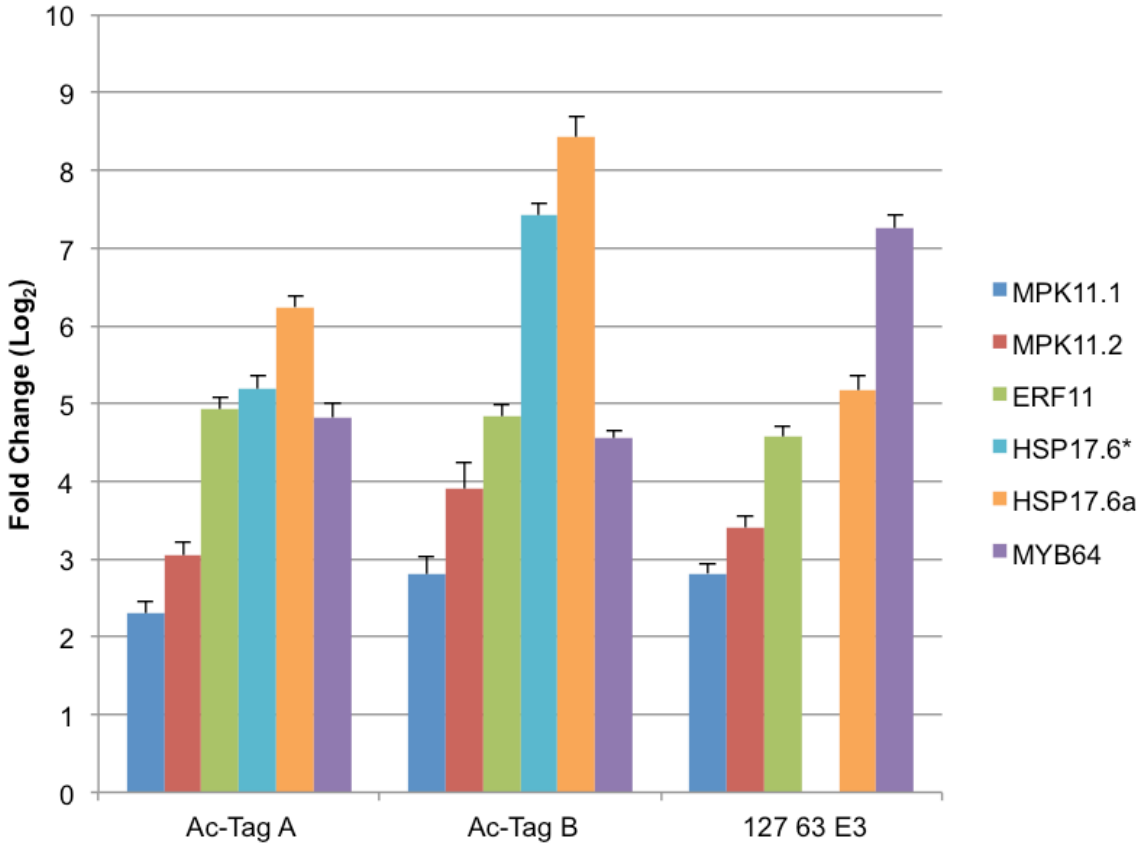


Figure 4.9 Expression of Genes of Interest in Two Independent Ac-Tag Thermotolerant Lines

qPCR was performed to measure the expression levels of the putative 'downstream' genes of interest in two independent, uncharacterised lines carrying activation-tags which were isolated from a thermotolerance screen for novel mutants. The plants sampled here were grown under normal conditions. The locations of the activation-tag inserts in these lines are not known.

*HSP17.6 qPCR results are not yet available for 35Spro:MYB64 line 127 – this should be repeated in further work.

4.10 Discussion

The NGS transcript profile of the 35Spro:MYB64 line showed that the list of the most highly upregulated transcripts were significantly enriched for sequences associated with stress response pathways. These sequences are known to be involved in both biotic and abiotic stresses, and the possibility that the plants had been infected was quickly discounted. There was no sign of pathogen attack; the soil had been pre-treated with insecticide to decrease the likelihood of fly larvae hatching and feeding on the roots; and *PR1* (pathogenesis-related 1; At2g14610), a classical marker of infection (Volko et al., 1998), was not upregulated. Furthermore, that the qPCR confirmation was performed on seedlings raised in the sterile environment of sealed agar plates provides perhaps the strongest evidence that infection is not responsible for the members of biotic stress response pathways noted in the NGS transcript profile.

The genes chosen for further investigation (*CZF2*, *CBF1*, *MPK11*, *ERF11*, *PP2C*) were selected on the basis that they were either well-characterised master regulators of response networks at a transcriptional level, or because they have the capability to alter protein activity by modulation of phosphorylation state. The expression levels of these downstream transcripts were elevated in line 141 which produced the NGS profile, and these transcripts were not elevated in line 127 under normal conditions but were highly upregulated when a heat stress was applied.

Application of ABA induced *MYB64* expression to a level 20 – 30-fold higher than was observed after a mock application, placing this MYB transcription factor in the ABA-dependent branch of the superfamily.

Overall, these results show that overexpression of *MYB64* leads to a clear change in the expression profile induced by thermal acclimation. The independent finding that these transcripts are upregulated in other activation-tagged thermotolerant mutants supports the conclusion that they make a contribution to the survival of the plant under conditions of abiotic stress. There is a wide range of response roles already ascribed to these genes, so future work should characterise the phenotype of the 35Spro:MYB64 lines under a variety of such stresses.

4.10.1 *MYB64* as a Potentiator of Stress Responses?

The transcript profiles of lines 127 and 141 as confirmed by qPCR reveal a difference in the way that MYB64 affects transcription of downstream genes in each line (Figure 4.4 and Figure 4.5). One explanation is that there are different positional effects in each of these independent transformants; differences attributable to disruption of a gene in one or both of the lines caused by the T-DNA insertion. If the average Arabidopsis gene is assumed to be 1-2 kb long and the total Arabidopsis genome is approximately 135 Mb encoding approximately 26,000 genes, then coding regions account for 19 – 39 % of the genome, and the rest is a mixture of regulatory regions and intergenic DNA (pseudogenes *etc*.). The likelihood, therefore, that one or two random T-DNA insertion events interrupted a region that contributes to the control of the *MYB64* regulon is small. This could be confirmed by generating and investigating other independent transformants, but the evidence available suggests that there has been a disruption to *MYB64* itself in one of the lines.

Some MYB TFs, such as MYBR2 and MYB44, are regulated by phosphorylation as part of an ABA-dependent signalling cascade (Pitzschke et al., 2009, Persak and Pitzschke, 2013). If such a regulatory domain were entirely missing or inactive because of a mutation (even as small as a single base substitution at a crucial position) the result would be abnormal transcription of the downstream sequences. This would account for the difference between lines 127 and 141. If it is assumed that line 127 produced the wild type version of MYB64, the failure of this overexpressing line to evoke a full stress response in the absence of a stress factor implies this TF requires modification at an activation domain by some other signalling mechanism. This conclusion is supported by the observation that when line 127 was exposed to heat stress the response was similar to, but significantly greater than in, wild type lines (Figure 4.5). In contrast, unacclimated plants of line 141 demonstrated a superior stress response than stressed wild type plants, suggesting the hypothetical regulatory domain on MYB64 was no longer functional / present, the protein was not dependent on its activation, and the additional signalling mechanism was, therefore, no longer required.

This model requires that the presence of MYB64 protein alone is not enough for expression of the regulon; there must also be an additional signal. The time-dependent qPCR investigation of the regulon in wild type (Figure 4.6) showed that while *MYB64* itself was induced to a much greater extent (approximately 18-fold) than some members of the regulon (with the exception of *CZF2*) many of them were expressed concomitantly, meaning *MYB64* protein was unlikely to have had sufficient time to accumulate and activate their transcription. Levels of *MYB64* began to rise again 900 minutes later, however, as did those of *PP2C*, *ERF11*, and both splice variants of *MPK11*.

The evidence discussed here supports the conclusion that the role of *MYB64* is perhaps to potentiate the expression of those genes designated here as the '*MYB64* regulon': the delay in activation of the regulon in wild type lines; the requirement for heat acclimation to activate the regulon in line 127; and the apparent requirement for a salt stimulus to cause upregulation of the smHSPs in activation-tagged line HT5.

5 MYB64 Overexpression and Stress Tolerance

5.1 Introduction

As discussed in section 1.4.3, in order to recapitulate the phenotype of activation-tagged halotolerant line HT5, a genetic construct was made consisting of the same tetramerised 35S enhancer elements fused directly to the TSS of the *MYB64* gene cloned from genomic DNA (gDNA). Transgenic Arabidopsis lines were created in a Col 0 background and these were experimentally shown to dramatically overexpress *MYB64* at the transcript level as measured by semi-quantitative RT-PCR (Price and Ramsay, unpublished). This line will be referred to as 35Spro:MYB64. Microarray analysis of the HT5 line grown under a long-term salt-stress had revealed that 15% of the 45 most upregulated genes belonged to the HSP families (Price, 2005). Preliminary experiments carried out by Price, and also by Nurcahyanti (2009), showed that there was a thermotolerant phenotype and that the smHSPs discussed in Chapter 3 were indeed upregulated. It was therefore of interest to investigate the thermotolerance profile of the new 35Spro:MYB64 lines.

5.2 Thermotolerance of 35Spro:MYB64 Lines

5.2.1 Thermotolerance of 35Spro:MYB64: Overall Morphology

Two independent 35Spro:MYB64 transformant lines (line 127 and 141, both homozygous) were tested for enhanced thermotolerance as compared with a wild type derived from the heterozygous, segregating progeny of the original transformation. In addition, a second, external wild type control was also used

to verify that the segregated wild type behaved similarly to that used in previous experiments where the growth of Col 0 acquired from an independent source had formed the baseline (*e.g.* section 3.5.1). Seeds were sown on agar plates and germinated horizontally under Normal Conditions. Plates were divided into 4 areas for different genotypes in order to remove as many sources of technical variation as possible and seeds were sown at a low enough density that leaves from neighbouring plants would not touch each other. Seedlings were allowed to grow for 2 weeks and were then subjected to the heat stress regime described above. Plates were qualitatively assessed at 2.5 days, 6 days, and 9 days post-stress. Plates were set up in triplicate and representative images are presented in Figure 5.1.

Col 0 and the segregated wild type seedlings (white coded quadrants) suffered an observable loss of green colouration in most of their leaves and did not develop any new growth in any of the 9 days following heat stress. The two homozygous 35Spro:MYB64 lines (green coded quadrants) suffered an equivalent loss of colouration and also did not develop any new growth. Any variation between these representative plates and the replicate plates that are not shown was between-plate variation; where a greater number of wild type seedlings in a quadrant had completely lost all colouration, an equivalent proportion of 35Spro:MYB64 seedlings on that plate also lost all colouration. The plates in Figure 5.1 were selected as they appear to represent the phenotype of plants subjected to a temperature just on the border of complete killing, and as such would theoretically have shown the most pronounced difference between a wild type and a thermotolerant line.

Figure 5.1 35Spro:MYB64 Lines Show Weak Thermotolerance at 2 Week Stage

Two independent homozygous transformants of 35Spro:MYB64 (line 127 and line 141 - green quadrants in top panel) were sown alongside segregated wild type (right-hand white quadrants) on duplicate plates. One of the duplicate plates also carried an independent wild type for comparison (Col 0 from a different source – upper left white quadrant). A third independent homozygous transformant was also included on the other duplicate plate (upper left green quadrant), but this seed stock failed to germinate. Seedlings were grown for 2 weeks under Normal Conditions. Plates were then exposed to thermal conditions previously determined to be sufficient to kill the segregated wild type: 1 hour acclimation at 37 ± 0.5 °C; 3 hour interval period at Normal Conditions; 3 hours at 44 ± 0.5 °C. Seedlings were photographed at the timepoints indicated until it was evident that no further change in viability, as measured by green leaf area, was taking place.

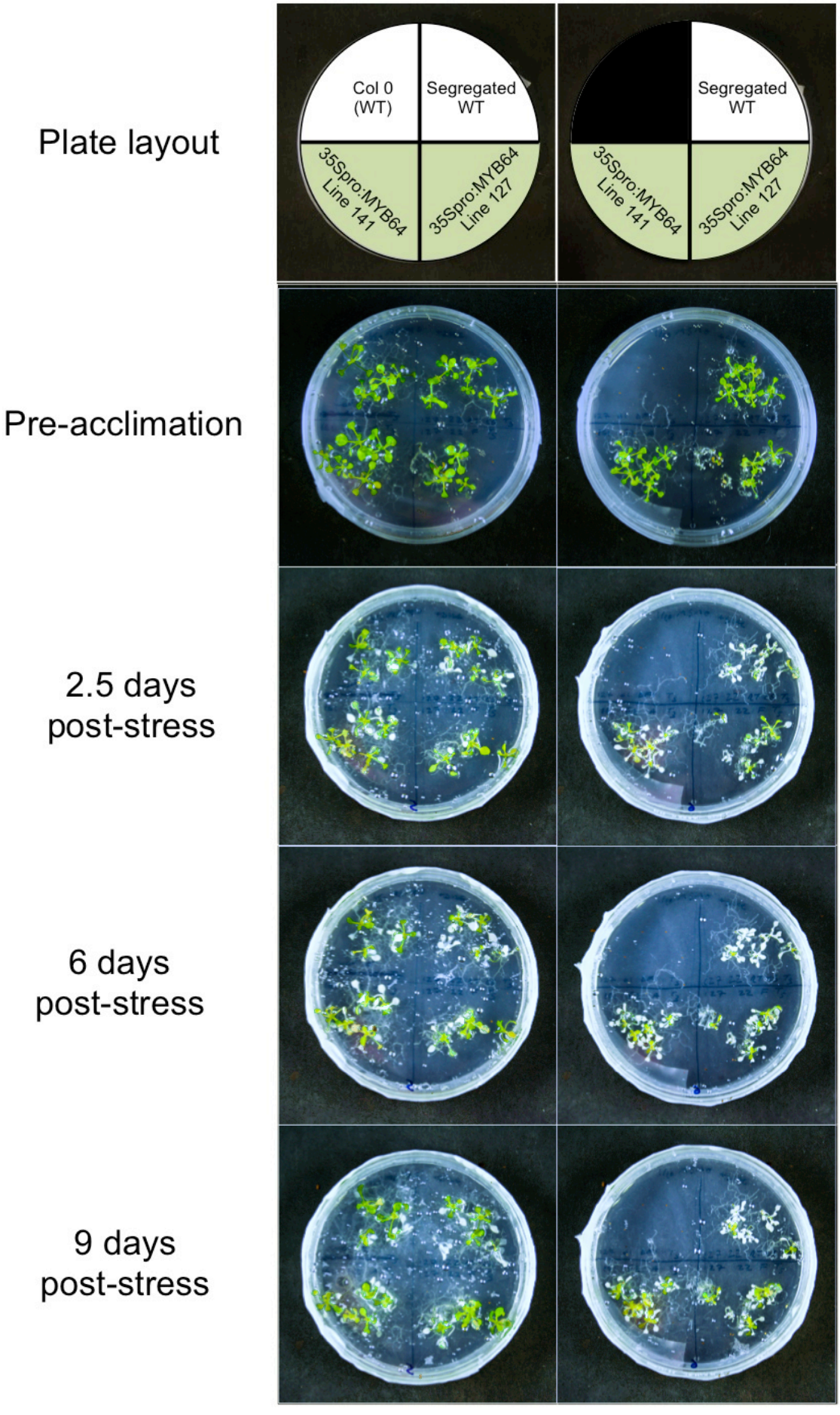


Plate layout
Col 0 (WT)
Segregated WT
35Spro:MYB64 Line 141
35Spro:MYB64 Line 127
Segregated WT
35Spro:MYB64 Line 141
35Spro:MYB64 Line 127
Pre-acclimation
2.5 days post-stress
6 days post-stress
9 days post-stress

From these results it was concluded that although increased *MYB64* expression had led to an increased abundance of HSP transcripts in the salt-stressed activation-tagged HT5 line, overexpression of a 35Spro:MYB64 transgene confers only a weak thermotolerant phenotype upon these 2-week old seedlings. Further investigations were carried out to look for thermotolerance at other stages of development.

5.2.2 Thermotolerance of 35Spro:MYB64: Germination

It is known that members of the small HSP gene family (smHSPs) are upregulated in siliques thereby implicating a functional role in seed maturation, possibly through dormancy to germination (Kotak et al., 2007b). Although it had not been proven experimentally, it is logical to assume that the near-constitutive (Weigel et al., 2000) 35Spro:MYB64 transgene would lead to overexpression of *MYB64* at this stage too. To investigate whether this overexpression might therefore confer a thermoprotective effect during germination, seeds were subjected to a heat stress immediately after sowing. Seeds of the two independent 35Spro:MYB64 transformant lines and of a segregated wild type were sown on agar plates (2.3.2) and incubated at 22, 44, 46 or 48 ± 0.5 °C for 3 hours immediately after the vernalisation step and immediately before transfer to the growth room under Normal Conditions. Germination, measured as cotyledon expansion, was recorded 1 week later (Figure 5.2).

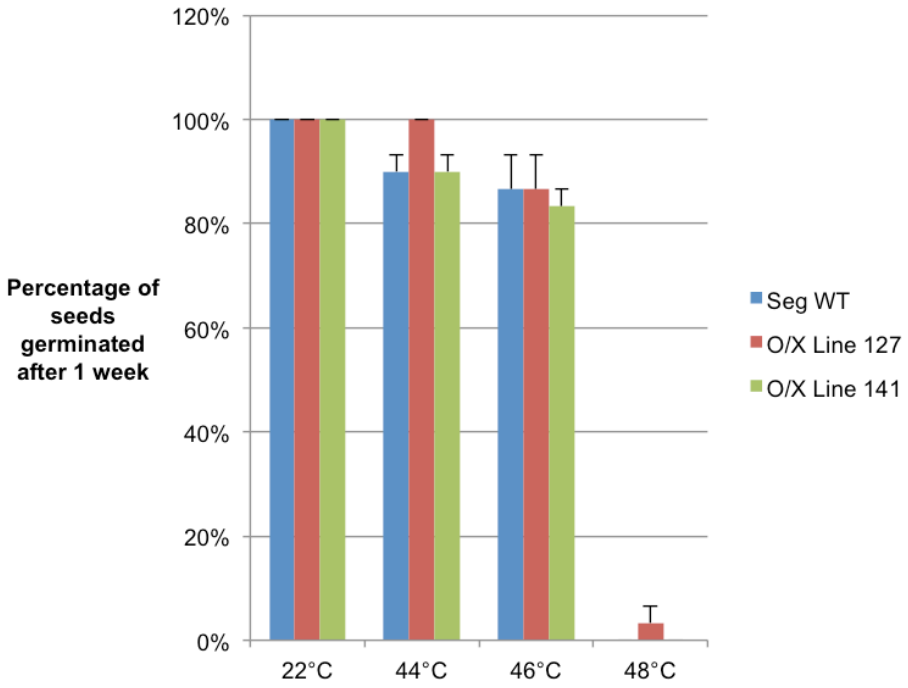


Figure 5.2 Germination of 35Spro:MYB64 Seeds Following Heat Stress

To investigate whether overexpression of MYB64 might confer a thermoprotective effect at the germination stage, seeds of two independent transformant lines were sown on tripartite agar plates, in duplicate, and incubated at 22, 44, 46, or 48 ± 0.5 °C for 3 hours immediately after the vernalisation step and immediately before transfer to the growth room under Normal Conditions. Germination, as scored by cotyledon expansion, was measured for each line one week later. Error bars represent means ± SE; *n* = 2 replicates with 15 plants each.

Each of the two 35Spro:MYB64 lines and the segregated wild type germinated normally at 22 °C and suffered a slight decline at 44 and 46 °C, but germination rates for all three lines was still greater than 80% and there were no significant differences (see Appendix ii–d). The ability of wild type seeds to germinate after a stress at 46 °C without the usual acclimation period provides evidence that the seeds have a higher basal thermotolerance than vegetative tissues in an established plant. All lines completely failed to germinate after a stress at 48 °C (with the exception of line 127: one outlying seed out of 30 was able to germinate).

5.2.3 Thermotolerance of 35Spro:MYB64: Root Development

Although publically-available transcript profiles do not strongly imply a role for *MYB64* in heat-stressed roots (section 3.3.2 - Data Mining of Open-Access Heat Acclimation Transcript Profiles) it was hypothesised that by constitutively overexpressing the gene we might confer a tolerance phenotype regardless of tissue type. Root development of the transgenic lines was thus measured with and without various heat stresses.

Fifteen seeds of segregated wild type were sown side-by-side with 15 seeds of 35Spro:MYB64 lines on bipartite agar plates (section 2.3.2) and germinated vertically under Normal Conditions. Roots grew down the surface of the agar and were measured after 5 days. Replicate plates were then subjected to one of six conditions: 22 ˚C (non-acclimated control); 22 ˚C -> 44 or 46 ± 0.5 ˚C (non-acclimated heat-stressed); 37 ˚C ± 0.5 °C (acclimation control); and 37 ± 0.5 ˚C -> 44 or 46 ± 0.5 ˚C (acclimated heat-stressed). Seedlings were also tested at a T_K of 46 ± 0.5 ˚C in this experiment for two reasons: basal thermotolerance of

wild type at this developmental stage had not yet been empirically determined; and a thermotolerant phenotype might not only lead to improved performance at damaging, yet sub-lethal, temperatures but might also lead to survival *above* the usual T_K. Plates were then returned to the growth room at Normal Conditions and allowed to recover and grow for 3 days before primary roots lengths were measured again. Differences between measurements made before and after heat stress for two independently transformed lines are presented in Figure 5.3.

Growth measured after incubation at 22 °C did not differ significantly from that measured after incubation at 37 ± 0.5 °C, demonstrating that acclimation itself has no deleterious effect on root extension, at least at this developmental stage. Measurements in Figure 5.3 are all normalised as appropriate against either of these controls. Non-acclimated, heat-stressed seedlings of all lines ceased all primary root growth immediately following the stress. Acclimated, heat-stressed seedlings of transgenic line 127 and the segregated wild type controls (Figure 5.3, **panel A**) also showed no further root growth. Acclimated, heat-stressed seedlings of both transgenic line 141 and the associated segregated wild type control (Figure 5.3, **panel B**) performed better on one of the replicate plates (approximately 60% of control root extension) and line 141 was not significantly different from segregated wild type. In replicate plate two, segregated wild type showed no root growth while line 127 showed a small degree of root extension (approximately 10% of control). This suggests the replicate plates reached slightly different temperatures within the oven and demonstrates the steepness of the response curve, and also suggests that replicate plate 2 perhaps represents the true differential phenotype revealed when the T_K is reached precisely.

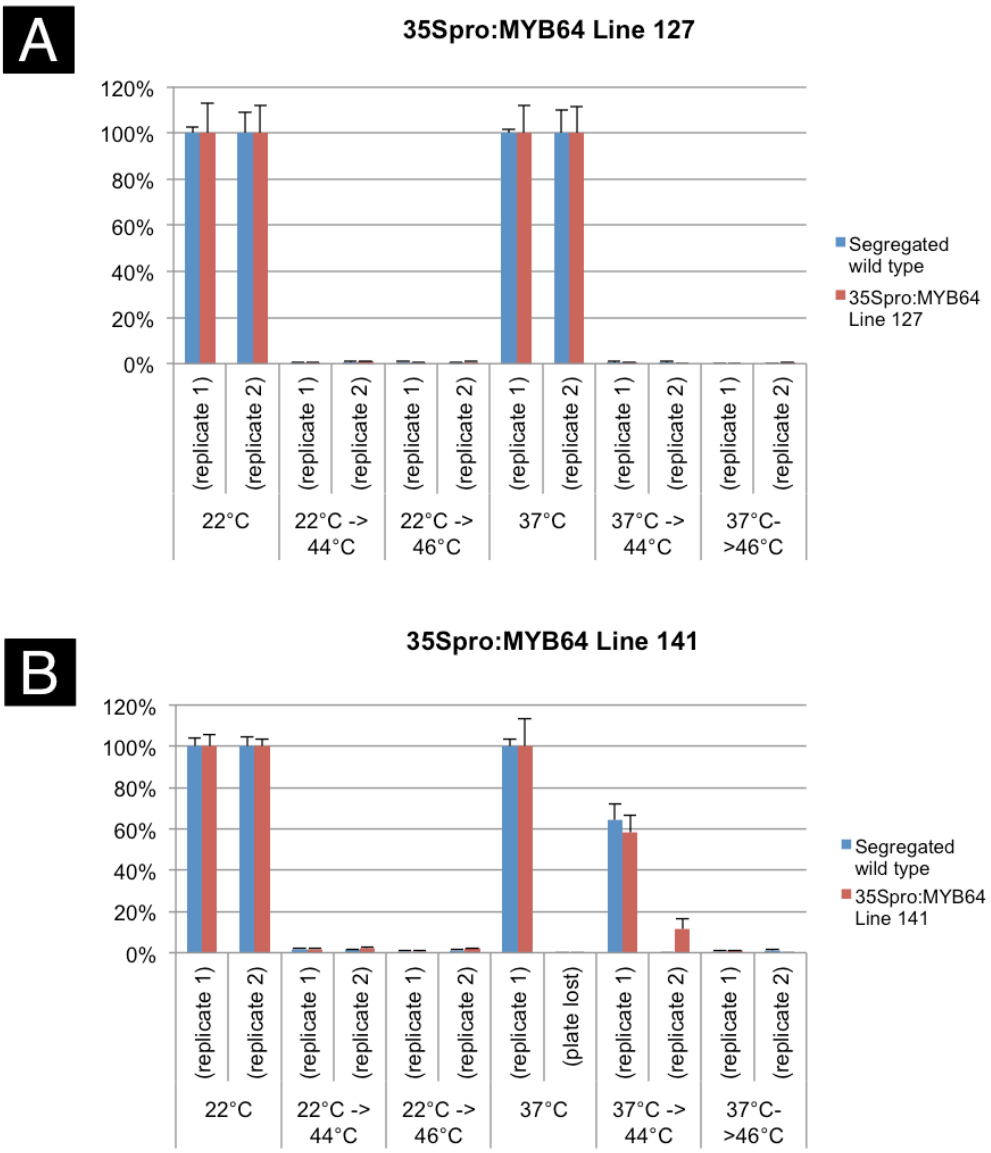


Figure 5.3 Effect of Heat Stress on Root Extension in 35Spro:MYB64 Lines

To investigate whether overexpression of *MYB64* might confer a thermoprotective effect shortly after the germination stage, wild type and transgenic lines were sown side-by-side on bipartite agar plates and grown vertically under Normal Conditions for 5 days. Root lengths were measured on day 5, then seedlings were subjected to a heat stress at 44 or 46 ± 0.5 °C either directly (left halves of graphs) or with an acclimation and recovery period prior to the stress (right halves of graphs). In all cases, root lengths were measured again 3 days later. The amount of root growth recorded after heat treatments is expressed here as a percentage of the growth recorded in the roots of the corresponding controls (i.e. non-acclimated (left halves of graphs) or acclimated (right halves of graphs)). **A** 35Spro:MYB64 Line 127 **B** 35Spro:MYB64 Line 141. Error bars represent means ± SE; *n* = 15.

These results indicate that overexpression of *MYB64* confers only a weak thermotolerant phenotype on the roots of Arabidopsis seedlings.

5.2.4 Thermotolerance of 35Spro:MYB64:Hypocotyl Extension

The next developmental stage investigated was the period shortly after germination. An experiment was performed to measure the length of dark-grown hypocotyls before and after a heat stress (the rationale behind employing commonly-used measurement is explained in 3.5.5).

Fifteen seeds each of segregated wild type and of the independent transformant lines 141 and 127 were sown on tripartite agar plates (section 2.3.2) and germinated vertically. Plates were incubated in the dark and exposed to stress temperatures at the appropriate times as also described previously in section 3.5.5. A further 2 days after stress the plates were unwrapped again and the post-stress hypocotyl extension was measured (Figure 5.4).

Again, an acclimation step alone had very little impact on growth; each line exhibited different hypocotyl extension rates but these were not significantly different between 22 °C controls and 37 ± 0.5 °C controls. As expected, acclimation conferred acquired thermotolerance upon segregated wild type and line 141 at 44 °C, but this acclimation was not evident in seedlings of line 127 which performed differently from the other two (statistically significant at $p<0.05$; see Appendix ii-e).

Neither line showed extension rates greater than segregated wild type. At 46 °C, non-acclimated wild type outperformed lines 127 and 141. These results indicate that overexpression of *MYB64* reduces growth rate of heat-stressed seedlings.

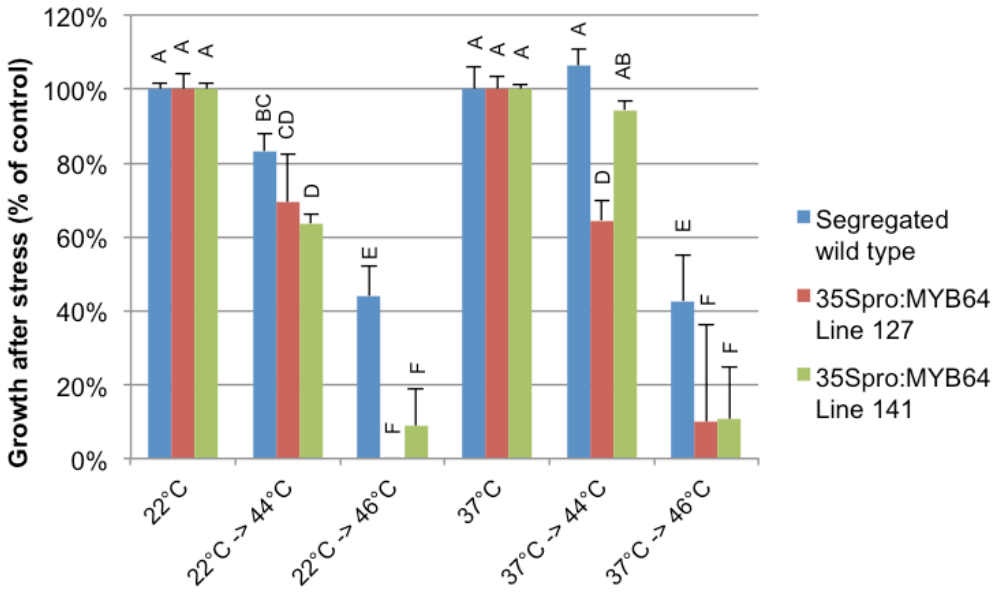


Figure 5.4 Effect of Heat Stress on Hypocotyl Extension in 35Spro:MYB64 Lines

Seeds of segregated wild type and of both transgenic lines were sown on tripartite agar plates and grown vertically at 22°C in the dark. Hypocotyl lengths were measured on day 3, then seedlings were subjected to a heat stress at 44 or 46 ± 0.5 °C either directly (left side) or after an acclimation and recovery period (right side). In all cases hypocotyl lengths were measured again 3 days later. Hypocotyl growth extension is expressed as a percentage of the corresponding controls (i.e. non-acclimated or acclimated). Error bars represent means ± SE; *n* = 15. Means that do not share a letter are significantly different; see Appendix ii-e.

5.3 Halotolerance of 35Spro:MYB64 Lines

The transcript profile of the original halotolerant mutant, HT5, suggested it might also be thermotolerant, but having thoroughly examined two independent overexpressing lines for a thermotolerance phenotype and having observed unconvincing results, a decision was made to return to the halotolerance phenotype. Experiments were designed to confirm *MYB64*'s role in the HT5 phenotype when constitutively overexpressed.

Fifteen seeds each of segregated wild type and of either line 127 or line 141 were sown on bipartite agar plates (section 2.3.2) and germinated vertically in the growth room under Normal Conditions. Plates contained normal growth medium with or without NaCl at a final concentration of 80 mM (the toxic concentration used to identify the original HT5 mutant from the pool of activation tagged seed (Price, 2005)). Plants were allowed to grow vertically for 2 weeks then seedlings were harvested and fresh weight measurements were recorded. The values presented in Figure 5.5 show the average fresh weight of each transgenic line expressed as a percentage of segregated wild type at the same salinity

Seedlings of both line 127 and line 141 achieved a higher average fresh weight on normal medium than those of segregated wild type (114 % and 119 %, respectively). They also achieved much higher fresh weights when grown on the saline medium (141 % and 176 % of segregated wild type). Tukey's Minimum Significant Difference test was performed; any line exhibiting a mean fresh weight greater than 159% of wild type when grown on 80 mM NaCl was

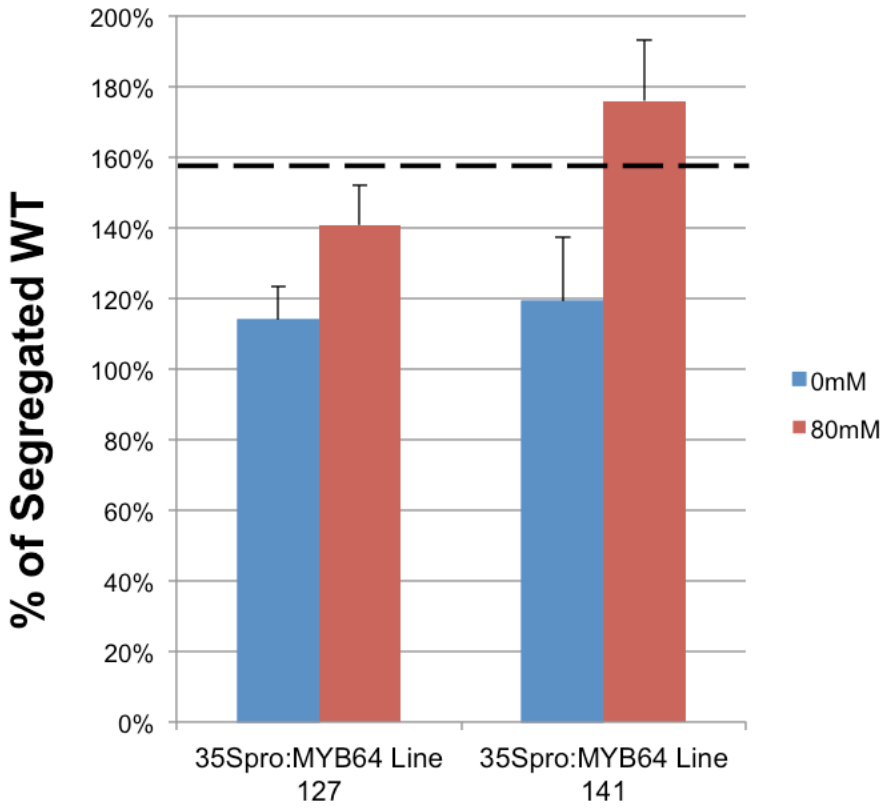


Figure 5.5 Effect of Salt Stress on Fresh Weight of 35Spro:MYB64 Lines

Seeds of 35Spro:MYB64 lines 127 and 141 were sown on bipartite plates alongside segregated wild type. Plates contained basal agar supplemented with 0mM NaCl (blue bars) or 80mM 80 NaCl (red bars) and were placed in the growth room vertically. Fresh weights were measured 2 weeks after germination. Ranked values were normalised on the ranked fresh weight of the segregated wild type seedlings at each NaCl concentration (*i.e.* wild type = 100%). Error bars represent the SE of n = 30. An Analysis of Variance test was performed and Tukey's Minimum Significant Difference ($p<0.05$) was calculated (broken line); treatment means above this line are significantly different from controls at $p<0.05$ (see Appendix ii-f).

significantly different (see Appendix ii-f). Thus, line 141 performed significantly better on salt, but line 127 did not.

Leaf number and shape were equivalent but leaves of segregated wild type were smaller (Figure 5.6). Roots of the wild type were longer with very little branching; primary roots of overexpressing lines were by comparison stunted in length, but had many more lateral roots. The salinity caused downward curling of the leaves of all lines as well as a twisted-root phenotype manifesting as a 45° tilt to the right as viewed in Figure 5.6.

5.4 ABA Response of 35Spro:MYB64 lines

Since many of the HT5 transcripts with the greatest increase in abundance had a functional annotation related to ABA responses, it was of interest to determine whether the expression of *MYB64* and of its 'downstream effectors' were influenced by the application of exogenous ABA. The hypothesis tested was that the lines overexpressing *MYB64* would be less sensitive than wild type as they would have had the chance to become acclimated to the already-high *MYB64* levels.

Thirty seedlings each of segregated wild type and of the two transgenic lines were sown on tripartite agar plates (section 2.3.2) supplemented with ABA to a final concentration of either 0, 0.5, or 1 μM. Except for the ABA, plants were grown under Normal Conditions for 14 days during which time seedlings were assessed as described below.

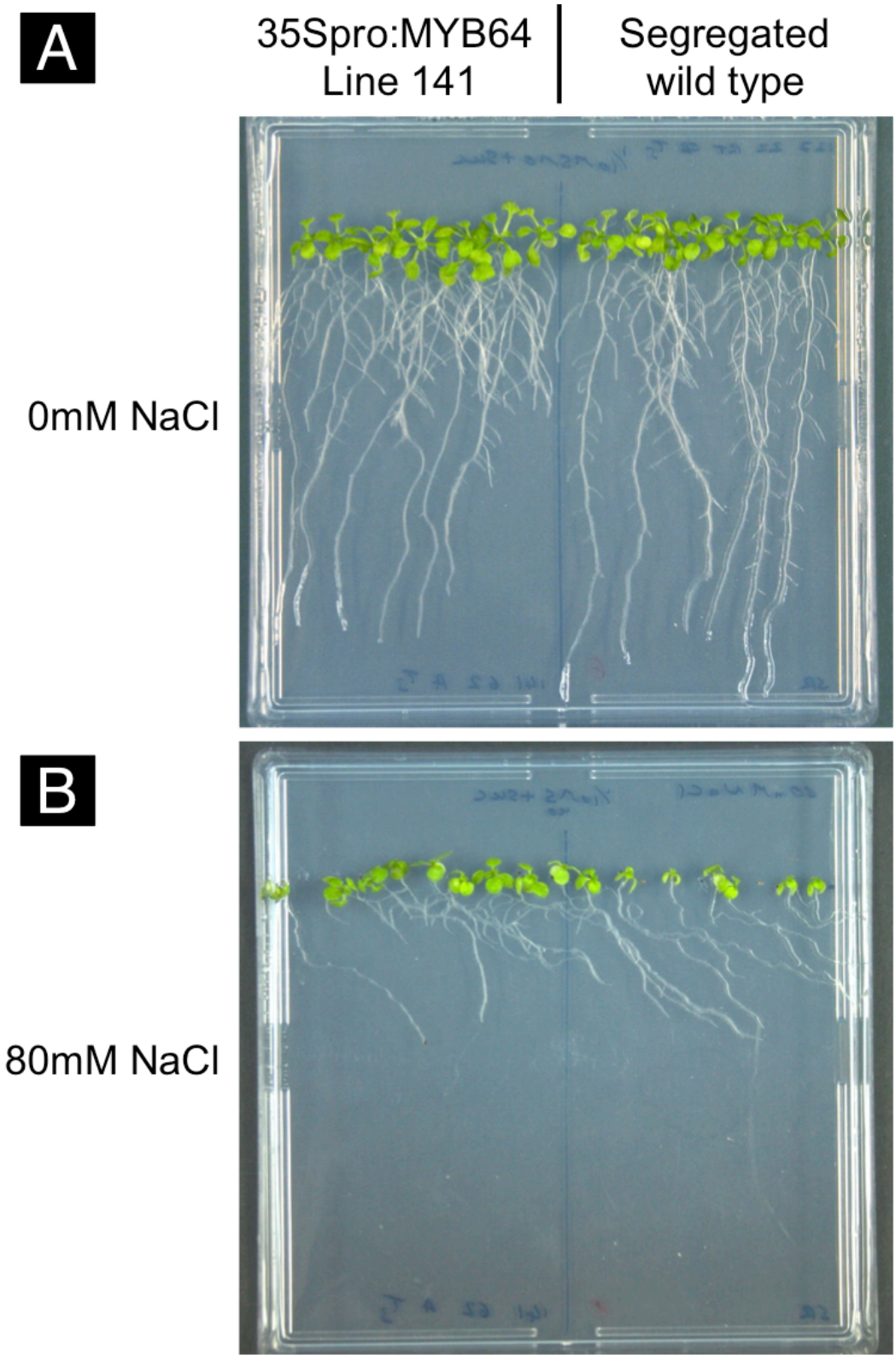


Figure 5.6 Effect of Salt Stress on Morphology of 35Spro:MYB64 Plants

Seeds of 35Spro:MYB64 line 141 were sown on bipartite plates alongside segregated wild type. Plates contained basal agar supplemented with either A 0mM NaCl or **B** 80mM NaCl. Plates were imaged 2 weeks after germination.

5.4.1 ABA Response of 35Spro:MYB64: Overall Morphology

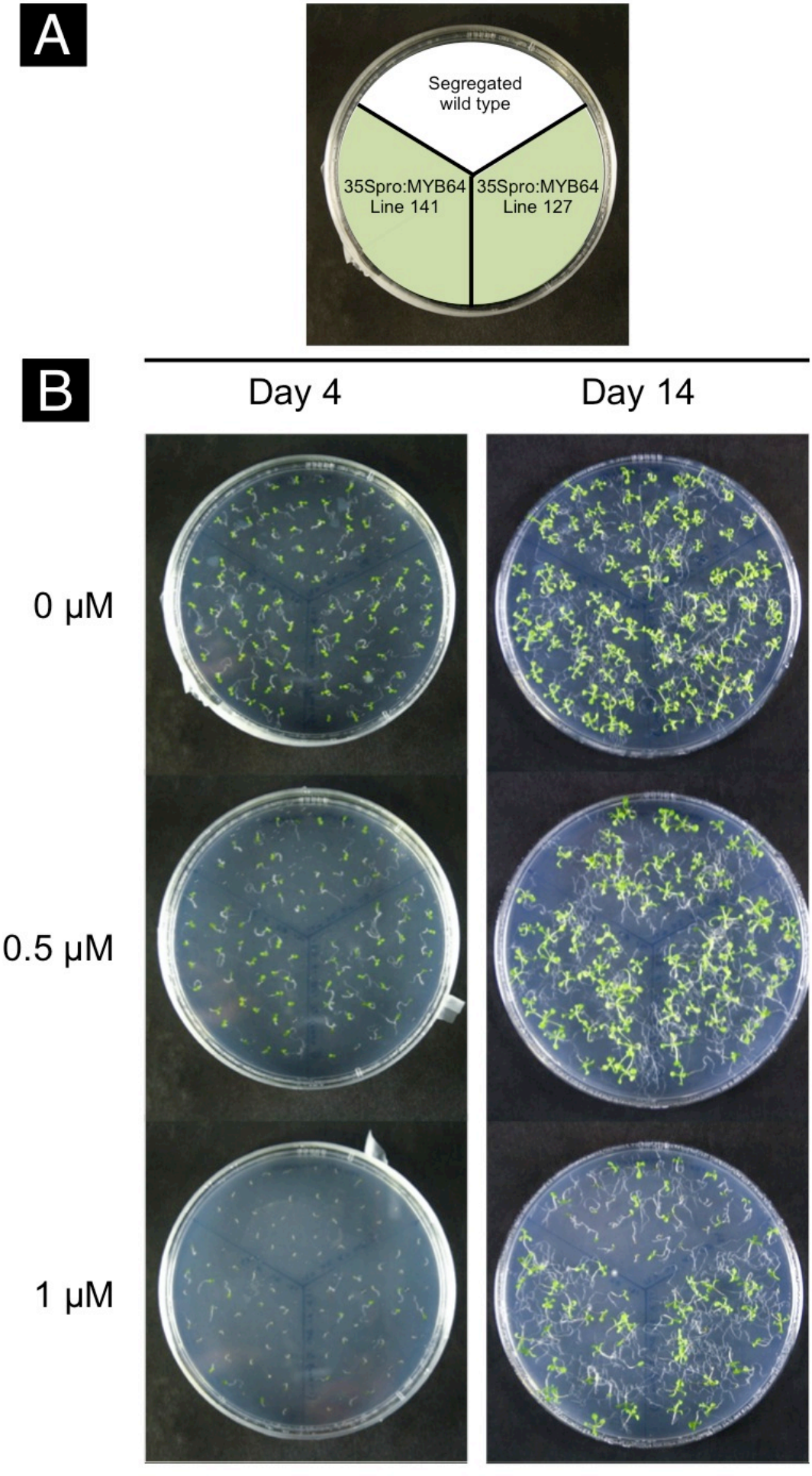
These plants were assessed qualitatively at day 4 and day 14 post germination and images are presented in Figure 5.7. At day 4 in the absence of ABA all three lines were performing equally well. In the presence of 0.5 μM ABA it was apparent that the segregated wild type developed more slowly and with smaller cotyledons than either of the two overexpressing lines. In the presence of 1 μM ABA there were almost no cotyledons apparent among the segregated wild type, while they had emerged from almost all seeds of lines 127 and 141. At day 14, in the presence of 0.5 μM ABA the root systems of line 127 and, to a lesser extent, line 141 appear denser than that of segregated wild type. In the presence of 1 μM ABA the two transgenic lines developed true leaves while the segregated wild type seedlings had either cotyledons only, or very small true leaves.

5.4.2 ABA Response of 35Spro:MYB64: Cotyledon Emergence

Cotyledon appearance among these plants was recorded daily from day 2 until day 14 (Figure 5.8). Appearance was defined as emergence from the seed coat and separation of the paired cotyledons as determined by viewing under a microscope, and scored as the percentage of seedlings that had reached this stage. In the absence of ABA both transgenic lines achieved 100 % cotyledon expansion within the first 3 days, while segregated wild type was not observed at this stage until day 5. This is consistent with the small growth advantage noted previously (Figure 5.5). In the presence of 0.5 and 1 μM ABA all three lines were slightly delayed in reaching maximum cotyledon expansion, though this was significantly slower in the segregated wild type ($p<0.001$; see Appendix ii-g). Under all conditions each line achieved a score of 80 % or greater within 5 days.

Figure 5.7 Effect of ABA on Development of 35Spro:MYB64 Lines

A Seeds of 35Spro:MYB64 lines 127 and 141 were sown alongside segregated wild type on tripartite plates containing basal agar supplemented with 0 µM, 0.5 µM, or 1 µM ABA and grown horizontally under Normal Conditions. **B** Photographs were taken at 4 and 14 days after germination.


A
Segregated wild type
35Spro:MYB64 Line 141
35Spro:MYB64 Line 127
B
Day 4
Day 14
0 µM
0.5 µM
1 µM

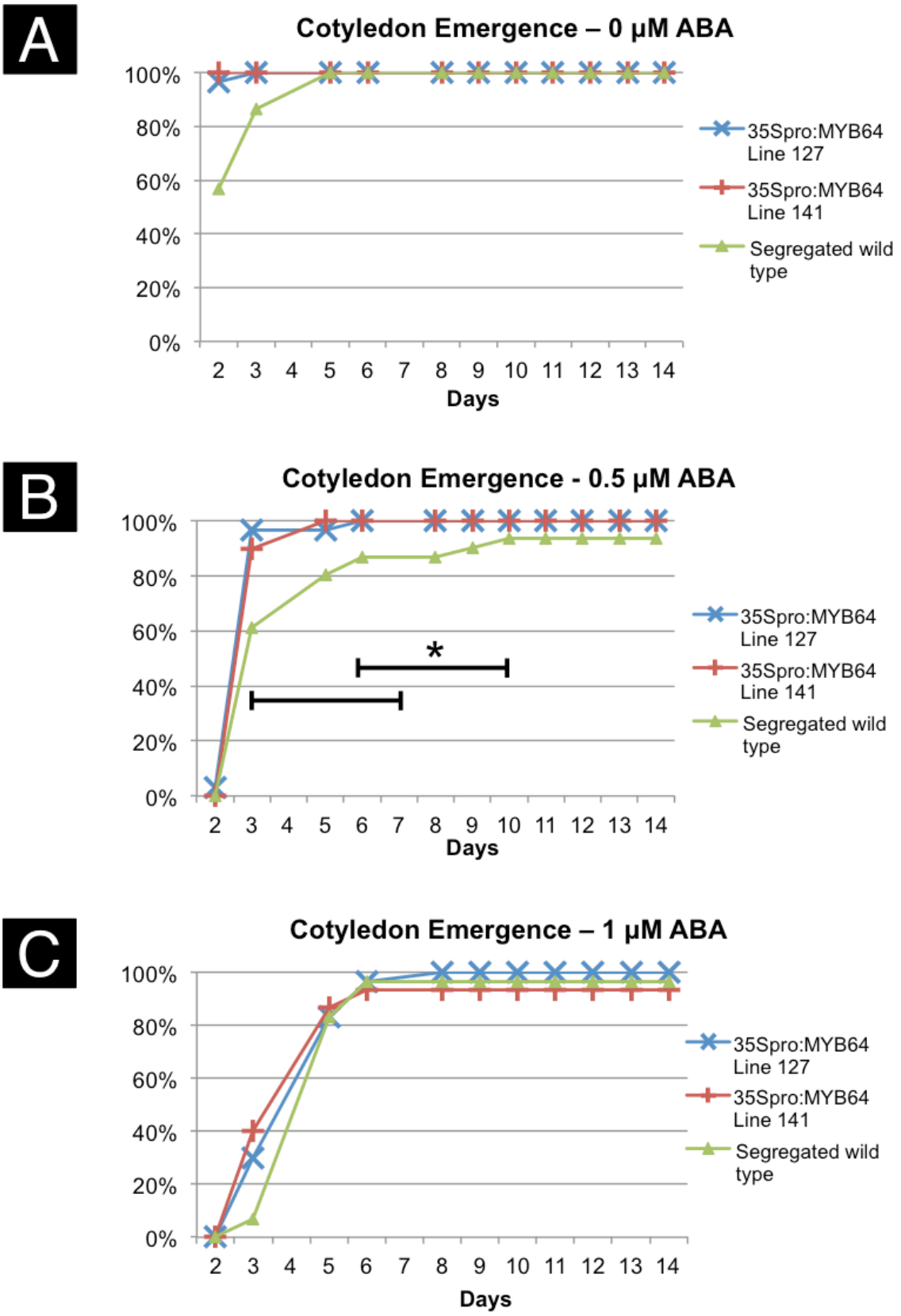


Figure 5.8 Effect of ABA on Cotyledon Expansion of 35Spro:MYB64 Lines

Seeds of 35Spro:MYB64 lines 127 and 141 were sown alongside segregated wild type on tripartite plates containing basal agar supplemented with either **A** 0 µM, **B** 0.5 µM, or **C** 1 µM ABA and grown horizontally under Normal Conditions. Seeds were counted every day for expansion of the cotyledons. At 0.5 µM ABA, segregated wild type cotyledon expansion was significantly slower than either of the transgenic lines ($p<0.001$; see **Error! Reference source not found.**) (pseudo-replicates generated by grouping data from days 3-7 and from days 6-10). Other time points were not significantly different. $n = 30$.

5.4.3 ABA Response of 35Spro:MYB64: Chlorophyll Production

The next observable stage of development was the production of green pigments in the cotyledons. This was measured over the same time period as above, determined by examination under a microscope and scored as the percentage of seedlings that had reached that stage (Figure 5.9). This simple method was chosen after consideration of more precise measurements, but the reliability of such assays has been disputed. Regardless of the technical limitations of methods of chlorophyll extraction and quantification, the aim here was simply to take a measure of the stage of development of each seedling and also to do this over a timecourse without destroying the plants.

In the absence of ABA, 100% of seedlings turned green within 5 days of germination in all three lines, though the segregated wild type did so with the slight delay already noted. In the presence of 0.5 μM ABA the two transgenic lines achieved their maximum within 5 days but segregated wild type did not reach its until day 10, though this was not significantly different (see Appendix ii-h). At 1 μM ABA the effect was such that segregated wild type still had not achieved a maximum score by the end of the two-week experiment. At every day of the experiment the two overexpressing lines had a significantly higher percentage of green seedlings than segregated wild type (also Appendix ii-h).

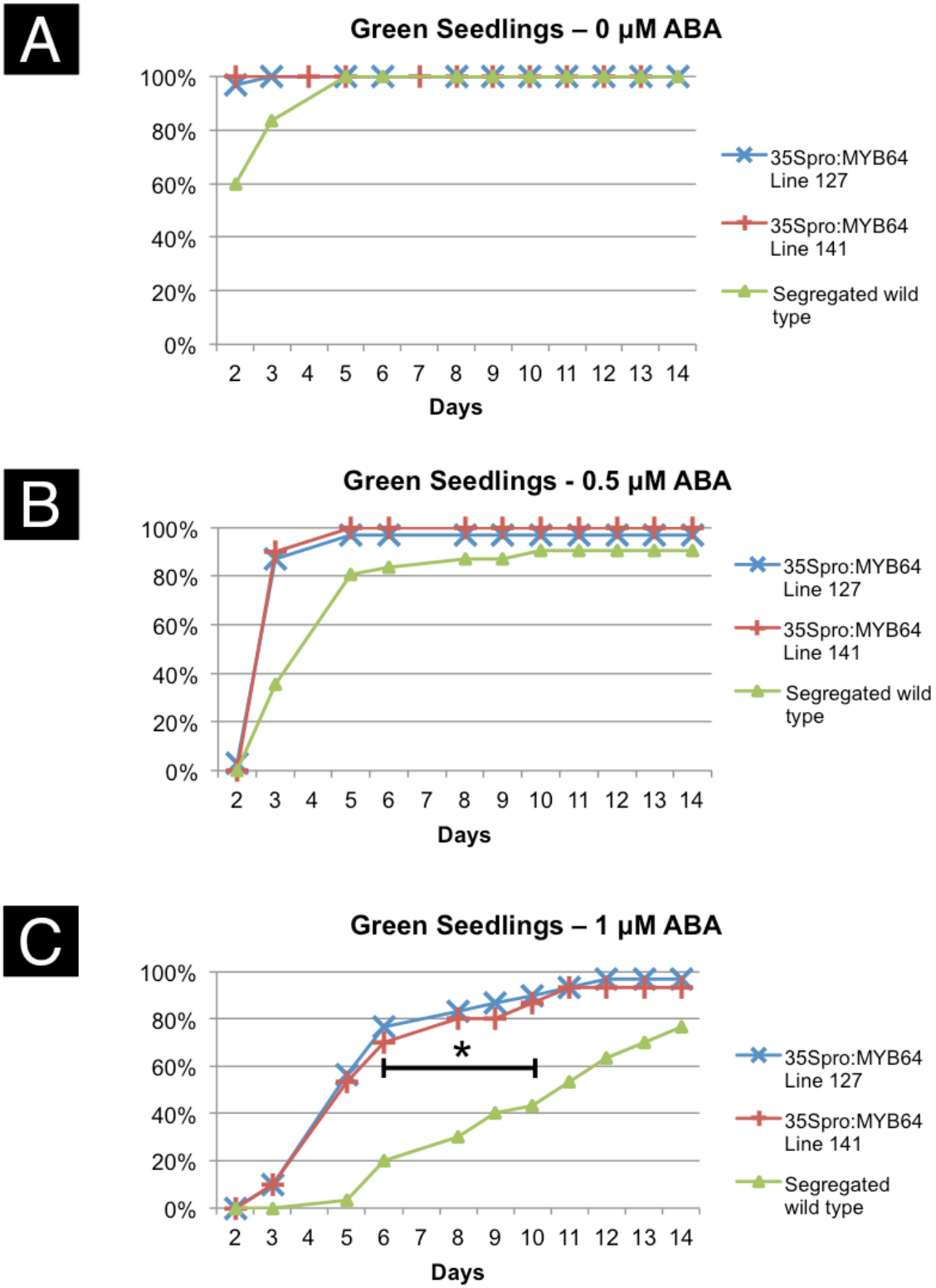


Figure 5.9 Effect of ABA on Visual Leaf Greening in 35Spro:MYB64 Lines

Seeds of 35Spro:MYB64 lines 127 and 141 were sown alongside segregated wild type on tripartite plates containing basal agar supplemented with either **A** 0 µM, **B** 0.5 µM, or **C** 1 µM ABA and grown horizontally under normal conditions. Seeds were counted every day for the appearance of greening (chlorophyll) in the cotyledons. At 1 µM ABA, chlorophyll production in segregated wild type was significantly slower than either of the transgenic lines ($p<0.001$; see **Error! Reference source not found.**) (pseudo-replicates generated by grouping data from days 6-10). Lines were not significantly different at other time points. $n = 30$.

5.4.4 ABA Response of 35Spro:MYB64: Fresh Weight

By day 14 there were no observable changes taking place on any sector of any plate (with the exception of seedling greening in wild type at 1 μM ABA) so it was assumed that the plants had reached maximum growth capacity on agar plates, and they were sacrificed in order to take fresh weight measurements (Figure 5.10).

As already noted in Figure 5.5 (Effect of Salt Stress on Fresh Weight of 35Spro:MYB64 Lines), under control conditions line 127 grew to a slightly higher fresh weight than segregated wild type (significant at $p<0.01$). In the presence of 0.5 μM ABA, fresh weight of segregated wild type was equivalent to that of its 0 μM control, but at 1 μM ABA it dropped significantly to 49 % of control ($p<0.01$; see Appendix ii-i). Transgenic line 127 suffered a slight inhibition of growth at 0.5 μM (81 % of control) and a further inhibition at 1 μM (73 %), but this was not as strong as the inhibition imposed on segregated wild type and was not a statistically significant reduction. Line 141 did not suffer any growth inhibition at any ABA concentration; in fact it achieved slightly (non-significant) *higher* fresh weights at both 0.5 μM (124 %) and 1 μM (119 %).

These results clearly demonstrate that overexpression of *MYB64* confers a resistance to ABA-mediated growth inhibition.

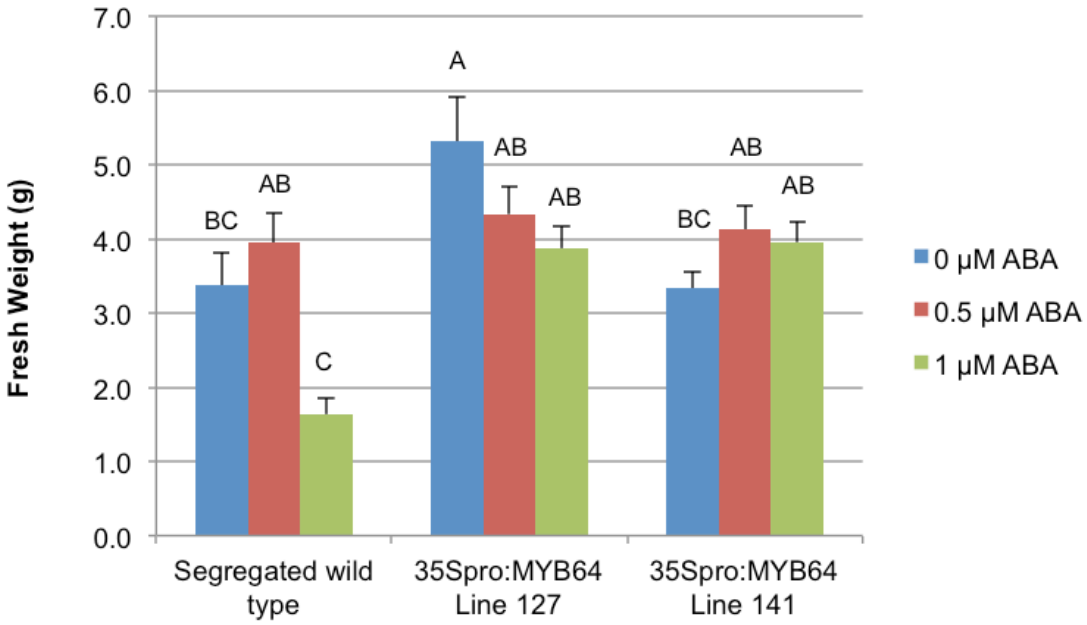


Figure 5.10 Effect of ABA on Fresh Weight of 35Spro:MYB64 Lines

Seeds of 35Spro:MYB64 lines 127 and 141 were sown alongside segregated wild type on tripartite plates containing basal agar supplemented with 0 µM, 0.5 µM, or 1 µM ABA and grown under Normal Conditions. Fresh weights were measured 2 weeks after germination. Means that do not share a letter are significantly different ($p<0.01$; see Appendix ii-i). Error bars represent the SE of $n = 30$.

5.5 Discussion

The aim of the experiments presented in this chapter was to establish whether the 35Spro:MYB64 lines were more thermotolerant than wild type and, following, that, how the lines responded to a variety of other abiotic stresses. The overwhelming answer to the first question is that there is only a weak thermotolerant phenotype observed under the temperature regimes described above.

Seedlings grown on agar plates and stressed at 2 weeks old showed no differences between lines. It may be that by this age the seedlings have reached a plateau in their growth rate and that they would not continue to grow any longer even in the absence of heat simply because of the constraints of growing in a plate. Were this experiment to be repeated, another suitable control would be a series of replicates kept in the growth room at Normal Conditions. When one observes the leaves that were able to retain their green pigmentation it becomes obvious that, across all lines, the greenest and therefore least damaged leaves are those in the centre of each rosette; the youngest. Leaf age thus appears to be negatively correlated with adaptability to a heat stress. Another reasonable alteration would therefore be to deliver the heat stress several days earlier so that all of the leaves were more metabolically active. Perhaps with this age-related greater capacity to regenerate, there might be a difference in the way the transgenics and the wild type recover.

Transgenic seeds given a heat stress at germination were also no more or less able than wild type to grow. This may appear to be an unusual experiment in the context of gene expression as the seeds are generally dormant, but since *MYB64*

was driven from a near-constitutively active promoter it would have been produced at all stages of development, including seed-setting. *MYB64* had been implicated in the expression, directly or indirectly, of smHSPs, which themselves are known to accumulate to high levels during seed dehydration (Kotak et al., 2007b). The hypothesis was that a higher level of both of these products might increase growth rate through a heat stress. This was shown not to be the case.

Remarkably, both transgenic lines showed a *lower* ability to extend hypocotyls after a stress than did wild type (Figure 5.4). The simplest explanation for this, is that the transgenic lines are suppressing growth in order to invest resources in adapting to the stress.

Thus, the hypothesis regarding the 35Spro:MYB64 line's enhanced thermotolerance was thoroughly investigated at various stages of development. The preliminary experiments carried out by Price (unpublished) and Nurcahyanti (2009) showed a stronger thermotolerance phenotype than demonstrated here, but despite exhaustive efforts they were not reproduced. This highlights the importance and difficulty of establishing and maintaining a reliable temperature regime suitable for investigation of such a steep biological response curve.

The next investigation concerned the responses of the two transgenic lines to other abiotic stresses. It was of interest to confirm whether the halotolerant phenotype observed in the activation-tagged mutant HT5, which expressed *MYB64* in a fundamentally different way (see sections 1.4.1 and 1.4.3), was reproducible in our engineered 35Spro:MYB64 lines. It was also hypothesised that since members of the MYB gene family has generally been characterised as ABA-dependent (Nakashima et al., 2009, Pieterse et al., 2009) there might be a

reduced response to the application of exogenous ABA as *MYB64* was already elevated.

Salt (NaCl) elicited a resistance phenotype in both transgenic lines, statistically significant in the case of line 141 (Figure 5.5 and Figure 5.6). This will be discussed in more detail in Chapter 6 ("General Discussion") after gene expression changes have been detailed in Chapter 4 (name here). ABA exposure also led to a clear and robust growth phenotype as measured by seedling size (Figure 5.7), developmental rate (Figure 5.8 and Figure 5.9) and by fresh weight (Figure 5.10).

5.5.1 Technical Considerations

5.5.1.1 Control of Temperature

Thermal gradients were the most difficult factor to control for in any of the experiments presented in this thesis. In order to investigate the heat-stress response of the plant, it was crucial to ensure that control was exerted over the temperature of the leaves (T_{LEAF}) rather than simply the temperature of the air (T_{AIR}).

T_{LEAF} is determined by the balance of thermal inputs and thermal losses. Heat is gained directly from the molecules in the atmosphere (T_{AIR}) and from incident light (Q). Heat is lost to moisture by transpiration, to the atmosphere by convection, and by infrared radiation. Each of these parameters had to be controlled in order to precisely control T_{LEAF}.

Incident light is perhaps less of a consideration in laboratory experiments than in the field as growth chamber illumination is often only a fraction of full sunlight. However, even at 1/3 sunlight strength, individual barley leaves have been shown to differ in temperature by as much as 8 °C from one end to the other as a result of partial shading from other leaves (Almalki, 2011, PhD thesis in preparation). While any such effect will not be as pronounced on the small leaves of very young Arabidopsis seedlings, it highlights the importance of precisely measuring T_{LEAF} as opposed to T_{AIR} in heat stress experiments generally. Incident light was eliminated as a confounding factor by incubating the plants in the dark.

Inputs to T_{LEAF} were therefore limited solely to the contribution made by T_{AIR}. This parameter was controlled by heating the plants in circulated-air cabinets. The cabinet controls allowed T_{AIR} to be specified to within 0.1 °C. After pilot experiments to determine the killing temperature produced inconsistent results, the accuracy of the thermal incubator was empirically determined by placing thermocouple probes at various positions around the interior. Even with a fan circulating the air, the upper and lower shelves were found to differ by up to 1 °C, so all heat treatments documented in this thesis are reported to be accurate to ± 0.5 °C.

While not normally in direct contact with the leaves, the growth substrate can also act as a thermal insulator. Parallel experiments carried out with thermal imaging cameras have shown that soil-grown plants placed in a circulated-air freezer cabinet can take up to 18 hours to reach the target temperature due to the specific heat capacity of the soil substrate (Wynne, 2012). (This has obvious implications for cold stress experiments that incorporate an overnight freezing

treatment.) It follows that the converse is also true: soil-grown plants placed in a high-temperature incubator will be subject to a heat-sink effect in the roots until the soil temperature has equilibrated with that of the cabinet. This was one factor contributing to the decision to focus on seedlings grown on plates.

Convective heat loss from these leaves was prevented by placing the leaves in a small, sealed environment; no air movement from the fans penetrated the plates to carry heat away, and the plants had no thermal input other than the exterior air temperature so there was no possibility of a convection current developing to cool the leaves.

Radiative losses were not preventable, but the thermostatically controlled incubator ensured that the air circulating around the plates replaced any heat lost in this way.

Transpiration is the major process by which plants are able to cool. This must be prevented in order to precisely control T_{LEAF} as opposed to T_{AIR}, and this is most easily done by ensuring there is no water vapour pressure deficit between the leaf and the atmosphere, *i.e.* achieving 100% relative humidity (RH). Pilot experiments were carried out on well-watered, soil-grown plants placed in containers constructed from sealed plastic bags. This meant that water could evaporate into the local atmosphere without being lost, thus reaching 100% RH. Efforts were made to ensure that the plastic did not come into contact with the leaves so as not to confound the results by prematurely allowing direct heat transfer from the air outside, but this had to be balanced against the need to have as little air inside the bag as possible so that 100% RH could be achieved in a period that was both short and relatively standardised across experiments. At 20 °C, a drop in RH from 100% to 99% corresponds to a water potential

difference of -1.36 megapascals (MPa) (Nobel, 1999, Table 2-1) which has a desiccating effect on leaves, and this impediment to stopping transpiration is compounded by the fact that RH drops dramatically as temperature increases (Nobel, 1999, Figure 2-13). This proved too difficult to control in a satisfactory and reproducible way, and this was the other major factor in the decision to avoid using soil-grown plants.

Seedlings on agar plates were therefore chosen as the primary tool with which to investigate thermotolerance. The temperature of the thin layer of atmosphere and the thin layer of agar substrate within a plate is able to equilibrate much more quickly than soil (within minutes). This means that a RH of 100% was achievable much more rapidly, thus bringing transpiration, the major determinant of leaf temperature in this experimental setup, to a halt within the first few minutes of incubation.

Even using plates, it became evident that seed density and plate position had a large effect on the reproducibility of results. When seeds were sown closely together the developing seedlings appeared to gain a measure of protection from incubation at T_K. For this reason protocols were adapted so that seedlings were spaced apart, ideally such that their leaves would not grow to touch each other. It also appeared that some areas of some plates were protected from the highest temperatures. These were deemed ‘edge effects’ and the cause has yet to be completely resolved, though it may be that proximity to the baffled fan combined with turbulence created by the other plates on the incubator shelves contributed to localised temperature gradients. Efforts were made to investigate this using a thermal imaging camera, but the glass interior door of the incubator was opaque to infrared light so this was not possible without opening the door

and allowing the warm air inside to mix with the cooler atmospheric air outside. The temperature of the plates inside dropped almost instantaneously with the fall in T_{AIR}, and the plastic of the plates was also opaque to infrared so the temperature of the plants (which would have remained stable for slightly longer) could not be directly measured this way either.

These temperature gradient observations may explain between-plate discrepancies revealed in some thermotolerance measurements presented in this thesis, demonstrating that it is absolutely crucial for the reliability of heat-stress results that the wild type control is sown on the same plate as the experimental line. This ensures the removal of as much variation in heat treatment as possible. The fact that all the seedlings in one plate on an incubator shelf can be killed at 44 °C while some of the seedlings in a replicate plate on the same shelf in the same incubator can survive also demonstrates the steepness of the curve of Arabidopsis survival to increasing temperatures. Nurcahyanti (2009) independently found that a temperature of 40 °C was not enough to kill wild type lines while 44 °C killed 93% of plants. These findings again reinforce the conclusion that control of T_{LEAF} rather than of T_{AIR} is crucial in heat stress experiments, particularly when the response curve is as steep as shown here.

Measurements of hypocotyl and root extension post-stress both revealed one surprising observation – that young seedlings of wild type or of either transgenic line were able to survive a stress at 44 °C, and to a much lesser extent 46 °C (Figure 5.3 and Figure 5.4). This leads to the conclusion that the T_K of wild type Arabidopsis should not be taken as an absolute value representative of all tissues at every stage of development.

6 General Discussion

6.1 Heat and Salt: Challenges to Growth, or to Survival?

Heat, drought and salinity are interrelated stresses that plants are unlikely to encounter in isolation outwith the controlled conditions of a laboratory. The typical consequence of each of them is that growth is arrested, and thus any experimental treatment that sustains growth through the stress period is typically regarded as a success. It is known, however, that many stress responses result in a state of relative metabolic stasis in order that resources can be invested in adaptive responses (Skirycz and Inze, 2010, Volaire et al., 2014). Such responses do not always require that the intensity of the response precisely matches the intensity of the stress in order to be effective; indeed it may be normal for certain stress responses to be over-activated at relatively low trigger levels (hypersensitivity), particularly if that stress would typically rise in intensity and persist for a long time. This early preparation would pre-empt chronic stresses such as an influx of salt into the local environment, which would not be removed on a timescale of hours in the same way as an acute stress induced by the diurnal heat cycle. One example of such a mechanism is seen in a knockout mutant carrying a T-DNA insertion in *APG7* (At5g45900, a positive regulator of senescence) identified by Price (unpublished). This mutant does not activate growth arrest following mild salt stress, whereas the wild type control is stunted. In perhaps a more realistic situation in the environment outside the lab, where levels of salt might eventually rise above a toxic threshold, this early growth-arrest response would clearly be advantageous for the plant, allowing

metabolism to focus on adaptation strategies rather than continued rapid growth.

Before interpreting and discussing the findings of this thesis it is therefore worth considering which strategy we are interested in investigating. If we are to bioengineer stress-tolerant crops to suit the changing climate models discussed in the introduction, we must first define what we mean by a 'successful' phenotype. The instinctive conclusion would be to regard larger plants as more successful, however it is important to briefly examine alternative stress responses exhibited by wild type plants and the mechanisms that underpin them.

One of three archetypal examples of altered growth rate in response to stress is that of the DELLA proteins, named for the protein domain that characterises the family. These have an inhibitory effect on growth in higher plants caused by an interaction between DELLAs and basic Helix-Loop-Helix (bHLH) transcription factors (Feng et al., 2008) and this inhibition is released when the plant produces the phytohormone gibberellic acid (GA). The GID1 (GA-Insensitive Dwarf 1) protein binds GA (Iuchi et al., 2007) and then targets DELLA proteins for polyubiquitination and subsequent degradation by the 26S proteasome (Dill et al., 2004). The DELLA-mediated inhibition of transcription is therefore lifted and the plant resumes growth.

The crucial part of this paradigm is the regulation of the production of GA. Only when conditions are favourable is GA synthesised; under stress conditions GA production decreases and growth is suppressed. As demonstrated by Zentella et al. (2007), failure to accumulate GA can be triggered by ABA, leading to a corresponding build-up of DELLA proteins. This has been shown to produce an observable root growth-restriction phenotype (Achard et al., 2006). Such growth

restrictions, whether caused either by exogenously applied ABA or by knockout mutation of the GA biosynthesis genes, have been correlated with improved tolerance of salt (Achard et al., 2006, Achard et al., 2008b, Magome et al., 2004), thus very clearly setting growth restriction in the context of a positive outcome for survival and adaptation.

Another example of a systematic shut-down of normal metabolic processes is the GCN (General Control Non-repressible) family of transcriptional / translational regulators, which also involves upregulation of stress-responsive pathways. *GCN2* (At3g59410) encodes a stress-activated kinase which phosphorylates eIF2α (Li et al., 2013), a component of the 40S translation initiation complex, altering the way the translation proceeds on a subset of the Arabidopsis transcriptome. Approximately two-thirds of all transcripts encode one main open reading frame (mORF) and the remaining third include one or more short upstream ORFS (uORFs) (Webb, 2008). Under normal conditions the 40S ribosome binds to the 7-methyl-cytosine cap of a mature messenger RNA molecule, scans along the sequence to the AUG codon, then the 60S ribosomal subunit is recruited and translation begins. At the end of the short uORF the ribosome disassociates and the short peptide produced may or may not go on to perform a functional role. Under stressed conditions, however, the 40S subunit, containing phosphorylated eIF2α, skips the uORF and translocates to the AUG of the mORF before the 60S subunit is recruited (Abwao, 2011). In this way the uORF transcripts are not expressed except in times of an appropriate stress. This uORF-skipping mechanism is not yet fully understood, but the effect is to halt translation of the normal transcriptome (mORF sequences) in preference of those sequences that encode stress-responsive proteins (uORF sequences).

A more physiological example of a mechanism for controlling growth is the expansin-mediated control of cell expansion. Plant growth is driven largely by cellular expansion in the zone just behind the meristems so the relatively inelastic cell wall must be able to adapt to this size increase. The expansins are a group of extracellular proteins located in the cell wall and their function is to disrupt the structural integrity of the cellulose microfibrils along the longitudinal axis (McQueenmason et al., 1992, Cosgrove, 2000). This allows the vacuole, and thus the entire cell, to expand under turgor pressure while constraining expansion along the axis of growth, rather than in all directions simultaneously. The cell wall is then reformed around the mature cell. The function of the expansins is regulated by acidity, which itself is determined by the action of ATPases in the plasma membrane, so this system is an example of a finely-tuned system of growth activation and suppression. While it has not been proven, it is possible that the function of the expansins could be halted upon sensing a stress that would negatively affect the water status of a plant, meaning that each cell would be constrained by its cell wall and thus the potential for further growth is effectively removed by an internally regulated system of adaptation.

Such a state would allow an organism to persist through a period of environmental stress that would be damaging to an actively growing cell. Perhaps, then, each phenotyping experiment carried out as part of this thesis should be characterised in the context of either a short-term response to an acute threat to *survival*, or a longer-term response to a chronic threat to continued *growth*.

6.1.1 Thermotolerance

Experiments on the thermal sensitivity of knockout lines of *HSP17.6* and *HSP17.6a* (AGI numbers At5g12020 and At5g12030, respectively) included post-heat-stress measurements of germination rate, root extension, hypocotyl extension, and seedling survival at 2 weeks old. Root and hypocotyl measurements relate to continual growth; a process which takes place over a period of days or weeks. Alternatively, germination and survival at two weeks were designed as acute interventions where the plants would either be killed or remain alive. The knockouts both showed weak or unconvincing thermosensitivity phenotypes under the two acute treatments scored by the binary dead-or-alive scheme (Figure 3.15 and Figure 3.16), although RT-PCR measurements of wild type lines showed that the transcription of these two genes was massively induced during thermal acclimation (Figure 3.5) suggesting that these sequences are indeed important in the response to high temperature. In the two experiments where tolerance was assessed by longer-term measures (hypocotyl and root extension over 3 days) the knockouts performed equivalent to, or slightly worse than, wild type. There was no significant difference in post-heat-stress root growth between either of the knockout lines and wild type (Figure 3.17), and the only significant difference in post-heat-stress hypocotyl extension was observed between wild type and *hsp17.6*, where wild type achieved 97 % of controls and the knockout achieved only 65 % (Figure 3.18).

These results suggest either that these two sequences are involved in medium- or long-term adaptation to high temperature as the effects of their loss took a few days to manifest, or, as the phenotypes were weak, they might be partially redundant (either with each other or with other members of the smHSP family).

At the moment this is difficult to assess as the two knockout lines studied here, *hsp17.6* and *hsp17.6a* (NASC stock numbers N507510 and N572448, respectively), each have a lesion in only one of the genes. The next stage in the investigation of these lines should incorporate a double-knockout, although the size of the smHSP family in Arabidopsis (21 members) and the mounting body of published evidence implies that the chance of finding a phenotype is relatively small (Dafny-Yelin et al., 2008, Sun and MacRae, 2005). Indeed, leading smHSP research groups have recently carried out similar experiments to those described here and also found weak phenotypes and thus pursued other lines of investigation (Elizabeth Vierling, University of Massachusetts Amherst, USA, personal communication).

Equivalent heat phenotyping experiments carried out on the 35Spro:MYB64 lines revealed a similarly weak thermotolerant phenotype when the plants were scored after exposure to two relatively acute stresses. Post-heat-stress survival after exposure to the killing temperature at two weeks old, scored by eye, revealed no clear survival advantage to overexpressing MYB64 (Figure 5.1). Post-heat-stress germination rates across all three lines were 90 – 100 % of control germination rates after heating to 44 °C, and 85 – 90% of control after heating to 46 °C, and there were no significant differences in either condition between wild type and overexpressing lines (Figure 5.2). When examined by the continuous assays of root growth and hypocotyl extension over the 3 days following a heat stress (Figure 5.3 and Figure 5.4) the wild type line grew equally well or better, but if *MYB64* overexpression leads to growth arrest as discussed previously, then this observed phenotype is exactly what would be expected as a beneficial response to stress. The next step in this line of investigation would be to examine if each line *recovers* from this period of notional stasis, and to

determine whether *MYB64* overexpression alters the duration of the stasis period, and the plant's vegetative and reproductive performance afterwards.

The NGS transcript profile data showed the elevation of some classic temperature-stress response regulators such as all three of the the DREB1/CBF family members (*CBF1* (17-fold), *CBF2* (5-fold), *CBF3* (7-fold)), and *DREB2A* (10-fold). It is known that overexpression of *CBF1* and *CBF3* confers freezing tolerance, and also that those plants perform much worse than wild type under normal growth conditions (Gilmour et al., 2000, Jaglo-Ottosen et al., 1998, Kasuga et al., 1999). Thus, it can be seen that constitutive expression of growth suppression effectors can be detrimental under permissive conditions but can actually improve survival in times of stress. It would be interesting, therefore, to compare the performance of the lines overexpressing *MYB64* with wild type over a protracted period of high (but not lethal) temperature stress.

It is noteworthy that the CBF genes are classically considered to be ABA-independent (Mizoi et al., 2012), yet this work reports evidence of an ABA-inducible transcription factor (MYB64) which, when overexpressed, induces transcription of all three CBF genes. The growth-suppression action of CBF1 is mediated by suppression of GA biosynthesis (Achard et al., 2008a), and thus MYB64 is potentially integrating the ABA and the GA / DELLA signalling pathways.

The Arabidopsis heat-stress literature to date has focussed on exploring survival at or above the killing temperature and efforts to improve thermotolerance have been geared towards increasing this threshold. Surprisingly little in the smHSP literature has been done to explore chronic responses or to devise a heat-stress regime to investigate the long-term detrimental effects on Arabidopsis. Here we

have shown two examples of sequences (smHSPs and *MYB64*) with no significant effect on immediate survival (even when given a pre-stress acclimation) but which may be contributing to the longer-term growth capacity of plants. Further work on these smHSP knockout lines and *MYB64* overexpressing lines should include investigations over a period of sustained sub-lethal high temperatures to determine the effects of sustained high temperatures across the duration of an entire growing season.

6.1.2 Halotolerance

Salt is an abiotic stress factor which presents a longer-term rather than an immediate survival problem, and the experiments on halotolerance reported in this thesis were designed to determine responses over a stress period of days and weeks. High levels of salt in the growth medium will lead to high levels of accumulated salt in the plant tissues. Various mechanisms operate to counteract this influx, such as sequestration (Zhu, 2001), efflux via antiporters such as SOS1 (Yang et al., 2009), apoptosis (Price, 2005), and exclusion at the exterior surface of the root (for more, see section 1.2.3) but the inevitable consequence of a sufficiently high concentration is that some salt will end up in the cytosol and subcellular compartments (Han et al., 2015). The effect of this is to increase the ionic strength of the cytosol, leading to altered secondary and tertiary protein structure. Proteins affected by this will be subject to the same threat as proteins in a heat-stressed cell; denaturation. The role of smHSPs and other chaperones in a salt-stress response thus becomes clearer.

It follows that reduced levels of smHSPs should negatively impact a cell's ability to cope with a saline environment. As argued above, however, continued growth

does not necessarily represent a more successful long-term survival strategy, suggesting one explanation for the improved growth observed in the knockout lines. HSP17.6 may be involved in suppressing growth, an adaptation that confers tolerance in the longer term. In the knockout lines, therefore, growth is sustained under stress conditions, at least in the short term. Since refolding of denatured proteins is an ATP-dependent process it would give an organism an evolutionary advantage to be able to reduce this energy cost by not producing new protein at a time when it was likely that it would be degraded. This would explain the large root-length reduction observed in the salt-stressed wild type, and the loss of smHSPs that are involved in such a signalling pathway would explain the much milder reduction seen in the knockouts (Figure 3.19). An alternative explanation, with a precedent discussed in the introduction to this thesis (Rhoads et al., 2005), is that expression of an HSP can lead to downregulation of other smHSPs in a negative feedback loop. The loss of a component of such a loop might lead to higher than normal levels of other members of the smHSP family, thus alleviating overall stress and accounting for the growth advantage observed in the knockout lines.

The phenotype exhibited by the salt-stressed 35Spro:MYB64 line, on the other hand, is a clear example of reduction in one measure of growth being complemented by a significant increase in another. The overexpressing line 141 exhibited stunted primary roots (Figure 5.5) while at the same time achieving a higher overall fresh weight (Figure 5.6). It would be interesting to repeat these experiments on the smHSP knockouts in order to supplement the root measurements, described in the paragraph above, with knockout fresh weight measurements and images of rosette sizes. The salt-stressed 35Spro:MYB64 line also had a greater number of lateral roots than segregated wild type. As

mentioned in the introduction to this thesis, salt sequestration in the vacuole is one mechanism for maintaining turgor pressure during times of salinity and dehydration stress. It is also known that lateral root production is stimulated on exposure to salt (Zolla et al., 2010) and that high Na^+ or low K^+ causes roots to 'corkscrew' and bend in what is perhaps a morphological adaptation that allows the plant to find patches of soil with a more favourable nutrient composition (Kellermeier et al., 2014). Root caps, being non-suberised and therefore more permeable to water, are the main root components where apoplastic nutrient uptake can occur (Sattelmach and Horst, 2007). These numerous short lateral roots effectively increase the amount of root cap area exposed to the salt-supplemented medium, and this would allow for greater ratio of apoplastic:symplastic transport of salt directly into the xylem from where it could reach the aerial tissues and be shuttled to the vacuole with a lower risk of causing intracellular toxicity.

With so many stress-response regulators upregulated in the transcript profile of this transgenic line it appears that the plants have become rather more comprehensively adapted to salt stress than if one simple end-effector protein had been overproduced, manifesting here as an enhanced rearrangement of root architecture. It would be of interest to examine in more detail the NGS transcript profile of this line when grown on high levels of salt in order to find out what other molecular changes might be contributing to the shortened root / higher fresh weight phenotype. Clearly, though, this line is able to adapt to the type of long-term stress that will become more prevalent with an increased demand for water in the field.

6.2 *MYB64* Overexpression

6.2.1 Interaction of MYB64 and smHSPs

The relationship between *MYB64* and smHSP expression was investigated by looking for promoter motifs in the 1 kb upstream of each TSS. No known motifs were enriched among these notional promoters of these three genes. If the MYB64 protein was binding directly to these regions it should be possible to find a common binding sequence; instead it is more likely that MYB64 binds to the promoters of other genes and activates these, which then directly or indirectly leads to upregulation of the smHSPs. This complexity was further discussed in section 4.3 (Transcript Profiling of 35Spro:MYB64 Line 141) where it is shown that a large and varied suite of stress-response transcripts was upregulated in the non-stressed 35Spro:MYB64 line, yet these smHSPs identified in the HT5 microarray were notably absent. It seems logical that downstream effector proteins, such as chaperones like the smHSPs, should be controlled by different promoter motifs from master regulators like kinases or transcription factors such as *CBF1* and *DREB2A* to eliminate the potential for inappropriate simultaneous regulation at these two different organisational levels of transcriptional cascades. A common MYB64 binding site might be found among the promoters of these high- or middle-level genes in the cascade instead. Perhaps, then, it is even possible that expression of each of these three smHSPs is activated by a different transcription factor.

Another explanation for the absence of any common promoter motifs is that the region assessed upstream of each TSS was too small. To expand the search would involve investigating sequences contained within, or distal to, other coding

regions. While distances of > 1kb between promoter binding site and TSS are not unknown, the increased distance would make concrete deductions difficult for two related reasons. First, the protein:protein interactions between a transcription factor bound at a distal motif and other transcription initiation factors bound at the TSS would depend on higher order DNA structures (loops, histones, *etc.*), theoretically allowing great latitude in the precise distance between the TSS and the MYB64 binding motif, and thus reducing the power of alignment algorithms. Second, conventional wisdom tells us that any putative motifs are less likely to be linked to the transcription of distant genes than to those in the local vicinity of the chromosome. Should any candidate motifs be found in a wider search they would need to be tested by functional *in vivo* studies (such as promoter mutagenesis for loss of function, transgenic promoter/reporter lines, chromatin immunoprecipitation), or *in vitro* studies (such as DNase footprinting) to establish any connection to expression of these smHSPs.

The observation that there are no promoter motifs evident in the 1 kb upstream of the TSS of each gene appears to be confirmed by the failure of the own-promoter-driven fluorescently-tagged constructs to express in *N. benthamiana*. Had there been any functional promoter motifs we might reasonably expect that, since the gene family is so highly conserved across species, *N. benthamiana* would have suitable factors to recognise and transcribe from them. The post-transcriptional modification and translation of these mRNAs evidently take place, as we can see expression of the fluorescently-tagged smHSPs driven by the constitutive 35S promoter (Figure 3.9, Figure 3.10, and Figure 3.11).

The discrepancy between the statistically significant halotolerant and phenotype and the weak or absent thermotolerant phenotypes exhibited by 35Spro:MYB64 line 141 also suggests that simply overexpressing *MYB64* may not be enough to upregulate smHSP transcripts, and that an appropriate environmental stimulus (*i.e.* other than heat) might also be required. In the only case where both *MYB64* and the smHSPs were shown to be upregulated (the HT5 microarray) the plants had been grown for 2 weeks on 80 mM NaCl. This evidence supports the hypothesis that a salt stimulus activates a signal which interacts cumulatively with the signals controlled by *MYB64*, resulting in the coordinated upregulation of the smHSPs.

6.2.2 Transcript Profiles of 35Spro:MYB64 Lines

Next-generation sequencing is still a relatively new method of interrogating the transcriptome activity and the insights it provides are significantly richer than anything available before, and thus the conclusions derived from the Illumina NGS experiment were judged to be more robust than those derived from the HT5 microarray. That microarray data, on which the *MYB64* investigation has been founded, reported fold-changes of less than one order of magnitude and only two transcripts were elevated by more than 4-fold (Table 1.1), whereas the greatest increase reported on the Illumina platform was 466-fold. While this is qualitatively true it should be noted that this is based on a ratio against a very low detection level in wild type and the margin of error for this calculation is therefore high. However, even when only considering sequences with ratios based on more robust detection values, *e.g.* >3 reads in the wild type sample and an RPKM of >1.0, the first transcripts that meet these criteria are still upregulated approximately 20-fold. For example, At3g53980 (encoding a lipid

transport protein with an uncharacterised role) is upregulated 27-fold, and the next (At1g28370 – encoding *ERF11*) is upregulated 20-fold. This may reflect a fundamental difference between technologies that measure analog signals (such as microarrays) and those that rely on digital signals, such as this NGS list of reads. Analog systems are subject to the limitations of signal saturation, minimum detection thresholds, and the nature of the dynamic range of values, and with microarrays there is also the possibility of inappropriate hybridisation. The only one of these limitations that affects a NGS transcriptome profiling is the minimum detection threshold. If it were necessary this could be avoided by increasing the input concentration of mRNA up to the limit of physical saturation for space on the surface of the flow cell and in addition more than one flow cell could be used to increase overall sequencing depth. The fact that Illumina sequencing is not subject to many of these limitations suggests it provides a more reliable estimate of these transcripts than the HT5 microarray and we propose that differences between results of the two types of experiment are to be expected.

It should be noted, however, that the HT5 microarray was performed on seedlings that had been germinated and grown on medium supplemented with NaCl while these NGS data came from transgenic tissues that were exposed to no stress. There is evidence that long-term exposure to a stress can cause a pattern of initial high-amplitude gene expression changes followed by a long decline as the organism adjusts (Young et al., 2013), so the relatively low level of gene expression changes in that long-term stress (3-4 week) experiment is perhaps not surprising. What is remarkable, though, is that the plants used to generate the Illumina transcript profile were grown for the same length of time on standard soil under normal environmental conditions and subject to no observable stress,

yet they still exhibited fold-increases in stress-response transcripts of one or two orders of magnitude. It can be speculated that the high level of *MYB64* expression simply drives consistently high levels of downstream gene expression, or it can be speculated that something about the analysis platform itself, which gives a fundamentally different measure of gene expression, exposes these larger changes in transcript levels. Since both the treatment and the measurement tool are different in these two experiments it is impossible to say definitively which is the cause, but it would be interesting to perform a microarray experiment in parallel with a next-generation sequencing analysis on the same tissue to see how closely the results matched.

One explanation for the variation between the HT5 microarray results and the NGS profile of the 35Spro:MYB64 line is perhaps the difference in the nature of the two lines. The HT5 insertional mutant carried the 35S enhancer elements more than 1kb upstream of the TSS of *MYB64*. The homozygous transgenic lines carried two copies of the wild type allele plus two extra copies of a construct with the 35S enhancers fused directly to the 5' end of the gene. Dosage effects therefore probably contribute to a large degree of the variability between the two transcript profiles. Nevertheless, similar qPCR findings regarding the stressed transcript profiles of independent transformant lines 127 and line 141, given appropriate stress conditions (Figure 4.4 and Figure 4.5), shows that these downstream genes are indeed modulated by MYB64.

Aside from their magnitude, the range of the stress responses represented in Table 4.1 was the most striking result. The annotation attached to At4g23680, the most highly 'upregulated' transcript, indicates that it encodes an uncharacterised, unnamed protein that is upregulated by biotic stimuli and

defence responses. This crossover between an initial abiotic stress phenotype and biotic response transcripts could be regarded as the major finding of these transcript profiling results. Of course, it is not unusual for a sequence to have a role in more than one process, but much has been made in the literature of categorising stimuli as biotic or abiotic and classifying the responsive genes accordingly (see section 1.2, "Crop Responses to Climate Change-Related Stresses: Prospects for Adaptation"). The systematic gene ontology analysis presented in Table 4.2 underscores the message that perhaps this paradigm should be reconsidered.

In reviews of abiotic stress-responsive plant hormone signalling networks, Nakashima et al. (2009) and Todaka et al. (2012) describe networks of ABA-dependent and ABA-independent abiotic stress responses in Arabidopsis and homologs in rice (*Oryza sativa*). The ABA-dependent pathway is dominated by transcription factors that bind and activate gene expression at ABA-Reponsive Elements (ABREs), and these TFS are collectively known as the AREB/ABF (ABA-Reponsive Element Binding / ABRE Binding Factor) family. These are proposed to be mainly activated by ABA in response to dehydration stress. The authors then describe an ABA-independent family, the Dehydration Responsive Element Binding (DREB) transcription factors, inducible by dehydration and cold stress. In two other recent reviews Pieterse et al. (2009, 2012) describe the current understanding of the networks of plant responses to a range of biotic stresses which have historically been centred around the jasmonic acid (JA), ethylene (ET) and salicylic acid (SA) biosynthesis pathways. Members of the R2R3 subfamily of MYB transcription factors have very recently been found to play an antagonistic role in SA responses while enhancing JA responses (Ambawat et al., 2013). Interestingly, expression of MYB96, another ABA-responsive R2R3 MYB

transcription factor, contributed to pathogen resistance by *increasing* the levels of salicylic acid (SA) biosynthesis (Seo and Park, 2010). This represents another example of a MYB transcription factor serving as a node between what are typically thought of as independent response pathways mediated by discrete signals.

These developments, published concurrently with the research described in this thesis, are beginning to provide a new understanding of the role of MYBs in connecting abiotic and biotic stress responses as well as ABA-dependent/independent pathways (Figure 6.1). Some MYBs have been shown to be ABA-inducible (e.g. MYB2, Abe et al., 2003) which is consistent with the findings regarding MYB64 (presented in Figure 4.7). Perhaps controversially, however, the list of the 50+1 most upregulated transcripts in the 35Spro:MYB64 line (Table 4.1) include DREB2A, a classic member of the ABA-independent response pathway. The HT5 mutant was isolated from an abiotic stress tolerance screen, though the transcript profiling again shows that 10% of the sequences listed in Table 4.1 are ethylene response factors. In total the list contains eight transcripts known to be biotic stress-responsive, five known to be abiotic stress-responsive, and four known to respond to both. When the genes of unknown function are removed these numbers equate to 25 % biotic stress, 15 % abiotic stress, and 12 % both. The other 48 % may nevertheless be involved in stress responses too. The gene ontology analysis performed on the 100 most highly upregulated transcripts Table 4.2 shows there is a statistically significant enrichment for genes encoding calcium-binding proteins. For example, At3g01830 (calmodulin-related protein; unknown role; ranked 3rd; 47-fold upregulated) and At4g29000 (EF-hand family protein; unknown role; ranked 29th; 16-fold upregulation) might play a role in calcium signalling, perhaps connected

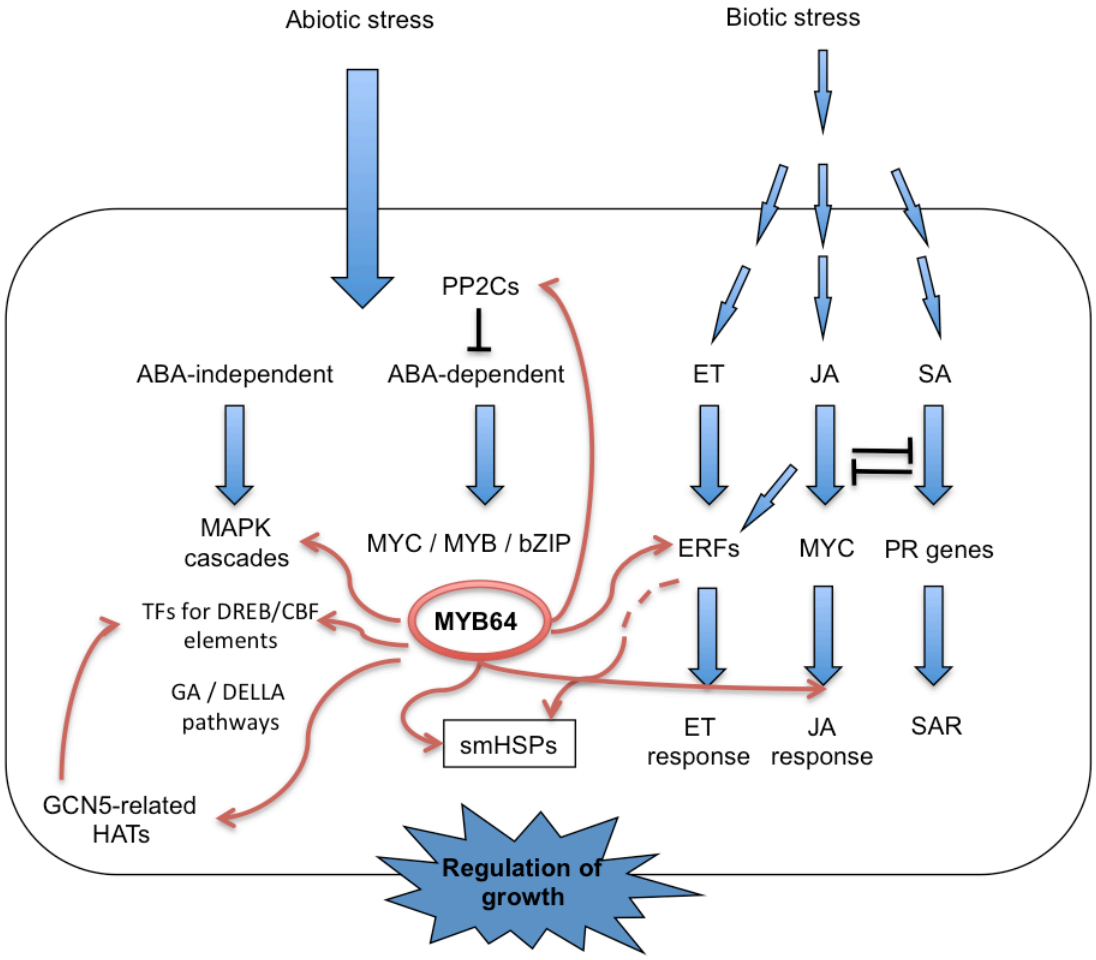


Figure 6.1 MYB64 as a Cross-Talk Node Between ABA Dependent / Independent Pathways and Abiotic / Biotic Stress Pathways

Historical literature has characterised pathways on the basis of the signals which activate them, leading to responses defined by high-level messengers such as ABA and other hormones. *MYB64* overexpression leads to upregulation (red arrows) of various components of a surprisingly diverse range of pathways. See main text for discussion.

to SOS3 activation. At3g21780 (ABA glycosyl-transferase; unknown role; ranked joint 34th; 15-fold upregulated) is likely to be involved in ABA metabolism in response to stress. At2g43030 (acetyl transferase, GCN5-related; unknown role; ranked 43rd; 13-fold upregulated) may be involved in suppression of growth by modulation of transcription. GCN5, a component of the GCN pathway mentioned previously (section 6.1.1), is a histone acetyl transferase that physically interacts with CBF1 and other transcriptional activator proteins (Mao et al., 2006), so it can be seen that At2g43030 is likely to act to change transcription patterns by the modification of chromatin.

Taken together, all of these results suggest that overexpression of MYB64 is contributing in a considerable way to the stress-readiness of the plant and there is considerable latitude for further investigation.

6.3 Independent Ac-Tag Mutants

The activation-tagged mutants identified from the thermotolerance screen on the Weigel lines (Alonso et al., 2003) were shown to have a very similar transcript profile to the 35Spro:MYB64 lines. This provides further evidence that the *MYB64* regulon is important in the thermotolerance response of Arabidopsis.

6.4 ABA and *MYB64*

Growth of both 35Spro:MYB64 lines was less inhibited than segregated wild type by exogenous ABA (Figure 5.7 to Figure 5.10). It has been shown that ABA inhibits growth by the activation of appropriate transcription factors (Golldack et al., 2013, Sreenivasulu et al., 2012). It was unexpected that overexpression of one such downstream transcription factor resulted in the opposite effect. One

explanation might be that these transgenic lines are constitutively mimicking exposure to ABA-related stress and may be acclimated or habituated, leading to a dampened detection when it is delivered exogenously. Such a mechanism might involve *MYB64*-directed breakdown of ABA as part of a normal negative feedback loop, which is enhanced in the transgenic lines leading to insensitivity.

Another explanation is that the upregulation of *PP2C* (At5g59220) might lead to the suppression of ABA-dependent responses in this situation. Recent developments in understanding of ABA detection and signal transduction show a family of proteins known as the PYR/PYL/RCAR (Pyrabactin Resitance/PYR1-Like/ Regulatory Components of ABA Receptors) group to play a role in modulating the activity of SnRK2s (Sucrose non-fermenting 1-Related protein Kinases); a group of kinases which phosphorylate and thus activate stress response TFs (Fujita et al., 2009). Under normal conditions PYR/PYL/RCAR proteins dimerise in an inactive form (Figure 6.2). The SnRK2s are held in an inactive, dephosphorylated state by PP2Cs. In the presence of ABA the PYR/PYL/RCAR proteins interact with ABA and adopt a conformation that enables them to bind into the pocket of the PP2Cs previously occupied by the SnRK2s . The SnRK2s are then free to autophosphorylate and then phosphorylate target TFs, which include AREBs (ABA-Responsive Element Binding proteins), WRKYs, and MYBs (Kim, 2014). This theory was recently supported by the results of Antoni et al. (2011) who showed that this PP2C is a negative regulator of ABA signalling. This further represents an example of cross-talk between biotic and abiotic stress responses.

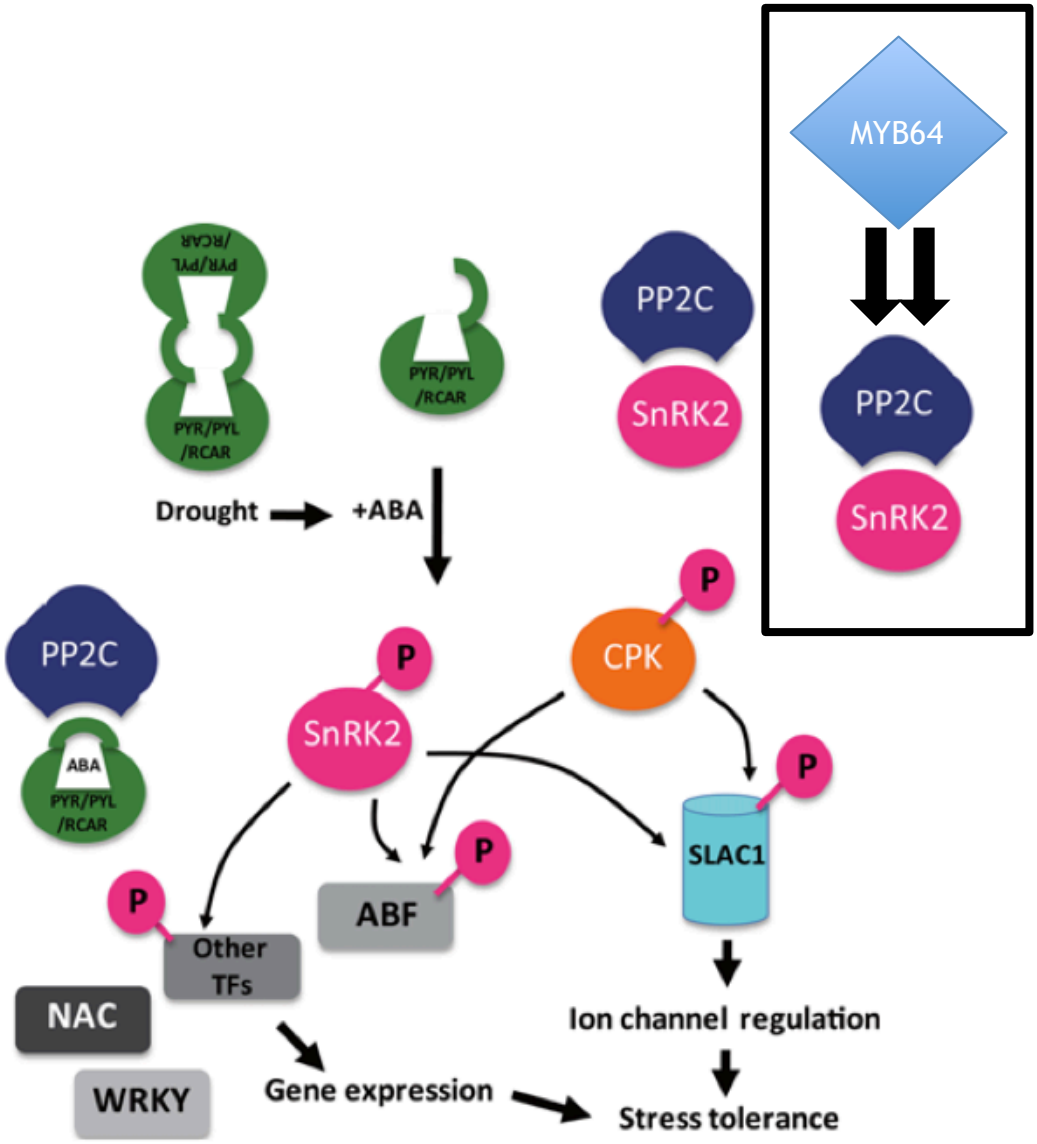


Figure 6.2 ABA Sensing in Response to Drought-related Stresses

The current model of ABA sensing includes a family of proteins known as the PYR/PYL/RCAR group, which bind ABA and inactivate the PP2C family. This leaves the snRK2 group of protein kinases free to auto-phosphorylate and activate other transcription factors responsible for stress responses, typically leading to ABA-directed growth suppression. In the 35Spro:MYB64 lines it appears that excess levels of a PP2C may be outcompeting exogenously applied ABA such that the PP2C is still able to bind and dephosphorylate SnRK2 (inset), thus maintaining an inhibitory effect on this pathway.

(Figure adapted from Kim, 2014)

As presented in this thesis, 35Spro:MYB64 lines are mildly insensitive to ABA-mediated growth suppression, indicating that the MYB64-mediated upregulation of *PP2C* noted in the NGS data results in its expression at levels in excess of the ABA concentrations used in the experiments presented here (section 5.4, "ABA Response of 35Spro:MYB64 lines"). PP2C therefore appears to be maintaining an inhibitory effect on the SnRK2s and suppressing the normal ABA growth arrest response in these transgenic lines.

6.5 Localisation and Function of the smHSPs

The finding that the tagged smHSPs (HSP17.6, HPS17.6a and HSP17.6b) localise in the cytosol confirms their compartmentalisation predicted in the literature (Scharf et al., 2001, Waters, 2013) and the punctate patterns deserve future investigation. Except where stated, leaves imaged by confocal microscopy were not subjected to any thermal acclimation or stress regimes, but they were illuminated by laser light for considerable periods of time, in some cases for an hour or more. It is feasible that this will have caused a localised heating effect in the leaf and this may actually have induced the smHSPs in some cells to undergo the same sort of multimerisation described in the literature (Baldwin et al., 2011b, Painter et al., 2008, Sobott et al., 2002, Stengel et al., 2010). These higher-order multimers are known to form within the first hour at 34 °C and within the first 15 minutes at 45 °C (Benesch et al., 2010). Since those studies were performed *in vitro* we know nothing about the localisation of the activated multimers in the cell; they may be disperse and homogenous or they may sequester semi-native proteins in clusters throughout the cytosol, held together in a manner analogous to inclusion bodies (though with a much smaller diameter

and in greater numbers throughout the cell). This clustering theory would explain the punctate pattern of fluorescence (see Figure 3.9 and Figure 3.10).

The timecourse video files complement the static images and show cytoplasmic streaming across the vacuole and around organelles, and demonstrate that the areas of complex smooth/punctate signal are highly dynamic and do not appear to move in a consistent direction, which would have indicated trafficking. This further confirms that these smHSPs are not bound to any particular membrane or organelle and are likely to be cytosolic.

The similar subcellular dynamics illustrations of HSP17.6 and HSP17.6b, the co-localisation analysis, and the weak thermosensitivity phenotype of the smHSP knockout lines reported in this thesis and by others all suggest that the smHSPs are highly redundant. The hydrophobicity profiles of HSP17.6, HSP17.6a, HSP17.6b and TaHSP16.9 (Figure 3.2) show that with the exception of HSP17.6 and HSP17.6a, which appear to be genetic duplicates and are thus very similar, the N-terminal is the only part with notable variation. Recent publications have supplied further evidence that this is the most important region for substrate binding and chaperone activity (Basha et al., 2012, 2013). The emerging model for smHSP function is that dimers, or perhaps the dodecamer, provide biological activity. Studies on dimerisation of smHSPs *in vitro* suggest all combinations of monomers are possible (Benesch et al., 2010, Stengel et al., 2010). It is feasible that specific heterodimers recognise and target particular domains of denatured proteins. An aggregation of smHSP dimers and possibly higher-order structures thus forms around the denatured target protein and HSP70 is then recruited to load the target protein onto HSP90 for refolding or tagging for degradation (Daugaard et al., 2007). Each of the 66 possible heterodimer combinations of the

11 cytosolic Arabidopsis smHSPs probably demonstrate different but overlapping affinities for denatured protein and this explains the partial functional redundancy observed in the smHSP knockout lines and their corresponding weak thermotolerance phenotype. In addition, one very recent study further complicated the heat shock protein expression paradigm by demonstrating that HSP17.6a and HSP101 were transcriptionally expressed upon exposure to the microgravity of spaceflight (Zupanska et al., 2013), possibly suggesting a role for these chaperones in mechanical cell support.

6.6 Final Conclusion

This thesis has presented the results of investigations of two Arabidopsis smHSPs regulated by MYB64, the contribution of MYB64 to abiotic stress tolerance, and the range of downstream effectors that might mediate those tolerant phenotypes.

Loss of the two smHSPs conferred a sensitivity phenotype upon seedlings given an acclimation and a stress when observed over several days, so these proteins may contribute to the plant's ability to maintain continued longer-term growth after heating. The knockout lines were less sensitive to moderate salt stress of 40 mM, however, suggesting that a homeostatic feedback mechanism might compensate by upregulating other smHSPs or other halotolerance responses. Further work on this topic should include assessments of adaptability to long-term high (sub-lethal) temperatures rather than relatively short periods at critical temperatures. It would also be interesting to create stably transformed lines expressing epitope-tagged versions of these proteins so that immunoprecipitation experiments could be carried out to investigate differences

in chaperone client specificity. These lines could also be used to look for *in vivo* interactions between smHSP family members both before and after heat stress; something that has not yet been published in the literature. It would also be of interest to repeat the knockout screens with double, rather than single, knockouts.

Lines overexpressing *MYB64* demonstrated suppression of growth when given a heat stress and observed over the following 3 days, again suggesting that this gene might contribute to the longer-term survival of the plant, in this case by over-activating innate growth suppression mechanisms. These lines exhibited a strong halotolerance phenotype when salt-stressed, achieving a higher fresh weight while at the same time reducing the length of the root system and increasing lateral root formation to a greater extent than the wild type line. The application of exogenous ABA (which activates *MYB64* expression) without any environmental stress factor did not induce the normal growth arrest response demonstrated by wild type, indicating that these overexpressing lines are either already acclimated to high levels of MYB64 or that something in the upregulated regulon is causing insensitivity; PP2C, for example. Further work on these lines could include transcript profiling of salt-stressed plants as these lines have such a stronger tolerant phenotype to salt than to high temperature.

In summary, a significant physiological phenotype was observed in one 35Spro:MYB64 line (141) on salt and ABA, and differences between the phenotype of this and transgenic line and another (127) were reflected in qPCR transcript profiles. While an initial transcript profile indicated a role for *MYB64* in thermotolerance, any such profile has yet to be conclusively identified.

The range of transcripts affected by *MYB64* overexpression, however, integrate several pathways that have classically been thought of as discrete (Mittler and Blumwald, 2010), but which more recent literature suggests act synergistically in more complex ways that once thought (AbuQamar et al., 2009, Narsai et al., 2013, Suzuki et al., 2005) and which it is sensible to expect would actually behave differently again upon challenge by multiple simultaneous stresses (Ramegowda and Senthil-Kumar, 2015). In keeping with the paradigm shift to a view of stress responses as largely interrelated, it should be noted that the MYB64-responsive transcript profile examined here: includes components of ABA-dependent (*e.g.* PP2C / SnRK2) and ABA-independent (*e.g.* MPK11) pathways; suggests a role in the regulation of GA / DELLA metabolism (*e.g.* CBF1); and even support the theory that biotic stress responses may be involved (*e.g.* ERF11). The action and regulation of the transcription factor MYB64 are therefore rich in opportunities for future phenotypic research.

Appendix i List of Transcripts Upregulated >6-fold in 35Spro:MYB64 Line 141

Gene Symbol	AGI Number	Fold Change	WT				p35S::AtMYB64-141				
			RPKM	Unique exon reads	Unique exon-exon reads	Unique intron-exon reads	RPKM	Unique exon reads	Unique exon-exon reads	Unique intron-exon reads	RPKM
MPK11	AT4G23680	**171.55**	1.03	3	2	0	177.24	928	83	3	177.24
MSN2.3	AT5G66650	**25.73**	1.01	5	1	0	25.88	232	12	2	25.88
T1F9.17	AT1G61340	**22.63**	1.34	5	0	1	30.23	204	26	7	30.23
F5K20.280	AT3G53980	**21.91**	1.75	4	1	0	38.28	158	14	0	38.28
ERF11	AT1G28370	**20.41**	2.66	10	0	0	54.3	368	0	0	54.3
F9F8.25	AT3G10930	**18.22**	2.58	7	0	0	46.93	230	0	0	46.93
BAP1	AT3G61190	**18.19**	1.4	5	0	0	25.49	164	0	0	25.49
BG3	AT3G57240	**17.75**	0.63	3	1	0	11.19	96	2	0	11.19
CZF2	AT5G04340	**17.49**	6.36	24	0	0	111.25	757	0	0	111.25
T13D8.8	AT1G60190	**17.44**	1.02	9	0	0	17.83	283	0	0	17.83
CBF1	AT4G25490	**17.13**	2.46	8	0	0	42.15	273	0	0	42.15
PBP1	AT5G54490	**16.54**	2.94	6	0	0	48.63	179	0	0	48.63
PP2C	AT3G29000	**15.97**	2.03	5	0	0	32.48	144	0	0	32.48
NF-YA10	AT5G06510	**15.67**	1.12	8	2	0	17.6	226	43	1	17.6
F13E17.22	AT3G02840	**15.61**	1.42	7	0	0	22.11	197	0	0	22.11
MYB51	AT1G18570	**15.49**	2.03	13	1	0	31.48	363	34	0	31.48
ERF5	AT5G47230	**15.19**	3.91	18	0	0	59.37	493	0	0	59.37
(None given)	AT4G27652	**14.23**	2.96	6	0	0	42.15	154	0	0	42.15
ERF-6-6	AT4G17490	**14.16**	10.65	43	0	0	150.85	1098	0	0	150.85
DIC1	AT2G22500	**13.48**	8.63	51	0	0	116.4	1240	0	0	116.4
CML38	AT1G76650	**12.70**	7.41	22	0	0	94.13	504	0	0	94.13
ERF MWD22	AT5G51190	**12.64**	7.76	25	0	0	98.15	570	0	0	98.15
RRTF1	AT4G34410	**12.25**	2.05	10	0	0	25.17	225	0	0	25.17
T16O11.2	AT3G09020	**12.20**	0.59	3	0	0	7.2	66	0	0	7.2
F37B13.20	AT4G29780	**12.02**	15.41	119	0	0	185.24	2579	0	0	185.24
K7B16.1	AT5G50800	**11.96**	1.45	7	1	0	17.3	151	32	0	17.3
WAKL2	AT1G16130	**11.46**	0.3	3	0	0	3.48	62	1	0	3.48
COBL8	AT3G16860	**11.40**	0.99	9	0	0	11.24	185	6	0	11.24
SHINE2	AT5G25390	**11.28**	0.79	3	0	0	8.91	61	4	1	8.91
MXK3.10	AT5G64870	**11.23**	0.59	4	0	0	6.62	81	1	0	6.62
MUF8.3	AT5G41750	**11.12**	2.42	34	0	1	26.9	727	10	10	26.9
T32G9.25	AT1G35210	**11.02**	2.76	8	0	0	30.37	159	0	0	30.37
(None given)	AT1G07135	**10.92**	35.33	100	0	0	385.82	1969	0	0	385.82
CNI1	AT5G27420	**10.81**	4.17	24	0	0	45.05	468	0	0	45.05
SZF2	AT2G40140	**10.79**	11.87	106	1	0	128.1	2063	41	2	128.1
WRKY46	AT2G46400	**10.60**	4.02	19	3	0	42.61	363	33	2	42.61
T6G21.2	AT5G22250	**10.59**	4.15	20	0	0	43.94	382	0	0	43.94
DREB2A	AT5G05410	**10.54**	0.53	3	0	0	5.58	57	3	0	5.58
F2H17.17	AT2G36220	**10.29**	7.21	34	0	0	74.25	631	0	0	74.25
F15E12.17	AT1G66090	**10.19**	5.32	32	1	0	54.24	588	29	6	54.24
F6N7.24	AT5G52750	**10.04**	8.49	21	2	1	85.16	380	60	5	85.16
T19K4.140	AT4G24570	**9.96**	34.28	158	0	0	341.38	2837	0	0	341.38
(None given)	AT4G36010	**9.92**	1.72	9	0	0	17.02	161	8	0	17.02
T10D17.50	AT3G44260	**9.81**	38.3	154	0	0	375.69	2724	0	0	375.69
STZ	AT1G27730	**9.81**	35.48	128	0	0	348.06	2264	0	0	348.06
F18B13.21	AT1G80130	**9.68**	2.45	13	1	0	23.68	227	31	2	23.68
T22D6.1	AT5G08030	**9.61**	0.62	3	1	0	5.96	52	12	1	5.96
F14D16.17	AT1G19020	**9.59**	8.38	17	0	0	80.32	294	0	0	80.32
GATL10	AT3G28340	**9.43**	1.61	10	0	0	15.18	170	0	0	15.18
F7O18.12	AT3G04640	**9.27**	15.96	48	0	0	147.86	802	7	0	147.86
(None given)	AT1G76955	**9.24**	1.28	3	1	0	11.8	50	20	0	11.8
WRKY40	AT1G80840	**9.20**	12.77	55	6	0	117.51	939	103	11	117.51
F8L21.110	AT4G11320	**9.12**	1.63	9	2	0	14.84	148	16	0	14.84

Gene Symbol	AGI Number	Fold Change	WT				p35S::AtMYB64-141				
			RPKM	Unique exon reads	Unique exon-exon reads	Unique intron-exon reads	RPKM	Unique exon reads	Unique exon-exon reads	Unique intron-exon reads	RPKM
					(continued from previous page)						
F25P17.10	AT2G24600	**9.10**	7.89	68	0	0	71.78	1116	40	13	71.78
(None given)	AT4G27654	**9.06**	1.61	3	0	0	14.58	49	0	0	14.58
WRKY33	AT2G38470	**9.04**	18.63	137	14	2	168.32	2232	280	44	168.32
DGAT	AT2G19450	**9.01**	0.5	4	2	0	4.5	65	25	0	4.5
	AT1G24145	**8.87**	1.98	4	0	0	17.55	64	6	2	17.55
	AT1G34060	**8.87**	0.65	4	0	0	5.73	59	8	0	5.73
RABH1C	AT4G39890	**8.69**	0.84	3	1	0	7.31	47	14	0	7.31
PP2C	AT2G30020	**8.63**	11.07	58	3	0	95.55	902	27	1	95.55
F3I6.6	AT1G24140	**8.60**	3.08	16	0	0	26.49	248	0	0	26.49
PP2-A5	AT1G65390	**8.60**	1.49	8	1	0	12.8	124	8	4	12.8
F3N23.12	AT1G72920	**8.59**	7.49	28	2	0	64.34	406	32	5	64.34
K5F14.7	AT5G54720	**8.56**	2.4	7	0	0	20.52	107	0	0	20.52
RHL41	AT5G59820	**8.53**	6.66	21	0	0	56.78	323	0	0	56.78
PK19	AT3G08720	**8.52**	1.37	11	3	0	11.67	169	17	3	11.67
(None given)	AT5G35735	**8.48**	20.13	112	2	2	170.75	1713	65	3	170.75
	AT2G02100	**8.47**	42.59	83	8	0	360.86	1268	256	17	360.86
F13F21.11	AT1G49450	**8.46**	0.51	4	0	0	4.34	61	0	0	4.34
	AT5G26920	**8.45**	6.47	52	2	0	54.67	792	70	7	54.67
F13I13.2	AT2G38790	**8.42**	6.49	27	0	0	54.65	410	0	0	54.65
AOC3	AT3G25780	**8.39**	2.05	8	1	0	17.23	118	5	1	17.23
BCS1	AT3G50930	**8.35**	4.14	37	0	0	34.53	557	0	0	34.53
NF-YA2	AT3G05690	**8.32**	5.12	29	7	0	42.62	435	79	0	42.62
PRR5	AT5G24470	**8.32**	0.57	5	1	0	4.77	75	12	1	4.77
	AT5G45340	**8.32**	6.12	45	4	0	50.9	675	88	1	50.9
NUDT21	AT1G73540	**8.30**	25.15	92	9	2	208.78	1377	113	13	208.78
	AT1G72520	**8.12**	2.64	31	0	0	21.41	454	31	2	21.41
ATNAS3	AT1G09240	**8.10**	1.03	5	0	0	8.38	72	0	0	8.38
ICS2	AT1G18870	**8.07**	1.47	11	5	0	11.89	160	76	2	11.89
F7J8.80	AT5G01100	**7.92**	0.77	7	2	0	6.07	100	20	0	6.07
	AT2G41640	**7.90**	2.29	16	1	0	18.07	228	11	0	18.07
	AT4G13395	**7.82**	11.51	19	0	0	90.02	268	0	0	90.02
	AT1G35830	**7.76**	1.71	6	0	0	13.25	84	0	0	13.25
	AT3G57450	**7.75**	16.14	43	0	0	125.11	601	0	0	125.11
	AT3G52400	**7.72**	16.26	84	5	0	125.51	1169	46	3	125.51
	AT5G56840	**7.65**	1.31	5	1	0	10.02	69	8	1	10.02
	AT1G64360	**7.49**	17	37	4	0	127.38	500	62	0	127.38
	AT5G64660	**7.48**	4.2	25	0	0	31.39	337	0	0	31.39
	AT5G57220	**7.42**	1.2	8	0	0	8.91	107	7	0	8.91
	AT4G25480	**7.41**	3.13	11	0	0	23.22	147	0	0	23.22
	AT4G29610	**7.41**	4.99	17	0	0	36.92	227	0	0	36.92
	AT3G46620	**7.40**	18.3	97	0	0	135.48	1295	0	0	135.48
	AT1G17380	**7.39**	1.37	6	3	0	10.13	80	14	1	10.13
	AT3G57540	**7.39**	0.62	3	0	0	4.58	40	2	2	4.58
	AT1G14870	**7.35**	1.4	4	0	0	10.27	53	19	2	10.27
	AT2G15390	**7.33**	1.93	14	0	0	14.17	181	9	0	14.17
	AT5G06860	**7.32**	1.11	0	0	0	8.1	4	4	0	8.1
	AT1G74450	**7.29**	8.61	59	0	0	62.81	776	0	0	62.81
	AT5G49280	**7.29**	2.12	7	0	0	15.42	92	0	0	15.42
	AT1G23710	**7.26**	11.88	54	0	0	86.23	707	0	0	86.23
	AT2G01180	**7.26**	6.11	45	1	0	44.37	589	27	0	44.37
	AT3G44300	**7.21**	1.09	6	1	0	7.86	71	10	0	7.86
	AT1G61800	**7.13**	1.15	7	2	0	8.23	90	12	0	8.23
	AT1G72910	**7.12**	23.68	95	0	0	168.69	1291	0	3	168.69
	AT1G56660	**7.11**	3.83	27	1	0	27.25	346	5	2	27.25
	AT3G16510	**7.07**	3.97	20	0	0	28.07	255	0	0	28.07
	AT2G20142	**7.07**	1	4	1	0	7.09	51	1	5	7.09
	AT4G14365	**7.06**	6.62	36	14	1	46.69	458	155	12	46.69

Gene Symbol	AGI Number	Fold Change	WT				p35S::AtMYB64-141				
			RPKM	Unique exon reads	Unique exon-exon reads	Unique intron-exon reads	RPKM	Unique exon reads	Unique exon-exon reads	Unique intron-exon reads	RPKM
			(continued from previous page)								
	AT3G55980	**7.04**	39.28	327	1	0	276.68	4153	69	4	276.68
	AT1G63720	**7.02**	1.02	6	0	0	7.15	76	1	0	7.15
	AT5G47850	**7.02**	0.33	3	0	0	2.29	38	0	0	2.29
	AT1G15010	**7.01**	5.31	11	0	0	37.2	139	0	0	37.2
	AT3G49530	**7.00**	6.6	46	7	0	46.25	581	84	6	46.25
	AT5G46710	**6.95**	4.88	21	3	0	33.92	263	33	5	33.92
	AT4G17670	**6.93**	1.24	4	0	0	8.63	50	3	0	8.63
	AT3G46090	**6.93**	2.44	6	0	0	16.94	75	0	0	16.94
	AT1G14520	**6.93**	1.45	8	4	0	10.05	100	30	8	10.05
	AT1G55760	**6.93**	0.59	4	1	0	4.11	50	5	0	4.11
	AT4G33040	**6.93**	1.34	4	0	0	9.31	50	0	0	9.31
	AT2G42530	**6.91**	16.24	41	8	0	112.24	511	69	2	112.24
	AT1G18300	**6.83**	14.57	60	0	0	99.53	739	14	4	99.53
	AT5G19240	**6.80**	54.72	161	5	0	372.06	1974	59	88	372.06
	AT1G67970	**6.73**	2.55	14	1	0	17.18	170	3	0	17.18
	AT3G02610	**6.65**	0.55	3	0	0	3.68	31	1	0	3.68
	AT5G17350	**6.65**	2.78	9	0	0	18.51	108	0	0	18.51
	AT4G02380	**6.65**	12.97	35	2	0	86.31	420	23	7	86.31
	AT5G47220	**6.62**	8.87	33	0	0	58.75	394	0	0	58.75
	AT5G41740	**6.56**	5.22	68	2	6	34.28	765	28	70	34.28

Table A1.1 Genes Upregulated >6.5-fold in 35Spro:MYB64 Line 141 and Detectable in Col-0

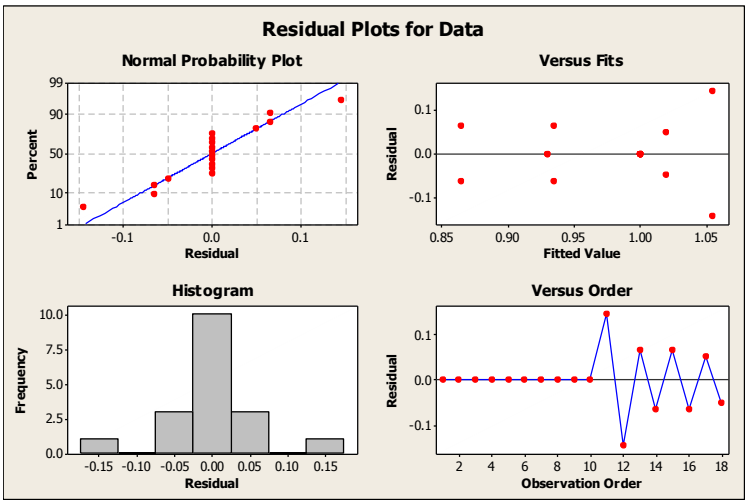
List of genes upregulated in the p35S::AtMYB64-141 lines compared with Col-0. Results have been filtered to show only genes upregulated >6.5-fold, and to exclude those represented by <3 sequencing hits in the Col-0 sample. Gene symbols are correct as of the TAIR9 genome release (www.arabiopsis.org). Gold symbols = chosen for further study; red symbols = abiotic associations; blue symbols = biotic associations; white symbols = no obvious relation to stress responses; grey symbols = uncharacterised gene.

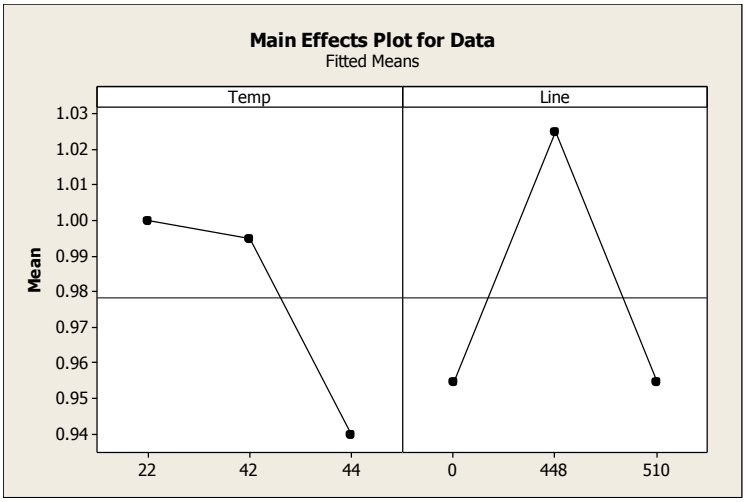
Appendix ii Statistical Analyses

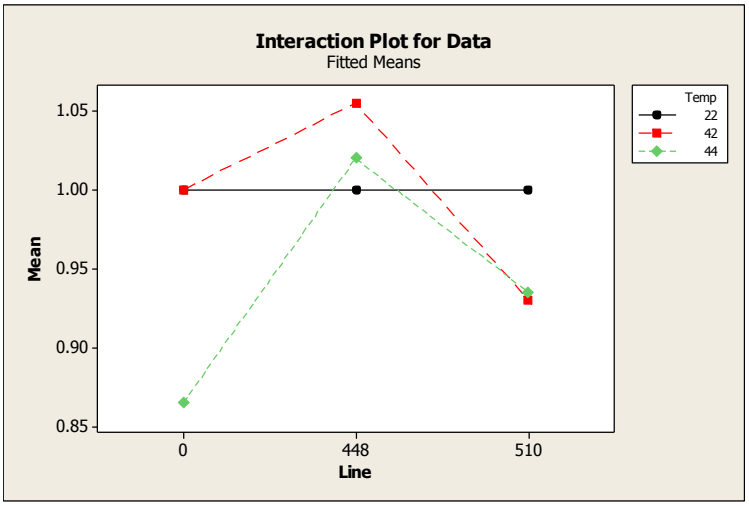
Appendix ii-a : Germination of Knockout Lines hsp17.6, hsp17.6a, and Wild Type Seedlings Following Heat Stress (Accompanies Figure 3.16)

General Linear Model: Data versus Temp, Line

```
Factor  Type   Levels  Values
Temp    fixed       3  22, 42, 44
Line    fixed       3  0, 448, 510


Analysis of Variance for Data, using Adjusted SS for Tests

Source      DF    Seq SS    Adj SS    Adj MS     F      P
Temp         2  0.013300  0.013300  0.006650  0.94  0.427
Line         2  0.019600  0.019600  0.009800  1.38  0.300
Temp*Line    4  0.020200  0.020200  0.005050  0.71  0.605
Error        9  0.063950  0.063950  0.007106
Total       17  0.117050


S = 0.0842945   R-Sq = 45.37%   R-Sq(adj) = 0.00%


Unusual Observations for Data

Obs     Data      Fit   SE Fit  Residual  St Resid
 11  1.20000  1.05500  0.05961   0.14500      2.43 R
 12  0.91000  1.05500  0.05961  -0.14500     -2.43 R

R denotes an observation with a large standardized residual.


Grouping Information Using Tukey Method and 95.0% Confidence

Temp  Line  N  Mean  Grouping
42     448  2   1.1  A
44     448  2   1.0  A
42       0  2   1.0  A
22     510  2   1.0  A
22     448  2   1.0  A
22       0  2   1.0  A
44     510  2   0.9  A
42     510  2   0.9  A
44       0  2   0.9  A

Means that do not share a letter are significantly different.
```


Residual Plots for Data
Normal Probability Plot
Percent
Residual
Versus Fits
Residual
Fitted Value
Histogram
Frequency
Residual
Versus Order
Residual
Observation Order


Main Effects Plot for Data
Fitted Means
Temp
Line
Mean


Interaction Plot for Data
Fitted Means
Temp
22
42
44
Mean
Line

Appendix ii-b : Analysis of Hypocotyl Extension in Knockout Lines hsp17.6 and hsp17.6a After Heat Stress (Accompanies Figure 3.18)

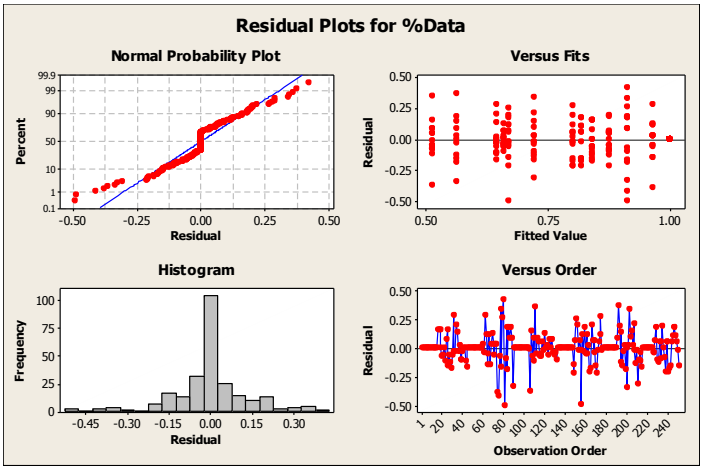
General Linear Model: %Data versus StressT, Acclim, Line

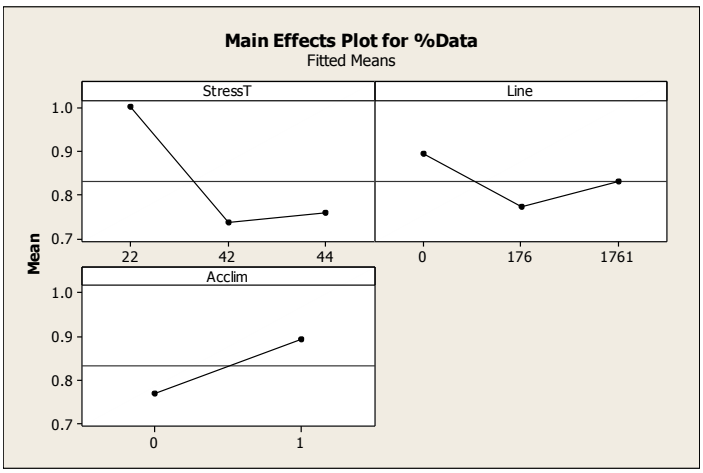
```
Factor   Type   Levels  Values
StressT  fixed       3  22, 42, 44
Acclim   fixed       2  0, 1
Line     fixed       3  0, 176, 1761


Analysis of Variance for %Data, using Adjusted SS for Tests

Source                DF    Seq SS   Adj SS   Adj MS      F      P
StressT                2   3.45114  3.51660  1.75830  99.44  0.000
Acclim                 1   0.93385  0.93424  0.93424  52.84  0.000
Line                   2   0.62254  0.63966  0.31983  18.09  0.000
StressT*Acclim         2   0.47258  0.47557  0.23778  13.45  0.000
StressT*Line           4   0.84343  0.85032  0.21258  12.02  0.000
Acclim*Line            2   0.01783  0.01705  0.00853   0.48  0.618
StressT*Acclim*Line    4   0.22975  0.22975  0.05744   3.25  0.013
Error                233   4.11996  4.11996  0.01768
Total                250  10.69107


S = 0.132974   R-Sq = 61.46%   R-Sq(adj) = 58.65%


Unusual Observations for %Data

Obs    %Data      Fit   SE Fit  Residual  St Resid
 31  0.92857  0.64294  0.03433   0.28563      2.22 R
 62  1.25000  0.96507  0.03433   0.28493      2.22 R
 74  0.58333  0.96507  0.03433  -0.38174     -2.97 R
 75  0.50000  0.91257  0.03433  -0.41257     -3.21 R
 77  1.25000  0.91257  0.03433   0.33743      2.63 R
 79  1.18182  0.91257  0.03433   0.26925      2.10 R
 80  1.33333  0.91257  0.03433   0.42076      3.28 R
 81  0.41667  0.91257  0.03433  -0.49591     -3.86 R
 89  0.58333  0.91257  0.03433  -0.32924     -2.56 R
106  0.14286  0.51050  0.03554  -0.36764     -2.87 R
111  0.86667  0.51050  0.03554   0.35616      2.78 R
151  0.92857  0.66805  0.03433   0.26052      2.03 R
156  0.17647  0.66805  0.03433  -0.49158     -3.83 R
175  1.07692  0.80057  0.03554   0.27636      2.16 R
192  0.93333  0.56026  0.04009   0.37307      2.94 R
201  0.22222  0.56026  0.04009  -0.33804     -2.67 R
203  1.06250  0.72008  0.03554   0.34242      2.67 R
211  0.41176  0.72008  0.03554  -0.30832     -2.41 R

R denotes an observation with a large standardized residual.


Grouping Information Using Tukey Method and 95.0% Confidence

Line   N  Mean  Grouping
   0  89   0.9  A
1761  75   0.8     B
 176  87   0.8        C

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

Acclim    N  Mean  Grouping
1       125   0.9  A
```

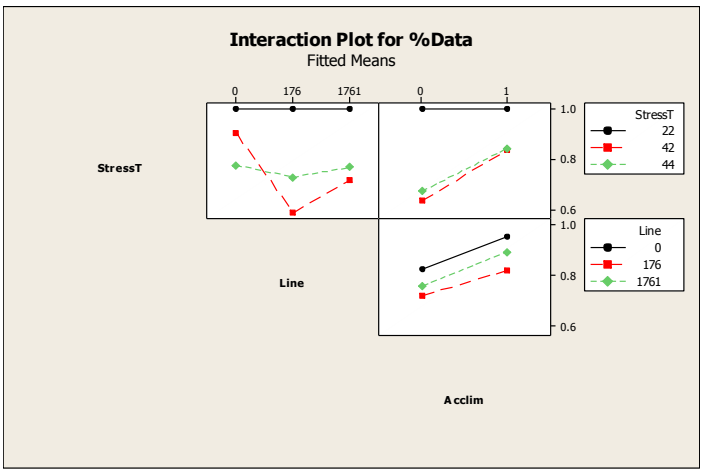
```
0       126   0.8    B

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

StressT   N  Mean  Grouping
22       84   1.0  A
44       83   0.8    B
42       84   0.7    B

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

Acclim  Line   N  Mean  Grouping
1          0  45   1.0  A
1       1761  36   0.9  A B
0          0  44   0.8    B C
1        176  44   0.8    B C
0       1761  39   0.8      C D
0        176  43   0.7        D

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

StressT  Line   N  Mean  Grouping
22         0  30   1.0  A
22       176  30   1.0  A
22      1761  24   1.0  A
42         0  29   0.9  A
44         0  30   0.8    B
44      1761  25   0.8    B
44       176  28   0.7    B
42      1761  26   0.7    B
42       176  29   0.6      C

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

StressT  Acclim   N  Mean  Grouping
22       1      40   1.0  A
22       0      44   1.0  A
44       1      40   0.8    B
42       1      45   0.8    B
44       0      43   0.7      C
42       0      39   0.6      C

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

StressT  Acclim  Line   N  Mean  Grouping
22       1         0  15   1.0  A
22       1       176  15   1.0  A
22       0      1761  14   1.0  A
22       0         0  15   1.0  A
22       1      1761  10   1.0  A
22       0       176  15   1.0  A
42       1         0  15   1.0  A B
44       1         0  15   0.9  A B
42       1      1761  15   0.9  A B C
42       0         0  14   0.8  A B C D
44       1      1761  11   0.8  A B C D E
44       1       176  14   0.8    B C D E
```

```
44         0         1761   14    0.7        C D E F
42         1          176   15    0.7          D E F G
44         0          176   14    0.7            E F G
44         0            0   15    0.6            E F G
42         0         1761   11    0.6              F G
42         0          176   14    0.5                G
```

Means that do not share a letter are significantly different.



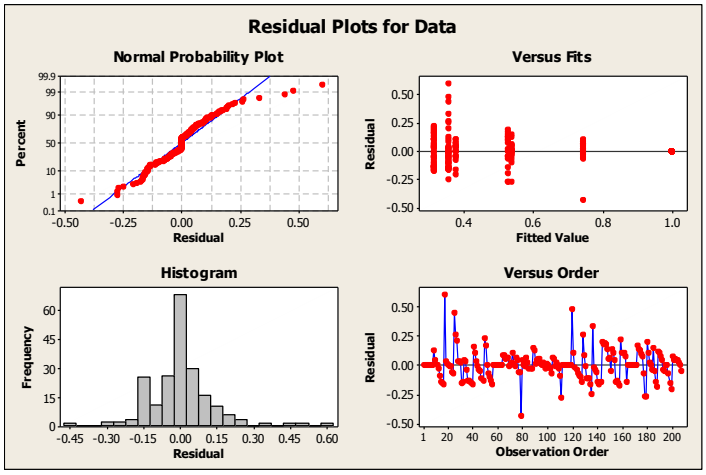


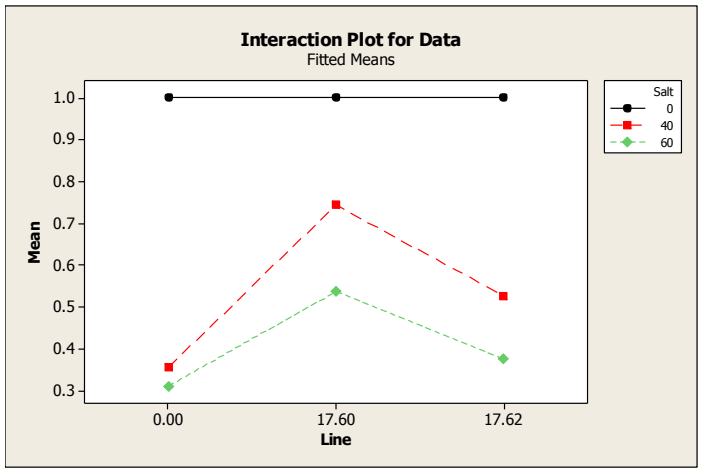


Appendix ii-c : Analysis of Fresh Weight of Knockout Lines hsp17.6 and hsp17.6a Grown in the Presence of 0, 40 or 60 mM NaCl (Accompanies Figure 3.19)

General Linear Model: Data versus Salt, Line

```
Factor  Type   Levels  Values
Salt    fixed       3  0, 40, 60
Line    fixed       3  0.00, 17.60, 17.62

Analysis of Variance for Data, using Adjusted SS for Tests

Grouping Information Using Tukey Method and 95.0% Confidence

Salt  Line     N  Mean  Grouping
 0     0.00   16   1.0  A
 0    17.62    8   1.0  A
 0    17.60    8   1.0  A
40    17.60   24   0.7     B
60    17.60   24   0.5        C
40    17.62   17   0.5        C
60    17.62   19   0.4           D
40     0.00   48   0.4           D
60     0.00   43   0.3           D

Means that do not share a letter are significantly different.
```





Appendix ii-d : Germination of 35Spro:MYB64 Seeds Following Heat Stress (Accompanies Figure 5.2)

General Linear Model: OX Data versus Temp2, Line2

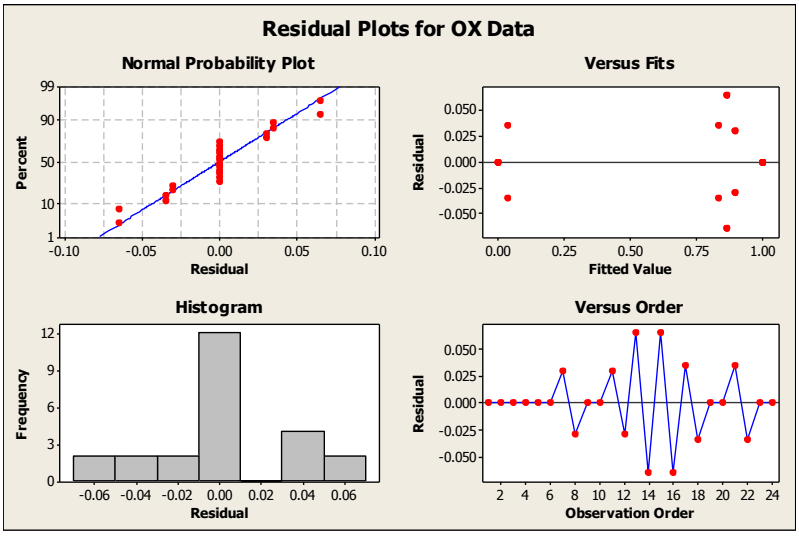
```
Factor   Type   Levels  Values
Temp2    fixed       4  22, 44, 46, 48
Line2    fixed       3  0, 127, 141
```

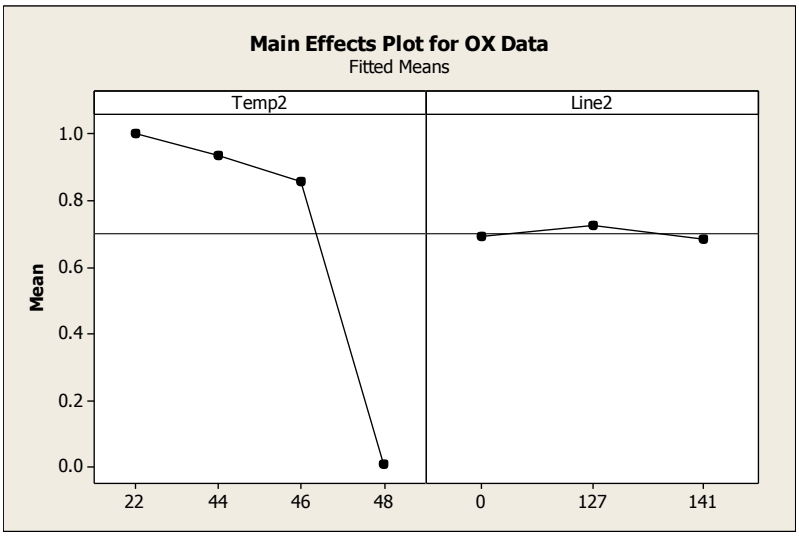
```
Analysis of Variance for OX Data, using Adjusted SS for Tests

Source        DF   Seq SS   Adj SS   Adj MS       F      P
Temp2          3  3.85363  3.85363  1.28454  606.87  0.000
Line2          2  0.00773  0.00773  0.00386    1.82  0.203
Temp2*Line2    6  0.00844  0.00844  0.00141    0.66  0.680
Error         12  0.02540  0.02540  0.00212
Total         23  3.89520

S = 0.0460072   R-Sq = 99.35%   R-Sq(adj) = 98.75%
```

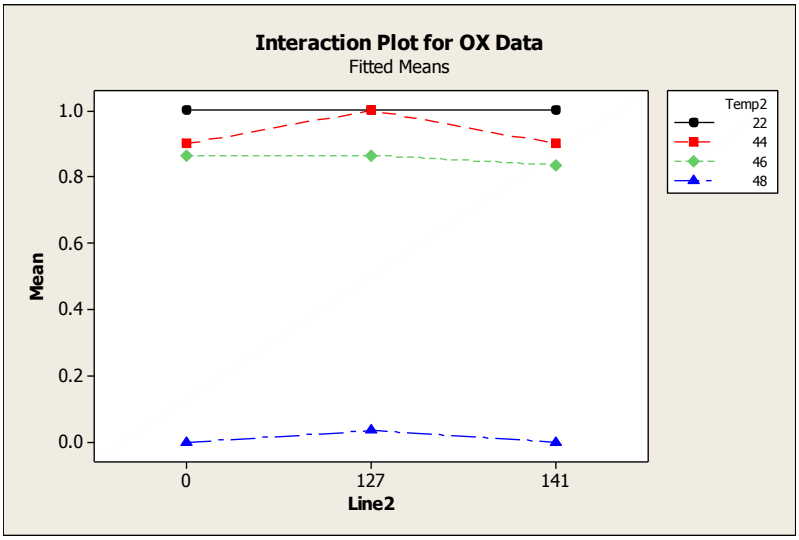
Grouping Information Using Tukey Method and 95.0% Confidence

Temp2	Line2	N	Mean	Grouping
44	127	2	1.0	A
22	127	2	1.0	A
22	141	2	1.0	A
22	0	2	1.0	A
44	141	2	0.9	A
44	0	2	0.9	A
46	0	2	0.9	A
46	127	2	0.9	A
46	141	2	0.8	A
48	127	2	0.0	B
48	0	2	0.0	B
48	141	2	-0.0	B

Means that do not share a letter are significantly different.




Main Effects Plot for OX Data
Fitted Means
Temp2
Line2
Mean
1.0
0.8
0.6
0.4
0.2
0.0
22
44
46
48
0
127
141


Interaction Plot for OX Data
Fitted Means
Mean
1.0
0.8
0.6
0.4
0.2
0.0
0
127
141
Line2
Temp2
22
44
46
48

Appendix ii-e : Effect of Heat Stress on Hypocotyl Extension in 35Spro:MYB64 Lines (Accompanies Figure 5.4)

General Linear Model: %OX versus StressTOX, LineOX, AcclimOX

```
Factor     Type   Levels  Values
StressTOX  fixed       3  22, 44, 46
LineOX     fixed       3  0, 127, 141
AcclimOX   fixed       2  0, 1


Analysis of Variance for %OX, using Adjusted SS for Tests

Source                      DF   Seq SS   Adj SS   Adj MS        F      P
StressTOX                    2  27.3865  25.0764  12.5382  1332.38  0.000
LineOX                       2   1.6910   1.7238   0.8619    91.59  0.000
AcclimOX                     1   0.2275   0.2339   0.2339    24.85  0.000
StressTOX*LineOX             4   1.0275   1.0322   0.2581    27.42  0.000
StressTOX*AcclimOX           2   0.2156   0.2647   0.1324    14.07  0.000
LineOX*AcclimOX              2   0.0895   0.0879   0.0439     4.67  0.010
StressTOX*LineOX*AcclimOX    4   0.4319   0.4319   0.1080    11.48  0.000
Error                      203   1.9103   1.9103   0.0094
Total                      220  32.9799


S = 0.0970071   R-Sq = 94.21%   R-Sq(adj) = 93.72%

Unusual Observations for %OX

Obs      %OX      Fit   SE Fit  Residual  St Resid
 13  1.18182  0.83228  0.03068   0.34953      3.80 R
 18  0.61538  0.83228  0.03068  -0.21690     -2.36 R
 21  0.66667  0.44044  0.03068   0.22622      2.46 R
 40  0.66667  1.06421  0.03234  -0.39755     -4.35 R
 41  0.69231  1.06421  0.03234  -0.37190     -4.07 R
 42  0.75000  1.06421  0.03234  -0.31421     -3.44 R
 43  0.72727  1.06421  0.03234  -0.33694     -3.68 R
 45  1.37500  1.06421  0.03234   0.31079      3.40 R
 46  1.50000  1.06421  0.03234   0.43579      4.76 R
 47  1.33333  1.06421  0.03234   0.26912      2.94 R
 48  1.33333  1.06421  0.03234   0.26912      2.94 R
 56  0.22222  0.42639  0.03234  -0.20417     -2.23 R
208  0.73333  0.94386  0.02690  -0.21053     -2.26 R

R denotes an observation with a large standardized residual.


Grouping Information Using Tukey Method and 95.0% Confidence

StressTOX   N  Mean  Grouping
22         74   1.0  A
44         74   0.8    B
46         73   0.2      C

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

LineOX   N  Mean  Grouping
  0     57   0.8  A
141     77   0.6    B
127     87   0.6      C

Means that do not share a letter are significantly different.
```

Grouping Information Using Tukey Method and 95.0% Confidence

AcclimOX	N	Mean	Grouping
1	110	0.7	A
0	111	0.6	B

Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

StressTOX	LineOX	N	Mean	Grouping
22	0	19	1.0	A
22	141	26	1.0	A
22	127	29	1.0	A
44	0	19	0.9	A
44	141	26	0.8	B
44	127	29	0.7	C
46	0	19	0.4	D
46	141	25	0.1	E
46	127	29	0.0	E

Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

StressTOX	AcclimOX	N	Mean	Grouping
22	1	37	1.0	A
22	0	37	1.0	A
44	1	37	0.9	B
44	0	37	0.7	C
46	1	36	0.2	D
46	0	37	0.2	D

Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

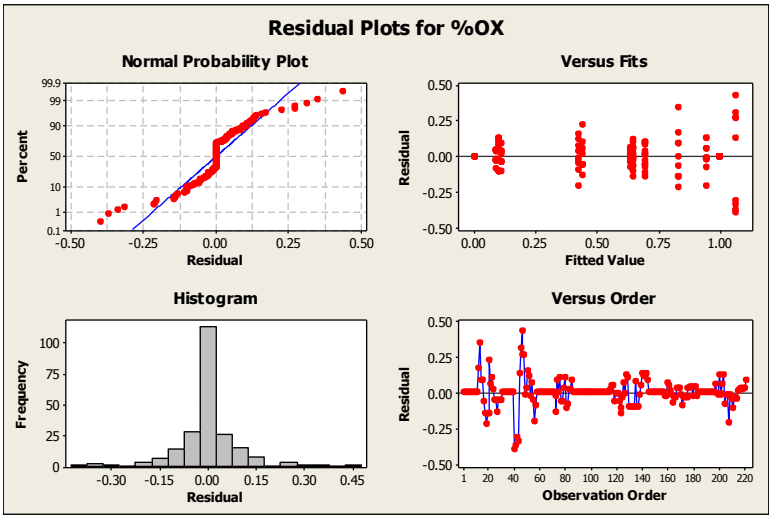
LineOX	AcclimOX	N	Mean	Grouping
0	1	27	0.8	A
0	0	30	0.8	A
141	1	38	0.7	B
127	1	45	0.6	C
141	0	39	0.6	C
127	0	42	0.6	C

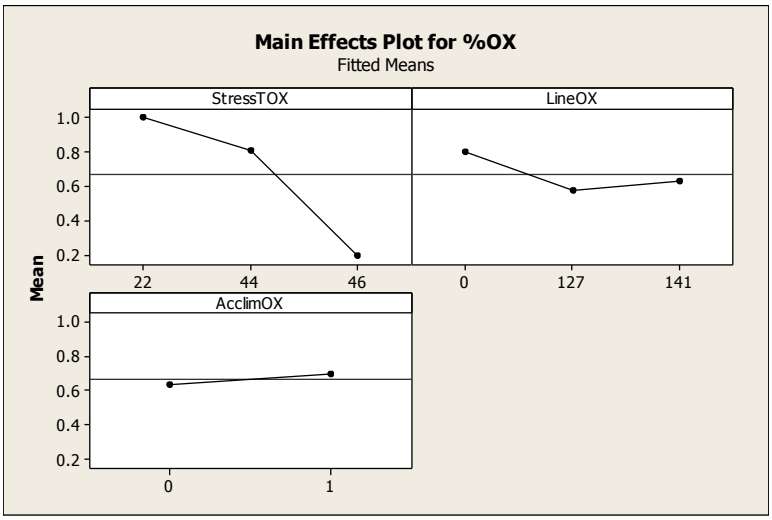
Means that do not share a letter are significantly different.

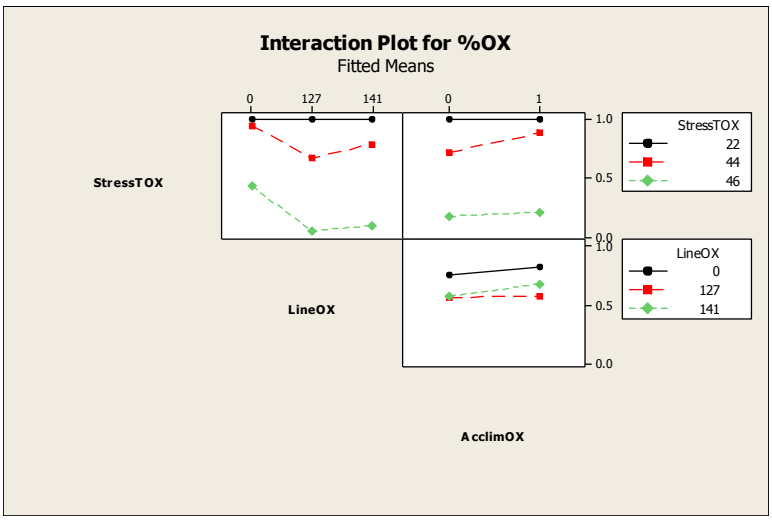
Grouping Information Using Tukey Method and 95.0% Confidence

StressTOX	LineOX	AcclimOX	N	Mean	Grouping
44	0	1	9	1.1	A
22	141	1	13	1.0	A
22	0	0	10	1.0	A
22	0	1	9	1.0	A
22	141	0	13	1.0	A
22	127	1	15	1.0	A
22	127	0	14	1.0	A
44	141	1	13	0.9	A B
44	0	0	10	0.8	B C
44	127	0	14	0.7	C D
44	127	1	15	0.6	D
44	141	0	13	0.6	D
46	0	0	10	0.4	E
46	0	1	9	0.4	E
46	141	1	12	0.1	F
46	127	1	15	0.1	F
46	141	0	13	0.1	F
46	127	0	14	-0.0	F

Means that do not share a letter are significantly different.







Appendix ii-f : Analysis of Fresh Weight of 35Spro:MYB64 Lines Grown in the Presence of 0 or 80 mM NaCl (Accompanies Figure 5.5)

General Linear Model: Data versus Line, Salt

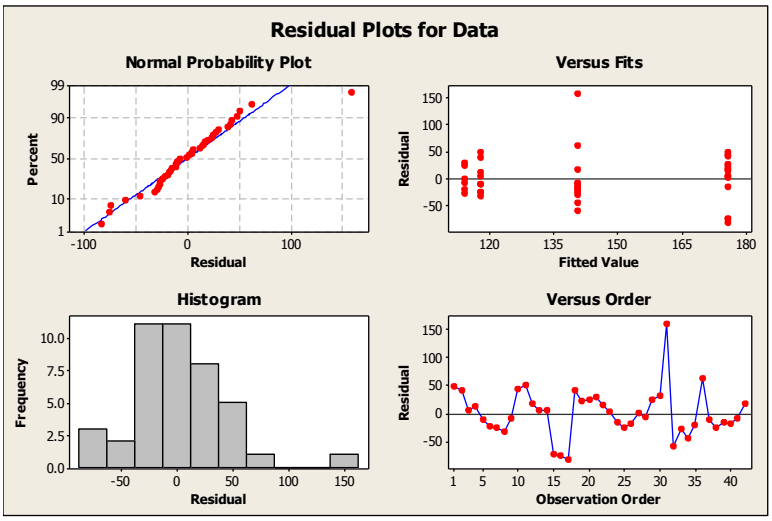
```
Factor  Type   Levels  Values
Line    fixed       2  127, 141
Salt    fixed       2  0, 80


Analysis of Variance for Data, using Adjusted SS for Tests

Source     DF  Seq SS  Adj SS  Adj MS     F      P
Line        1    5138    3576    3576  1.83  0.184
Salt        1   19555   16837   16837  8.60  0.006
Line*Salt   1    2287    2287    2287  1.17  0.287
Error      38   74383   74383    1957
Total      41  101364


S = 44.2430   R-Sq = 26.62%   R-Sq(adj) = 20.82%


Unusual Observations for Data

Obs     Data      Fit  SE Fit  Residual  St Resid
 31  300.000  140.704  12.772   159.296      3.76 R

R denotes an observation with a large standardized residual.


Grouping Information Using Tukey Method and 95.0% Confidence

Line   N   Mean  Grouping
141   24  146.9  A
127   18  127.3  A

Means that do not share a letter are significantly different.


Tukey Simultaneous Tests
Response Variable Data
All Pairwise Comparisons among Levels of Line
Line = 127  subtracted from:

      Difference       SE of           Adjusted
Line    of Means  Difference  T-Value   P-Value
141        19.56       14.47    1.352    0.1845


Grouping Information Using Tukey Method and 95.0% Confidence

Salt   N   Mean  Grouping
80    27  158.3  A
 0    15  115.9    B

Means that do not share a letter are significantly different.


Tukey Simultaneous Tests
Response Variable Data
All Pairwise Comparisons among Levels of Salt
Salt =  0  subtracted from:

      Difference       SE of           Adjusted
Salt    of Means  Difference  T-Value   P-Value
80         42.43       14.47    2.933    0.0057
```

Grouping Information Using Tukey Method and 95.0% Confidence

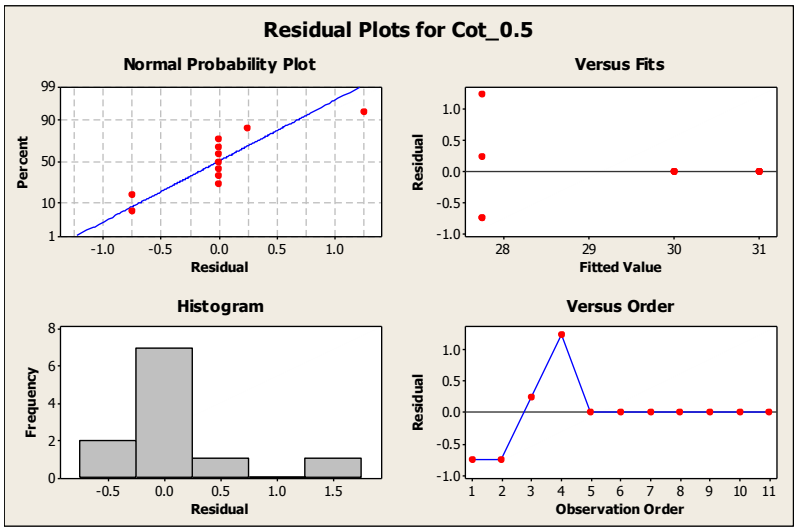
Line	Salt	N	Mean	Grouping
141	80	15	175.9	A
127	80	12	140.7	A B
141	0	9	117.8	B
127	0	6	113.9	B

Means that do not share a letter are significantly different.

Tukey's MSD = 58.06



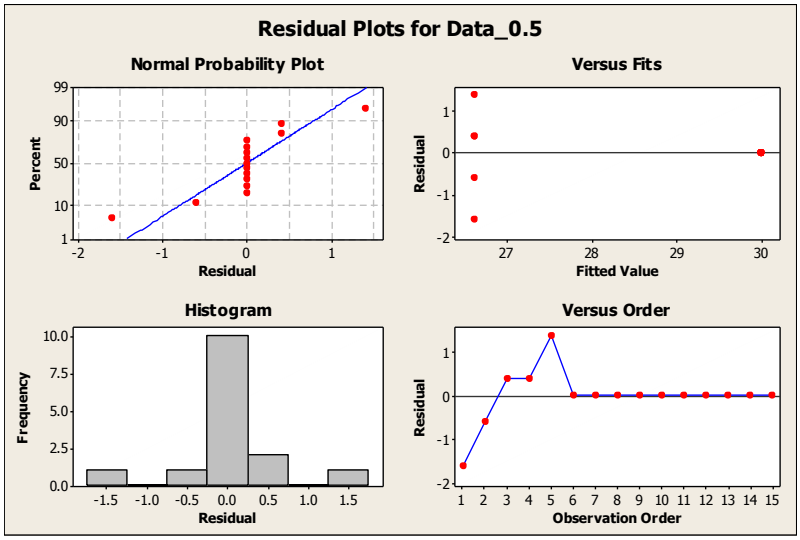
Appendix ii-g : Cotyledon Expansion in 35Spro:MYB64 Lines and Wild Type at 0.5μM ABA (Days 6-10) (Accompanies Figure 5.8)



General Linear Model: Cot_0.5 versus Line_0.5

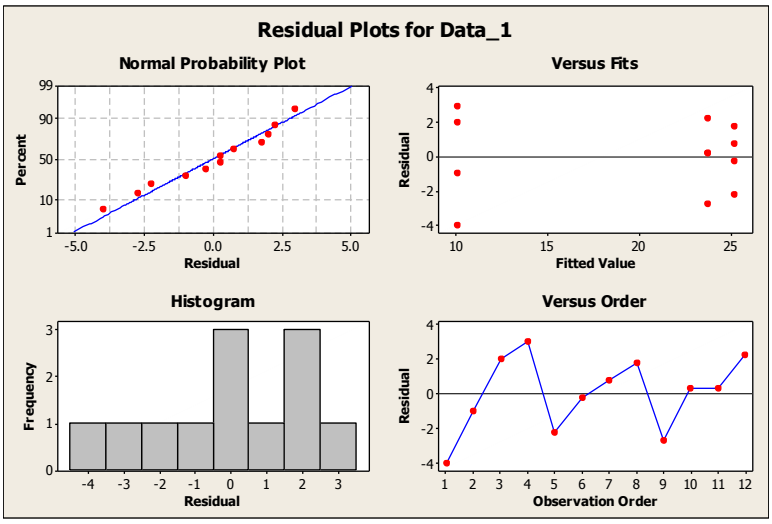
```
Factor     Type   Levels  Values
Line_0.5   fixed       3  0, 127, 141


Analysis of Variance for Cot_0.5, using Adjusted SS for Tests

Source     DF  Seq SS  Adj SS  Adj MS      F      P
Line_0.5    2  21.977  21.977  10.989  31.97  0.000
Error       8   2.750   2.750   0.344
Total      10  24.727


S = 0.586302   R-Sq = 88.88%   R-Sq(adj) = 86.10%


Unusual Observations for Cot_0.5

Obs  Cot_0.5      Fit  SE Fit  Residual  St Resid
  4  29.0000  27.7500  0.2932    1.2500      2.46 R

R denotes an observation with a large standardized residual.
```

Appendix ii-h : Analysis of Chlorophyll Production in 35Spro:MYB64 Lines and Wild Type in the Presence of 0.5 or 1μM ABA (Accompanies Figure 5.9)



General Linear Model: Data_0.5 versus Line_0.5

```
Factor     Type   Levels  Values
Line_0.5   fixed       3  0, 127, 141


Analysis of Variance for Data_0.5, using Adjusted SS for Tests

Source     DF  Seq SS  Adj SS  Adj MS      F      P
Line_0.5    2  38.533  38.533  19.267  44.46  0.000
Error      12   5.200   5.200   0.433
Total      14  43.733


S = 0.658281   R-Sq = 88.11%   R-Sq(adj) = 86.13%


Unusual Observations for Data_0.5

Obs  Data_0.5      Fit  SE Fit  Residual  St Resid
  1   25.0000  26.6000  0.2944   -1.6000     -2.72 R
  5   28.0000  26.6000  0.2944    1.4000      2.38 R

R denotes an observation with a large standardized residual.


Grouping Information Using Tukey Method and 99.9% Confidence

Line_0.5  N  Mean  Grouping
141       5  30.0  A
127       5  30.0  A
  0       5  26.6     B

Means that do not share a letter are significantly different.
```



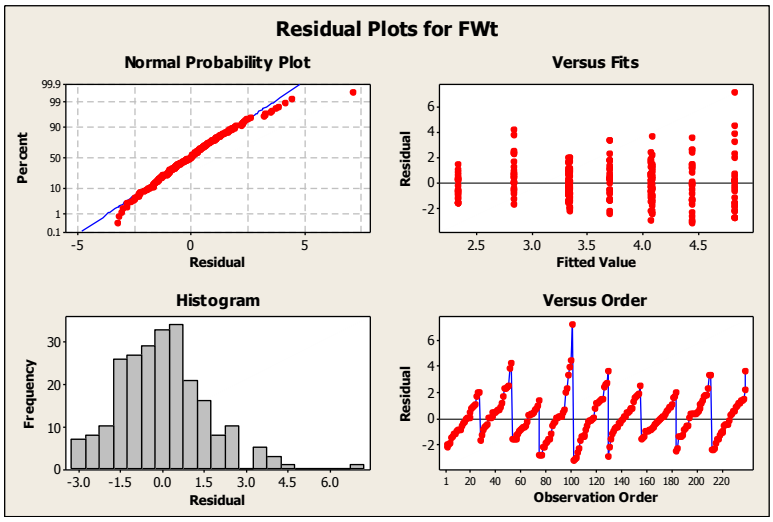
General Linear Model: Data_1 versus Line_1

```
Factor  Type   Levels  Values
Line_1  fixed       3  0, 127, 141


Analysis of Variance for Data_1, using Adjusted SS for Tests

Source  DF  Seq SS  Adj SS  Adj MS      F      P
Line_1   2  565.17  565.17  282.58  49.38  0.000
Error    9   51.50   51.50    5.72
Total   11  616.67


S = 2.39212   R-Sq = 91.65%   R-Sq(adj) = 89.79%


Grouping Information Using Tukey Method and 99.9% Confidence

Line_1  N  Mean  Grouping
127     4  25.2  A
141     4  23.7  A
  0     4  10.0     B

Means that do not share a letter are significantly different.
```

Appendix ii-i : Analysis of Fresh Weight of 35Spro:MYB64 Lines and Wild Type Germinated in the Presence of 0, 0.5, or 1μM ABA (Accompanies Figure 5.10)

General Linear Model: FWt versus CodeABA, CodeLine

```
Factor     Type   Levels  Values
CodeABA    fixed       3  0.0, 0.5, 1.0
CodeLine   fixed       3  0, 127, 141


Analysis of Variance for FWt, using Adjusted SS for Tests

Source              DF   Seq SS   Adj SS  Adj MS      F      P
CodeABA              2   10.876   14.368   7.184   2.99  0.052
CodeLine             2   99.434  103.430  51.715  21.54  0.000
CodeABA*CodeLine     4   41.313   41.313  10.328   4.30  0.002
Error              230  552.177  552.177   2.401
Total              238  703.800


S = 1.54944   R-Sq = 21.54%   R-Sq(adj) = 18.81%


Unusual Observations for FWt

Obs      FWt     Fit  SE Fit  Residual  St Resid
 51   6.6000  3.5600  0.3099    3.0400      2.00 R
 52   7.0000  3.5600  0.3099    3.4400      2.27 R
 98   8.1000  4.9222  0.2982    3.1778      2.09 R
 99   8.7000  4.9222  0.2982    3.7778      2.48 R
100   9.3000  4.9222  0.2982    4.3778      2.88 R
101  12.0000  4.9222  0.2982    7.0778      4.65 R
102   1.2100  4.3004  0.2928   -3.0904     -2.03 R
129   8.0000  4.3004  0.2928    3.6996      2.43 R
130   1.1000  4.1538  0.3039   -3.0538     -2.01 R
239   7.7000  3.9571  0.2928    3.7429      2.46 R

R denotes an observation with a large standardized residual.

Grouping Information Using Tukey Method and 99.9% Confidence

CodeABA  CodeLine   N  Mean  Grouping
0.0      127       27   4.9  A
0.5      127       28   4.3  A B
1.0      127       26   4.2  A B
0.5      141       28   4.0  A B
1.0      141       28   4.0  A B
0.5        0       25   3.6  A B C
0.0      141       28   3.2  A B C
0.0        0       27   3.0    B C
1.0        0       22   1.9      C

Means that do not share a letter are significantly different.
```


Residual Plots for FWt
Normal Probability Plot
Percent
99.9
99
90
50
10
1
0.1
-5
0
5
Residual
Versus Fits
Residual
6
4
2
0
-2
2.5
3.0
3.5
4.0
4.5
Fitted Value
Histogram
Frequency
30
20
10
0
-3.0
-1.5
0.0
1.5
3.0
4.5
6.0
Residual
Versus Order
Residual
6
4
2
0
-2
1
20
40
60
80
100
120
140
160
180
200
220
Observation Order

References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell,* 15**,** 63-78.

AbuQamar, S., Luo, H., Laluk, K., Mickelbart, M. V. & Mengiste, T. 2009. Crosstalk between biotic and abiotic stress responses in tomato is mediated by the AIM1 transcription factor. *Plant Journal,* 58**,** 347-360.

Abwao, S. I. 2011. *Translational control of abiotic stress responses in Arabidopsis thaliana*. PhD, University of Glasgow.

Achard, P., Cheng, H., De Grauwe, L., Decat, J., Schoutteten, H., Moritz, T., Van der Straeten, D., Peng, J. R. & Harberd, N. P. 2006. Integration of plant responses to environmentally activated phytohormonal signals. *Science,* 311**,** 91-94.

Achard, P., Gong, F., Cheminant, S., Alioua, M., Hedden, P. & Genschik, P. 2008a. The cold-inducible CBF1 factor-dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. *Plant Cell,* 20**,** 2117-2129.

Achard, P., Renou, J.-P., Berthome, R., Harberd, N. P. & Genschik, P. 2008b. Plant DELLAs restrain growth and promote survival of adversity by reducing the levels of reactive oxygen species. *Current Biology,* 18**,** 656-660.

Adie, B., Chico, J. M., Rubio-Somoza, I. & Solano, R. 2007. Modulation of plant defenses by ethylene. *Journal of Plant Growth Regulation,* 26**,** 160-177.

Almalki, N. 2011. Investigation of temperature gradients across shaded and unshaded barley leaves. Glasgow: University of Glasgow.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C. & Ecker, J. R. 2003. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science,* 301**,** 653-657.

Ambawat, S., Sharma, P., Yadav, N. R. & Yadav, R. C. 2013. MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants,* 19**,** 307-321.

Antoni, R., Rodriguez, L., Gonzalez-Guzman, M., Pizzio, G. A. & Rodriguez, P. L. 2011. News on ABA transport, protein degradation, and ABFs/WRKYs in ABA signaling. *Current Opinion in Plant Biology,* 14**,** 547-553.

Arrigo, A. P. 2013. Human small heat shock proteins: Protein interactomes of homo- and hetero-oligomeric complexes: An update. *Febs Letters,* 587**,** 1959-1969.

Atkin, O. K. & Macherel, D. 2009. The crucial role of plant mitochondria in orchestrating drought tolerance. *Annals of Botany,* 103**,** 581-597.

Bachmair, A., Novatchkova, M., Potuschak, T. & Eisenhaber, F. 2001. Ubiquitylation in plants: a post-genomic look at a post-translational modification. *Trends in Plant Science,* 6**,** 463-70.

Baldwin, A. J., Lioe, H., Robinson, C. V., Kay, L. E. & Benesch, J. L. P. 2011a. alpha B-Crystallin Polydispersity Is a Consequence of Unbiased Quaternary Dynamics. *Journal of Molecular Biology,* 413**,** 297-309.

Baldwin, A. J., Lioe, H., Robinson, C. V., Kay, L. E. & Benesch, J. L. P. 2011b. αB-Crystallin polydispersity is a consequence of unbiased quaternary dynamics. *Journal of Molecular Biology,* 413**,** 297-309.

Barnabas, B., Jaeger, K. & Feher, A. 2008. The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell and Environment,* 31**,** 11-38.

Basha, E., Jones, C., Blackwell, A. E., Cheng, G., Waters, E. R., Samsel, K. A., Siddique, M., Pett, V., Wysocki, V. & Vierling, E. 2013. An unusual dimeric small heat shock protein provides insight into the mechanism of this class of chaperones. *Journal of Molecular Biology,* 425**,** 1683-96.

Basha, E., O'Neill, H. & Vierling, E. 2012. Small heat shock proteins and alpha-crystallins: Dynamic proteins with flexible functions. *Trends in Biochemical Sciences,* 37**,** 106-17.

Battisti, D. S. & Naylor, R. L. 2009. Historical warnings of future food insecurity with unprecedented seasonal heat. *Science,* 323**,** 240-244.

Bazzini, A. A., Mongelli, V. C., Hopp, H. E., del Vas, M. & Asurmendi, S. 2007. A practical approach to the understanding and teaching of RNA silencing in plants. *Electronic Journal of Biotechnology* [Online], 10.

Bechtold, U., Albihlal, W. S., Lawson, T., Fryer, M. J., Sparrow, P. A. C., Richard, F., Persad, R., Bowden, L., Hickman, R., Martin, C., Beynon, J. L., Buchanan-Wollaston, V., Baker, N. R., Morison, J. I. L., Schoeffl, F., Ott, S. & Mullineaux, P. M. 2013. Arabidopsis HEAT SHOCK TRANSCRIPTION FACTORA1b overexpression enhances water productivity, resistance to drought, and infection. *Journal of Experimental Botany,* 64**,** 3467-3481.

Benesch, J. L. P., Aquilina, J. A., Baldwin, A. J., Rekas, A., Stengel, F., Lindner, R. A., Basha, E., Devlin, G. L., Horwitz, J., Vierling, E., Carver, J. A. & Robinson, C. V. 2010. The quaternary organization and dynamics of the molecular chaperone HSP26 are thermally regulated. *Chemistry & Biology,* 17**,** 1008-1017.

Bertani, G. 2004. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *Journal of Bacteriology,* 186**,** 595-600.

Berthomieu, P., Conejero, G., Nublat, A., Brackenbury, W. J., Lambert, C., Savio, C., Uozumi, N., Oiki, S., Yamada, K., Cellier, F., Gosti, F., Simonneau, T., Essah, P. A., Tester, M., Very, A. A., Sentenac, H. & Casse, F. 2003. Functional analysis of AtHKT1 in Arabidopsis shows that Na+ recirculation by the phloem is crucial for salt tolerance. *Embo Journal,* 22**,** 2004-2014.

Bethke, G., Pecher, P., Eschen-Lippold, L., Tsuda, K., Katagiri, F., Glazebrook, J., Scheel, D. & Lee, J. 2012. Activation of the Arabidopsis thaliana Mitogen-Activated Protein Kinase MPK11 by the Flagellin-Derived Elicitor Peptide, flg22. *Molecular Plant-Microbe Interactions,* 25**,** 471-480.

Blakeslee, J. J., Zhou, H. W., Heath, J. T., Skottke, K. R., Barrios, J. A., Liu, S. Y. & DeLong, A. 2008. Specificity of RCN1-mediated protein phosphatase 2A regulation in meristem organization and stress response in roots. *Plant Physiology,* 146**,** 539-53.

Blum, A. 2005. Drought resistance, water-use efficiency, and yield potential - are they compatible, dissonant, or mutually exclusive? *Australian Journal of Agricultural Research,* 56**,** 1159-1168.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72**,** 248-254.

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

Canella, D., Gilmour, S. J., Kuhn, L. A. & Thomashow, M. F. 2010. DNA binding by the Arabidopsis CBF1 transcription factor requires the PKKP/RAGRxKFxETRHP signature sequence. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms,* 1799**,** 454-462.

Cell, P. 2013. *Instructions for Authors* [Online]. Available: http://www.plantcell.org/site/misc/ifora.xhtml#Nomenclature and Terminology [Accessed 1st Dec 2013].

Chaves, M. M., Flexas, J. & Pinheiro, C. 2009. Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany,* 103**,** 551-560.

Chen, Q., Lauzon, L. M., DeRocher, A. E. & Vierling, E. 1990. Accumulation, stability, and localization of a major chloroplast heat-shock protein. *Journal of Cell Biology,* 110**,** 1873-83.

Cho, E. K. & Choi, Y. J. 2009. A nuclear-localized HSP70 confers thermoprotective activity and drought-stress tolerance on plants. *Biotechnology Letters,* 31**,** 597-606.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A. D., Friedlingstein, P., Grunwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T. & Valentini, R. 2005. Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature,* 437**,** 529-533.

Clark, A. R., Lubsen, N. H. & Slingsby, C. 2012. sHSP in the eye lens: Crystallin mutations, cataract and proteostasis. *International Journal of Biochemistry & Cell Biology,* 44**,** 1687-1697.

Clarke, S. M., Mur, L. A. J., Wood, J. E. & Scott, I. M. 2004. Salicylic acid dependent signaling promotes basal thermotolerance but is not essential for acquired thermotolerance in Arabidopsis thaliana. *The Plant Journal,* 38**,** 432-447.

Clough, S. J. & Bent, A. F. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J,* 16**,** 735-43.

Cosgrove, D. J. 2000. Loosening of plant cell walls by expansins. *Nature,* 407**,** 321-326.

Crafts-Brandner, S. J. & Salvucci, M. E. 2002. Sensitivity of photosynthesis in a C4 plant, maize, to heat stress. *Plant Physiology,* 129**,** 1773-1780.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W.-R. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiology,* 139**,** 5-17.

DaCosta, M. & Huang, B. R. 2006. Osmotic adjustment associated with variation in bentgrass tolerance to drought stress. *Journal of the American Society for Horticultural Science,* 131**,** 338-344.

Dafny-Yelin, M., Tzfira, T., Vainstein, A. & Adam, Z. 2008. Non-redundant functions of sHSP-CIs in acquired thermotolerance and their role in early seed development in Arabidopsis. *Plant Molecular Biology,* 67**,** 363-373.

Dai, A. G. 2011. Drought under global warming: a review. *Wiley Interdisciplinary Reviews-Climate Change,* 2**,** 45-65.

Daugaard, M., Rohde, M. & Jaattela, M. 2007. The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *Febs Letters,* 581**,** 3702-3710.

Devaiah, B. N., Nagarajan, V. K. & Raghothama, K. G. 2007. Phosphate homeostasis and root development in Arabidopsis are synchronized by the zinc finger transcription factor ZAT6. *Plant Physiology,* 145**,** 147-159.

Dill, A., Thomas, S. G., Hu, J., Steber, C. M. & Sun, T.-P. 2004. The Arabidopsis F-Box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. *Plant Cell,* 16**,** 1392-1405.

EPA, U. S. E. P. A.-U. n.d. *Overview of Greenhouse Gases: Carbon Dioxide Emissions* [Online]. United States Environmental Protection Agency. Available: http://www.epa.gov/climatechange/ghgemissions/gases/co2.html [Accessed 15th Feb 2014].

FAO, F. a. A. O. o. t. U. N.-. 2013. The Post 2015 Water Thematic Consultation - Water Resource Availability - Framing Paper. Rome, Italy: Food and Agriculture Organisation of the United Nations.

FAO, F. a. A. O. o. t. U. N.-. 2014. *Water Development and Management Unit - Topics - Water Quality* [Online]. Rome, Italy: Food and Agriculture Organisation of the United Nations. Available: http://www.fao.org/nr/water/topics_qual_marginal_water.html [Accessed 13th Feb 2014].

Feng, S. H., Martinez, C., Gusmaroli, G., Wang, Y., Zhou, J. L., Wang, F., Chen, L. Y., Yu, L., Iglesias-Pedraz, J. M., Kircher, S., Schafer, E., Fu, X. D., Fan, L. M. & Deng, X. W. 2008. Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. *Nature,* 451**,** 475-U9.

Flowers, T. J. 2004. Improving crop salt tolerance. *Journal of Experimental Botany,* 55**,** 307-319.

Fowler, S. & Thomashow, M. F. 2002. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell,* 14**,** 1675-1690.

Fujita, Y., Nakashima, K., Yoshida, T., Katagiri, T., Kidokoro, S., Kanamori, N., Umezawa, T., Fujita, M., Maruyama, K., Ishiyama, K., Kobayashi, M., Nakasone, S., Yamada, K., Ito, T., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2009. Three SnRK2 protein kinases are the main positive regulators of abscisic acid signaling in response to water stress in Arabidopsis. *Plant Cell Physiology,* 50**,** 2123-32.

Gaxiola, R. A., Rao, R., Sherman, A., Grisafi, P., Alper, S. L. & Fink, G. R. 1999. The Arabidopsis thaliana proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 1480-1485.

Gilmour, S. J., Sebolt, A. M., Salazar, M. P., Everard, J. D. & Thomashow, M. F. 2000. Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiology,* 124**,** 1854-1865.

Golldack, D., Li, C., Mohan, H. & Probst, N. 2013. Gibberellins and abscisic acid signal crosstalk: living and developing under unfavorable conditions. *Plant Cell Reports,* 32**,** 1007-1016.

Gourdji, S. M., Sibley, A. M. & Lobell, D. B. 2013. Global crop exposure to critical high temperatures in the reproductive period: historical trends and future projections. *Environmental Research Letters,* 8**,** 10.

Han, Y., Yin, S. Y. & Huang, L. 2015. Towards plant salinity tolerance-implications from ion transporters and biochemical regulation. *Plant Growth Regulation,* 76**,** 13-23.

Haslbeck, M., Franzmann, T., Weinfurtner, D. & Buchner, J. 2005. Some like it hot: the structure and function of small heat-shock proteins. *Nature Structural & Molecular Biology,* 12**,** 842-846.

Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S. & Mullineaux, P. M. 2000. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Molecular Biology,* 42**,** 819-832.

Helm, K. W., Schmeits, J. & Vierling, E. 1995. An endomembrane-localized small heat-shock protein from Arabidopsis thaliana. *Plant Physiology,* 107**,** 287-288.

Hettenhausen, C., Baldwin, I. T. & Wu, J. 2012. Silencing MPK4 in Nicotiana attenuata enhances photosynthesis and seed production but compromises abscisic acid-induced stomatal closure and guard cell-mediated resistance to Pseudomonas syringae pv tomato DC3000. *Plant Physiology,* 158**,** 759-76.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. & Zimmermann, P. 2008. Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Advances in Bioinformatics,* 2008.

Huang, B. R., Rachmilevitch, S. & Xu, J. C. 2012. Root carbon and protein metabolism associated with heat tolerance. *Journal of Experimental Botany,* 63**,** 3455-3465.

Huang, D. W., Sherman, B. T. & Lempicki, R. A. 2008. Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nature Protocols,* 4**,** 44-57.

Illumina, I. 2010. *Technology Spotlight: Illumina Sequencing*. California, USA.

Ishitani, M., Liu, J. P., Halfter, U., Kim, C. S., Shi, W. M. & Zhu, J. K. 2000. SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. *Plant Cell,* 12**,** 1667-1677.

Ito, M., Araki, S., Matsunaga, S., Itoh, T., Nishihama, R., Machida, Y., Doonan, J. H. & Watanabe, A. 2001. G2/M-phase-specific transcription during the plant cell cycle is mediated by c-Myb-like transcription factors. *Plant Cell,* 13**,** 1891-1905.

Iuchi, S., Suzuki, H., Kim, Y. C., Iuchi, A., Kuromori, T., Ueguchi-Tanaka, M., Asami, T., Yamaguchi, I., Matsuoka, M., Kobayashi, M. & Nakajima, M. 2007. Multiple loss-of-function of Arabidopsis gibberellin receptor AtGID1s completely shuts down a gibberellin signal. *Plant Journal,* 50**,** 958-66.

Jacobsen, J. V. & Shaw, D. C. 1989. Heat-stable proteins and abscisic acid action in barley aleurone cells. *Plant Physiology,* 91**,** 1520-1526.

Jaglo-Ottosen, K. R., Gilmour, S. J., Zarka, D. G., Schabenberger, O. & Thomashow, M. F. 1998. Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science,* 280**,** 104-106.

Jagtap, V., Bhargava, S., Streb, P. & Feierabend, J. 1998. Comparative effect of water, heat and light stresses on photosynthetic reactions in Sorghum bicolor (L.) Moench. *Journal of Experimental Botany,* 49**,** 1715-1721.

Jinn, T.-L., Wu, S.-H., Yeh, C.-H., Hsieh, M.-H., Yeh, Y.-C., Chen, Y.-M. & Lin, C.-Y. 1993. Immunological kinship of Class I low molecular weight heat

shock proteins and thermostabilization of soluble proteins in vitro among plants. *Plant Cell Physiol.,* 34**,** 1055-1062.

Johnston, A. J., Meier, P., Gheyselinck, J., Wuest, S. E., Federer, M., Schlagenhauf, E., Becker, J. D. & Grossniklaus, U. 2007. Genetic subtraction profiling identifies genes essential for Arabidopsis reproduction and reveals interaction between the female gametophyte and the maternal sporophyte. *Genome Biology,* 8**,** R204.

Karimi, M., Inzé, D. & Depicker, A. 2002. GATEWAY(TM) vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science,* 7**,** 193-195.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. 1999. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology,* 17**,** 287-291.

Katiyar-Agarwal, S., Agarwal, M. & Grover, A. 2003. Heat-tolerant basmati rice engineered by over-expression of hsp101. *Plant Molecular Biology,* 51**,** 677-686.

Kellermeier, F., Armengaud, P., Seditas, T. J., Danku, J., Salt, D. E. & Amtmann, A. 2014. Analysis of the root system architecture of arabidopsis provides a quantitative readout of crosstalk between nutritional signals. *Plant Cell,* 26**,** 1480-1496.

Kendrick, M. D. & Chang, C. 2008. Ethylene signaling: new levels of complexity and regulation. *Current Opinion in Plant Biology,* 11**,** 479-485.

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg-Bauer, E., Kudla, J. & Harter, K. 2007. The AtGenExpress Global Stress Expression data set: Protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal,* 50**,** 347-363.

Kim, K. H., Alam, I., Kim, Y. G., Sharmin, S. A., Lee, K. W., Lee, S. H. & Lee, B. H. 2012. Overexpression of a chloroplast-localized small heat shock protein OsHSP26 confers enhanced tolerance against oxidative and heat stresses in tall fescue. *Biotechnology Letters,* 34**,** 371-377.

Kim, K. K., Kim, R. & Kim, S.-H. 1998. Crystal structure of a small heat-shock protein. *Nature,* 394**,** 595-599.

Kim, T. H. 2014. Mechanism of ABA signal transduction: Agricultural highlights for improving drought tolerance. *Journal of Plant Biology,* 57**,** 1-8.

Klempnauer, K. H., Gonda, T. J. & Bishop, J. M. 1982. Nucleotide-sequence of the retroviral leukemia gene V-MYB and its cellular progenitor C-MYB - the architecture of a transduced oncogene. *Cell,* 31**,** 453-463.

Knowledge, I. W. o. 2012. *Journal Citation Reports* [Online]. Available: http://admin-apps.webofknowledge.com/JCR/JCR [Accessed 1st Dec 2013].

Kodaira, K. S., Qin, F., Tran, L. S., Maruyama, K., Kidokoro, S., Fujita, Y., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2011. Arabidopsis Cys2/His2 zinc-finger proteins AZF1 and AZF2 negatively regulate abscisic acid-repressive and auxin-inducible genes under abiotic stress conditions. *Plant Physiology,* 157**,** 742-56.

Kotak, S., Larkindale, J., Lee, U., von Koskull-Doring, P., Vierling, E. & Scharf, K. D. 2007a. Complexity of the heat stress response in plants. *Current Opinion in Plant Biology,* 10**,** 310-316.

Kotak, S., Vierling, E., Baumlein, H. & Koskull-Doring, P. v. 2007b. A novel transcriptional cascade regulating expression of heat stress proteins during seed development of Arabidopsis. *Plant Cell,* 19**,** 182-195.

Kumar, S. V., Lucyshyn, D., Jaeger, K. E., Alos, E., Alvey, E., Harberd, N. P. & Wigge, P. A. 2012. Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature,* 484**,** 242-U127.

Kyte, J. & Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* 157**,** 105-32.

Lambert, W., Koeck, P. J. B., Ahrman, E., Purhonen, P., Cheng, K., Elmlund, D., Hebert, H. & Emanuelsson, C. 2011. Subunit arrangement in the dodecameric chloroplast small heat shock protein Hsp21. *Protein Science,* 20**,** 291-301.

Larkindale, J. & Vierling, E. 2008. Core genome responses involved in acclimation to high temperature. *Plant Physiology,* 146**,** 748-761.

Lei, Y. B., Yin, C. Y. & Li, C. Y. 2006. Differences in some morphological, physiological, and biochemical responses to drought stress in two contrasting populations of Populus przewalskii. *Physiologia Plantarum,* 127**,** 182-191.

Lenne, C. & Douce, R. 1994. A low-molecular-mass heat-shock protein is localized to higher-plant mitochondria. *Plant Physiology,* 105**,** 1255-1261.

Li, M. W., AuYeung, W. K. & Lam, H. M. 2013. The GCN2 homologue in Arabidopsis thaliana interacts with uncharged tRNA and uses Arabidopsis eIF2alpha molecules as direct substrates. *Plant Biol (Stuttg),* 15**,** 13-8.

Li, Z., Zhang, L., Yu, Y., Quan, R., Zhang, Z., Zhang, H. & Huang, R. 2011. The ethylene response factor AtERF11 that is transcriptionally modulated by the bZIP transcription factor HY5 is a crucial repressor for ethylene biosynthesis in Arabidopsis. *Plant Journal,* 68**,** 88-99.

Lindquist, S. 1986. The heat-shock response. *Annual Review of Biochemistry,* 55**,** 1151-1191.

Liu, J. G., Qin, Q. L., Zhang, Z., Peng, R. H., Xiong, A. S., Chen, J. M. & Yao, Q. H. 2009. OsHSF7 gene in rice, Oryza sativa L., encodes a transcription factor that functions as a high temperature receptive and responsive factor. *Bmb Reports,* 42**,** 16-21.

Liu, J. P., Ishitani, M., Halfter, U., Kim, C. S. & Zhu, J. K. 2000. The Arabidopsis thaliana SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 3730-3734.

Liu, J. P. & Zhu, J. K. 1998. A calcium sensor homolog required for plant salt tolerance. *Science,* 280**,** 1943-1945.

Liu, X. M., Nguyen, X. C., Kim, K. E., Han, H. J., Yoo, J., Lee, K., Kim, M. C., Yun, D. J. & Chung, W. S. 2013. Phosphorylation of the zinc finger transcriptional regulator ZAT6 by MPK6 regulates Arabidopsis seed germination under salt and osmotic stress. *Biochem Biophys Res Commun,* 430**,** 1054-9.

Lobell, D. B., Baldos, U. L. C. & Hertel, T. W. 2013a. Climate adaptation as mitigation: the case of agricultural investments. *Environmental Research Letters,* 8**,** 12.

Lobell, D. B. & Gourdji, S. M. 2012. The influence of climate change on global crop productivity. *Plant Physiology,* 160**,** 1686-1697.

Lobell, D. B., Hammer, G. L., McLean, G., Messina, C., Roberts, M. J. & Schlenker, W. 2013b. The critical role of extreme heat for maize production in the United States. *Nature Climate Change,* 3**,** 497-501.

Lobell, D. B., Sibley, A. & Ortiz-Monasterio, J. I. 2012. Extreme heat effects on wheat senescence in India. *Nature Climate Change,* 2**,** 186-189.

Lohmann, C., Eggers-Schumacher, G., Wunderlich, M. & Schoffl, F. 2004. Two different heat shock transcription factors regulate immediate early

expression of stress genes in Arabidopsis (vol 271, pg 11, 2004). *Molecular Genetics and Genomics,* 271, 376-376.

Love, A. J., Yun, B. W., Laval, V., Loake, G. J. & Milner, J. J. 2005. Cauliflower mosaic virus, a compatible pathogen of Arabidopsis, engages three distinct defense-signaling pathways and activates rapid systemic generation of reactive oxygen species. *Plant Physiology,* 139, 935-948.

Magome, H., Yamaguchi, S., Hanada, A., Kamiya, Y. & Oda, K. 2004. dwarf and delayed-flowering 1, a novel Arabidopsis mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor. *Plant Journal,* 37, 720-729.

Mao, Y. P., Pavangadkar, K. A., Thomashow, M. F. & Triezenberg, S. J. 2006. Physical and functional interactions of Arabidopsis ADA2 transcriptional coactivator proteins with the acetyltransferase GCN5 and with the cold-induced transcription factor CBF1. *Biochimica Et Biophysica Acta-Gene Structure and Expression,* 1759, 69-79.

Martin, C. & Paz-Ares, J. 1997. MYB transcription factors in plants. *Trends in Genetics,* 13, 67-73.

Maruyama, K., Sakuma, Y., Kasuga, M., Ito, Y., Seki, M., Goda, H., Shimada, Y., Yoshida, S., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2004. Identification of cold-inducible downstream genes of the Arabidopsis DREB1A/CBF3 transcriptional factor using two microarray systems. *Plant Journal,* 38, 982-993.

McQueenmason, S., Durachko, D. M. & Cosgrove, D. J. 1992. 2 endogenous proteins that induce cell-wall extension in plants. *Plant Cell,* 4, 1425-1433.

Meinke, D. & Koornneef, M. 1997. Community Standards for Arabidopsis Genetics. *The Plant Journal,* 12, 247-253.

Mittler, R. & Blumwald, E. 2010. Genetic engineering for modern agriculture: Challenges and perspectives. *Annual Review of Plant Biology,* 61, 443-462.

Mittler, R., Kim, Y., Song, L. H., Coutu, J., Coutu, A., Ciftci-Yilmaz, S., Lee, H., Stevenson, B. & Zhu, J. K. 2006. Gain- and loss-of-function mutations in Zat10 enhance the tolerance of plants to ablotic stress. *Febs Letters,* 580, 6537-6542.

Mizoi, J., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2012. AP2/ERF family transcription factors in plant abiotic stress responses. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms,* 1819, 86-96.

Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E. & Bukau, B. 2003. Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Molecular Microbiology,* 50, 585-595.

Montero-Barrientos, M., Hermosa, R., Cardoza, R. E., Gutierrez, S., Nicolas, C. & Monte, E. 2010. Transgenic expression of the Trichoderma harzianum hsp70 gene increases Arabidopsis resistance to heat and other abiotic stresses. *Journal of Plant Physiology,* 167, 659-665.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T. & Kimura, T. 2007. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J Biosci Bioeng,* 104, 34-41.

Nakashima, K., Ito, Y. & Yamaguchi-Shinozaki, K. 2009. Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. *Plant Physiology,* 149, 88-95.

Narsai, R., Wang, C., Chen, J., Wu, J., Shou, H. & Whelan, J. 2013. Antagonistic, overlapping and distinct responses to biotic stress in rice (Oryza sativa) and interactions with abiotic stress. *Bmc Genomics,* 14.

Nelson, G. C., Rosegrant, M. W., Koo, J., Robertson, R., Sulser, T., Zhu, T., Ringler, C., Msangi, S., Palazzo, A., Batka, M., Magalhaes, M., Valmonte-Santos, R., Ewing, M. & Lee, D. 2009. Climate change: Impact on agriculture and costs of adaptation. Washington D.C.: International Food Policy Research Institute (IFPRI).

Neta-Sharir, I., Isaacson, T., Lurie, S. & Weiss, D. 2005. Dual role for tomato heat shock protein 21: Protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. *Plant Cell,* 17**,** 1829-1838.

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

Nobel, P. S. 1999. *Physicochemical & Environmental Plant Physiology,* London, UK, Academic Press.

Nurcahyanti, A. D. R. 2009. *Studies on the genetic basis for thermotolerance in Arabidopsis thaliana.* MSc, University of Glasgow.

O'Connor, T. R., Dyreson, C. & Wyrick, J. J. 2005. Athena: a resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences. *Bioinformatics,* 21**,** 4411-4413.

Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S. & Nishimura, Y. 1992. Solution structure of a DNA-binding unit of MYB - A helix turn helix-related motif with conserved tryptophans forming a hydrophobic core. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 6428-6432.

Ogawa, D., Yamaguchi, K. & Nishiuchi, T. 2007. High-level overexpression of the Arabidopsis HsfA2 gene confers not only increased themotolerance but also salt/osmotic stress tolerance and enhanced callus growth. *Journal of Experimental Botany,* 58**,** 3373-3383.

Ono, K., Hibino, T., Kohinata, T., Suzuki, S., Tanaka, Y., Nakamura, T. & Takabe, T. 2001. Overexpression of DnaK from a halotolerant cyanobacterium Aphanothece halophytica enhances the high-temperature tolerance of tobacco during germination and early growth. *Plant Science,* 160**,** 455-461.

Osteryoung, K. W. & Vierling, E. 1994. Dynamics of small heat-shock protein distribution within the chloroplasts of higher-plants. *Journal of Biological Chemistry,* 269**,** 28676-28682.

Painter, A. J., Jaya, N., Basha, E., Vierling, E., Robinson, C. V. & Benesch, J. L. P. 2008. Real-time monitoring of protein complexes reveals their quaternary organization and dynamics. *Chemistry & Biology,* 15**,** 246-253.

Panchuk, II, Volkov, R. A. & Schoffl, F. 2002. Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in Arabidopsis. *Plant Physiology,* 129**,** 838-853.

Pavlova, E. L., Rikhvanov, E. G., Tauson, E. L., Varakina, N. N., Gamburg, K. Z., Rusaleva, T. M., Borovskii, G. B. & Voinikov, V. K. 2009. Effect of salicylic acid on the development of induced Thermotolerance and induction of heat shock protein synthesis in the Arabidopsis thaliana cell culture. *Russian Journal of Plant Physiology,* 56**,** 68-73.

Pazour, G. J., Ta, C. N. & Das, A. 1992. Constitutive mutations of Agrobacterium tumefaciens transcriptional activator virG. *J Bacteriol,* 174**,** 4169-74.

Persak, H. & Pitzschke, A. 2013. Tight Interconnection and Multi-Level Control of Arabidopsis MYB44 in MAPK Cascade Signalling. *PLOS ONE,* 8.

Pieterse, C. M. J., Leon-Reyes, A., Van der Ent, S. & Van Wees, S. C. M. 2009. Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology,* 5**,** 308-316.

Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A. & Van Wees, S. C. M. 2012. Hormonal modulation of plant immunity. *Annual Review of Cell and Developmental Biology,* 28**,** 489-521.

Pitzschke, A., Djamei, A., Teige, M. & Hirt, H. 2009. VIP1 response elements mediate mitogen-activated protein kinase 3-induced stress gene expression. *PNAS,* 106**,** 18414-18419.

Pongratz, J., Lobell, D. B., Cao, L. & Caldeira, K. 2012. Crop yields in a geoengineered climate. *Nature Climate Change,* 2**,** 101-105.

Price, J. 2005. *An Investigation into Halotolerance Mechanisms in Arabidopsis thaliana.* Ph.D, University of Glasgow.

Price, J. & Ramsay, S. Unpublished work.

Qi, Y. C., Wang, H. J., Zou, Y., Liu, C., Liu, Y. Q., Wang, Y. & Zhang, W. 2011. Over-expression of mitochondrial heat shock protein 70 suppresses programmed cell death in rice. *Febs Letters,* 585**,** 231-239.

Queitsch, C., Hong, S. W., Vierling, E. & Lindquist, S. 2000. Heat shock protein 101 plays a crucial role in thermotolerance in arabidopsis. *Plant Cell,* 12**,** 479-492.

Rabiger, D. S. & Drews, G. N. 2013. MYB64 and MYB119 are required for cellularization and differentiation during female gametogenesis in Arabidopsis thaliana. *PLOS Genetics,* 9.

Ramegowda, V. & Senthil-Kumar, M. 2015. The interactive effects of simultaneous biotic and abiotic stresses on plants: Mechanistic understanding from drought and pathogen combination. *Journal of Plant Physiology,* 176**,** 47-54.

Rhoads, D. M., White, S. J., Zhou, Y., Muralidharan, M. & Elthon, T. E. 2005. Altered gene expression in plants with constitutive expression of a mitochondrial small heat shock protein suggests the involvement of retrograde regulation in the heat stress response. *Physiologia Plantarum,* 123**,** 435-444.

Romero, I., Fuertes, A., Benito, M. J., Malpica, J. M., Leyva, A. & Paz-Ares, J. 1998. More than 80 R2R3-MYB regulatory genes in the genome of Arabidopsis thaliana. *Plant J.,* 14**,** 273-84.

Rosegrant, M. W., Koo, J., Cenacchi, N., Ringler, C., Robertson, R., Fisher, M., Cox, C., Garrett, K., Perez, N. D. & Sabbagh, P. 2014. Food security in a world of natural resource scarcity: The role of agricultural technologies. Washington D.C.: International Food Policy Research Institute (IFPRI).

Sakamoto, H., Maruyama, K., Sakuma, Y., Meshi, T., Iwabuchi, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2004. Arabidopsis Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiology,* 136**,** 2734-2746.

Sakuma, Y., Maruyama, K., Qin, F., Osakabe, Y., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2006. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 18822-18827.

Sambrook, J. & Russell, D. W. 2001. *Molecular Cloning: A Laboratory Manual, Third Edition,* Cold Spring Harbor, New York, USA, Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America,* 74**,** 5463-5467.

Sattelmach, B. & Horst, W. J. 2007. *The apoplast of higher plants: Compartment of storage, transport and reactions,* Dordrecht, The Netherlands, Springer.

Schaffer, R., Landgraf, J., Accerbi, M., Simon, V., Larson, M. & Wisman, E. 2001. Microarray analysis of diurnal and circadian-regulated genes in Arabidopsis. *Plant Cell,* 13**,** 113-123.

Scharf, K. D., Siddique, M. & Vierling, E. 2001. The expanding family of Arabidopsis thaliana small heat stress proteins and a new family of proteins containing alpha-crystallin domains (Acd proteins). *Cell Stress & Chaperones,* 6**,** 225-37.

Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods,* 9**,** 671-675.

Scholl, R. L., May, S. T. & Ware, D. H. 2000. Seed and molecular resources for Arabidopsis. *Plant Physiology,* 124**,** 1477-80.

Schramm, F., Ganguli, A., Kiehlmann, E., Englich, G., Walch, D. & 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 Molecular Biology,* 60**,** 759-772.

Schweighofer, A., Hirt, H. & Meskiene, L. 2004. Plant PP2C phosphatases: emerging functions in stress signaling. *Trends in Plant Science,* 9**,** 236-243.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y. & Shinozaki, K. 2001. Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell,* 13**,** 61-72.

Seo, P. J. & Park, C.-M. 2009. Auxin homeostasis during lateral root development under drought condition. *Plant signaling & behavior,* 4**,** 1002-4.

Seo, P. J. & Park, C.-M. 2010. MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in Arabidopsis. *New Phytologist,* 186**,** 471-483.

Shi, H. Z., Ishitani, M., Kim, C. S. & Zhu, J. K. 2000. The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na+/H+ antiporter. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 6896-6901.

Siddique, M., Gernhard, S., von Koskull-Doering, P., Vierling, E. & Scharf, K.-D. 2008. The plant sHSP superfamily: five new members in Arabidopsis thaliana with unexpected properties. *Cell Stress & Chaperones,* 13**,** 183-197.

Skirycz, A. & Inze, D. 2010. More from less: plant growth under limited water. *Current Opinion in Biotechnology,* 21**,** 197-203.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. & Hood, L. E. 1986. Fluorescence Detection in Automated DNA Sequence Analysis. *Nature,* 321**,** 674-679.

Sobott, F., Benesch, J. L. P., Vierling, E. & Robinson, C. V. 2002. Subunit exchange of multimeric protein complexes - Real-time monitoring of subunit exchange between small heat shock proteins by using electrospray mass spectrometry. *Journal of Biological Chemistry,* 277**,** 38921-38929.

Song, L. L., Jiang, Y. L., Zhao, H. Q. & Hou, M. F. 2012. Acquired thermotolerance in plants. *Plant Cell Tissue and Organ Culture,* 111**,** 265-276.

Sreenivasulu, N., Harshavardhan, V. T., Govind, G., Seiler, C. & Kohli, A. 2012. Contrapuntal role of ABA: Does it mediate stress tolerance or plant growth retardation under long-term drought stress? *Gene,* 506**,** 265-273.

State, U. S. D. o. 2014. 2014 U.S. Climate Action Report to the UN Framework Convention on Climate Change. *In:* Affairs, B. o. O. a. I. E. a. S. (ed.). Washington D.C.

Stengel, F., Baldwin, A. J., Painter, A. J., Jaya, N., Basha, E., Kay, L. E., Vierling, E., Robinson, C. V. & Benesch, J. L. P. 2010. Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proceedings of the National Academy of Sciences,* 107**,** 2007-2012.

Stone, P. 2001. The effects of heat stress on cereal yield and quality. *In:* Basra, A. S. (ed.) *Crop responses and adaptations to temperature stress.* Binghamton, N.Y.: Food Products Press.

Stracke, R., Werber, M. & Weisshaar, B. 2001. The R2R3-MYB gene family in Arabidopsis thaliana. *Current Opinion in Plant Biology,* 4**,** 447-456.

Sun, L. P., Liu, Y., Kong, X. P., Zhang, D., Pan, J. W., Zhou, Y., Wang, L., Li, D. Q. & Yang, X. H. 2012. ZmHSP16.9, a cytosolic class I small heat shock protein in maize (Zea mays), confers heat tolerance in transgenic tobacco. *Plant Cell Reports,* 31**,** 1473-1484.

Sun, Y. & MacRae, T. H. 2005. Small heat shock proteins: molecular structure and chaperone function. *Cellular and Molecular Life Sciences,* 62**,** 2460-2476.

Suzuki, N., Rizhsky, L., Liang, H. J., Shuman, J., Shulaev, V. & Mittler, R. 2005. Enhanced tolerance to environmental stress in transgenic plants expressing the transcriptional coactivator multiprotein bridging factor 1c. *Plant Physiology,* 139**,** 1313-1322.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P. & Huala, E. 2008. The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Research,* 36**,** D1009-14.

TAIR. *The Arabidopsis Information Resource (TAIR) [www.arabidopsis.org]* [Online].

TAIR. 2007. *TAIR Nomenclature Guidelines* [Online]. Available: http://www.arabidopsis.org/portals/nomenclature/guidelines.jsp [Accessed 1st Dec 2013].

Teixeira, E. I., Fischer, G., van Velthuizen, H., Walter, C. & Ewert, F. 2013. Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology,* 170**,** 206-215.

Todaka, D., Nakashima, K., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2012. Toward understanding transcriptional regulatory networks in abiotic stress responses and tolerance in rice. *Rice,* 5.

Tonsor, S. J., Scott, C., Boumaza, I., Liss, T. R., Brodsky, J. L. & Vierling, E. 2008. Heat shock protein 101 effects in A. thaliana: genetic variation, fitness and pleiotropy in controlled temperature conditions. *Mol Ecol,* 17**,** 1614-26.

Uchida, A., Hibino, T., Shimada, T., Saigusa, M., Takabe, T., Araki, E. & Kajita, H. 2008. Overexpression of DnaK chaperone from a halotolerant

cyanobacterium Aphanothece halophytica increases seed yield in rice and tobacco. *Plant Biotechnology,* 25**,** 141-150.

UNEP, U. N. E. P.-. 2014. *UNEP - Climate Change - Home* [Online]. Available: http://www.unep.org/climatechange/ [Accessed 13th Feb 2014].

Urao, T., Yamaguchishinozaki, K., Urao, S. & Shinozaki, K. 1993. An Arabidopsis MYB homolog is induced by dehydration stress and its gene-product binds to the conserved MYB recognition sequence. *Plant Cell,* 5**,** 1529-1539.

Vacca, R. A., de Pinto, M. C., Valenti, D., Passarella, S., Marra, E. & De Gara, L. 2004. Production of reactive oxygen species, alteration of cytosolic ascorbate peroxidase, and impairment of mitochondrial metabolism are early events in heat shock-induced programmed cell death in tobacco bright-yellow 2 cells. *Plant Physiology,* 134**,** 1100-1112.

Vain, P., Harvey, A., Worland, B., Ross, S., Snape, J. & Lonsdale, D. 2004. The effect of additional virulence genes on transformation efficiency, transgene integration and expression in rice plants using the pGreen/pSoup dual binary vector system. *Transgenic Research,* 13**,** 593-603.

van Montfort, R. L. M., Basha, E., Friedrich, K. L., Slingsby, C. & Vierling, E. 2001. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nature Structural Biology,* 8**,** 1025-1030.

Vierling, E. 1991. The roles of heat shock proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology,* 42**,** 579-620.

Vierling, E. 2009. *RE: Personal communication.*

Vierling, E., Mishkind, M. L., Schmidt, G. W. & Key, J. L. 1986. Specific heat shock proteins are transported into chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America,* 83**,** 361-365.

Volaire, F., Barkaoui, K. & Norton, M. 2014. Designing resilient and sustainable grasslands for a drier future: Adaptive strategies, functional traits and biotic interactions. *European Journal of Agronomy,* 52**,** 81-89.

Volko, S. M., Boller, T. & Ausubel, F. M. 1998. Isolation of new Arabidopsis mutants with enhanced disease susceptibility to Pseudomonas syringae by direct screening. *Genetics,* 149**,** 537-548.

Ward, J. M. 2001. Identification of novel families of membrane proteins from the model plant Arabidopsis thaliana. *Bioinformatics,* 17**,** 560-563.

Waters, E. R. 2013. The evolution, function, structure, and expression of the plant sHSPs. *Journal of Experimental Botany,* 64**,** 391-403.

Waters, E. R., Aevermann, B. D. & Sanders-Reed, Z. 2008. Comparative analysis of the small heat shock proteins in three angiosperm genomes identifies new subfamilies and reveals diverse evolutionary patterns. *Cell Stress & Chaperones,* 13**,** 127-142.

Waters, E. R., Lee, G. J. & Vierling, E. 1996. Evolution, structure and function of the small heat shock. *Journal of Experimental Botany,* 47**,** 325-338.

Waters, E. R. & Rioflorido, I. 2007. Evolutionary analysis of the small heat shock proteins in five complete algal genomes. *Journal of Molecular Evolution,* 65**,** 162-174.

Waters, E. R. & Vierling, E. 1999. The diversification of plant cytosolic small heat shock proteins preceded the divergence of mosses. *Molecular Biology and Evolution,* 16**,** 127-139.

Webb, S. 2008. *Prediction of translationally regulated stress response proteins in Arabidopsis thaliana.* MSc, University of Glasgow.

Weigel, D., Ahn, J. H., Blazquez, M. A., Borevitz, J. O., Christensen, S. K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E. J., Neff,

M. M., Nguyen, J. T., Sato, S., Wang, Z. Y., Xia, Y., Dixon, R. A., Harrison, M. J., Lamb, C. J., Yanofsky, M. F. & Chory, J. 2000. Activation tagging in Arabidopsis. *Plant Physiol,* 122**,** 1003-13.

Wilkinson, S. & Davies, W. J. 2010. Drought, ozone, ABA and ethylene: new insights from cell to plant to community. *Plant Cell and Environment,* 33**,** 510-525.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V. & Provart, N. J. 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE,* 2**,** e718.

Wu, S. J., Ding, L. & Zhu, J. K. 1996. SOS1, a genetic locus essential for salt tolerance and potassium acquisition. *Plant Cell,* 8**,** 617-627.

Wuest, S. E., Vijverberg, K., Schmidt, A., Weiss, M., Gheyselinck, J., Lohr, M., Wellmer, F., Rahnenfuhrer, J., von Mering, C. & Grossniklaus, U. 2010. Arabidopsis female gametophyte gene expression map reveals similarities between plant and animal gametes. *Current Biology,* 20**,** 506-512.

Wynne, C. 2012. *An investigation into the cold-stress response of various Arabidopsis small heat shock protein gene knockouts.* BSc Undergraduate Thesis, University of Glasgow.

Xin, H. B., Zhang, H., Chen, L., Li, X. X., Lian, Q. L., Yuan, X., Hu, X. Y., Cao, L., He, X. L. & Yi, M. F. 2010. Cloning and characterization of HsfA2 from Lily (Lilium longiflorum). *Plant Cell Reports,* 29**,** 875-885.

Xiong, L. M., Wang, R. G., Mao, G. H. & Koczan, J. M. 2006. Identification of drought tolerance determinants by genetic analysis of root response to drought stress and abscisic acid. *Plant Physiology,* 142**,** 1065-1074.

Xue, G. P., McIntyre, C. L., Jenkins, C. L. D., Glassop, D., van Herwaarden, A. F. & Shorter, R. 2008. Molecular dissection of variation in carbohydrate metabolism related to water-soluble carbohydrate accumulation in stems of wheat. *Plant Physiology,* 146**,** 441-454.

Xue, Y., Peng, R., Xiong, A., Li, X., Zha, D. & Yao, Q. 2010. Over-expression of heat shock protein gene hsp26 in Arabidopsis thaliana enhances heat tolerance. *Biologia Plantarum,* 54**,** 105-111.

Yang, Q., Chen, Z.-Z., Zhou, X.-F., Yin, H.-B., Li, X., Xin, X.-F., Hong, X.-H., Zhu, J.-K. & Gong, Z. 2009. Overexpression of SOS (Salt Overly Sensitive) genes increases salt tolerance in transgenic Arabidopsis. *Molecular Plant,* 2**,** 22-31.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G. & Li-Jia, Q. 2013. The MYB transcription factor superfamily of Arabidopsis: Expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology,* 60**,** 107-124.

Yeh, C. H., Chang, P. F. L., Yeh, K. W., Lin, W. C., Chen, Y. M. & Lin, C. Y. 1997. Expression of a gene encoding a 16.9-kDa heat-shock protein, Oshsp16.9, in Escherichia coli enhances thermotolerance. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 10967-10972.

Yoshida, T., Sakuma, Y., Todaka, D., Maruyama, K., Qin, F., Mizoi, J., Kidokoro, S., Fujita, Y., Shinozaki, K. & Yamaguchi-Shinozaki, K. 2008. Functional analysis of an Arabidopsis heat-shock transcription factor HsfA3 in the transcriptional cascade downstream of the DREB2A stress-regulatory system. *Biochemical and Biophysical Research Communications,* 368**,** 515-521.

Young, J. C., Krysan, P. J. & Sussman, M. R. 2001. Efficient screening of Arabidopsis T-DNA insertion lines using degenerate primers. *Plant Physiology,* 125**,** 513 - 518.

Young, J. W., Locke, J. C. W. & Elowitz, M. B. 2013. Rate of environmental change determines stress response specificity. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 4140-4145.

Yu, E. Y., Kim, S. E., Kim, J. H., Ko, J. H., Cho, M. H. & Chung, I. K. 2000. Sequence-specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1. *Journal of Biological Chemistry,* 275**,** 24208-24214.

Zentella, R., Zhang, Z.-L., Park, M., Thomas, S. G., Endo, A., Murase, K., Fleet, C. M., Jikumaru, Y., Nambara, E., Kamiya, Y. & Sun, T.-P. 2007. Global analysis of DELLA direct targets in early gibberellin signaling in Arabidopsis. *Plant Cell,* 19**,** 3037-3057.

Zhang, J., Li, X., He, Z., Zhao, X., Wang, Q., Zhou, B., Yu, D., Huang, X., Tang, D., Guo, X. & Liu, X. 2013. Molecular character of a phosphatase 2C (PP2C) gene relation to stress tolerance in Arabidopsis thaliana. *Mol Biol Rep,* 40**,** 2633-44.

Zhang, J. L. & Shi, H. 2013. Physiological and molecular mechanisms of plant salt tolerance. *Photosynth Res,* 115**,** 1-22.

Zhao, S. Z. & Qi, X. Q. 2008. Signaling in plant disease resistance and symbiosis. *Journal of Integrative Plant Biology,* 50**,** 799-807.

Zhou, Y. L., Chen, H. H., Chu, P., Li, Y., Tan, B., Ding, Y., Tsang, E. W. T., Jiang, L. W., Wu, K. Q. & Huang, S. Z. 2012. NnHSP17.5, a cytosolic class II small heat shock protein gene from Nelumbo nucifera, contributes to seed germination vigor and seedling thermotolerance in transgenic Arabidopsis. *Plant Cell Reports,* 31**,** 379-389.

Zhu, J.-K. 2001. Plant salt tolerance. *Trends in Plant Science,* 6**,** 66-71.

Zhu, Y., Wang, Z., Jing, Y. J., Wang, L. L., Liu, X., Liu, Y. X. & Deng, X. 2009. Ectopic over-expression of BhHsf1, a heat shock factor from the resurrection plant Boea hygrometrica, leads to increased thermotolerance and retarded growth in transgenic Arabidopsis and tobacco. *Plant Molecular Biology,* 71**,** 451-467.

Ziska, L. H., Blumenthal, D. M., Runion, G. B., Hunt, E. R. & Diaz-Soltero, H. 2011. Invasive species and climate change: an agronomic perspective. *Climatic Change,* 105**,** 13-42.

Zolla, G., Heimer, Y. M. & Barak, S. 2010. Mild salinity stimulates a stress-induced morphogenic response in Arabidopsis thaliana roots. *Journal of Experimental Botany,* 61**,** 211-224.

Zupanska, A. K., Denison, F. C., Ferl, R. J. & Paul, A. L. 2013. Spaceflight engages heat shock protein and other molecular chaperone genes in tissue culture cells of *Arabidopsis thaliana*. *American Journal of Botany,* 100**,** 235-248.