Identification of Genetic Markers Associated with Chronic Moderate Hyperosmotic Stress Responses in Spring Varieties of *Triticum aestivum*

Ellis Harrison Smith

Masters by Research Project

Department of Biology

2019

Project Supervisor: Dr Andrea Harper

Abstract:

Triticum aestivum (bread wheat) contributes approximately a fifth of global daily protein and calorie intake, and annually is cultivated on an area of 220 million ha, being at present the most extensively grown stable food crop worldwide (Shiferaw, 2013; FAO, 2017). Abiotic stresses such as drought and heat can dramatically reduce bread wheat productivity by impeding photosynthetic and metabolic processes necessary for optimal growth and development (Gregorová, 2015; Zhang, 2017). Wheat yields have been shown to be suffering greater decreases as a result of drought and heat stress incidents over the last decade, compared to several of the last previous decades (Lesk, 2016).

Associative Transcriptomics (AT) is a statistical genomics method which identifies associations between phenotypic trait data, and sequence variation (single nucleotide polymorphisms) and transcript abundance (gene expression markers) data. Here, 135 *Triticum aestivum* landraces, belonging to the YoGI biodiversity panel (Harper and Bancroft, unpublished data), were phenotyped for traits measuring biomass accumulation in response to hyperosmotic stress. Samples were cultured hydroponically, and treated with 7.5% polyethylene glycol (PEG) 8000. An AT analysis was conducted utilising RNA-seq derived transcriptomic data taken from the YoGI panel, and the phenotypic trait data generated in this study.

The AT analysis identified 21 gene expression markers, and sequence variation of one genomic region, as significantly associating with measured variation in fresh root biomass. TraesCS4B01G319100.1, identified by NCBI pBLAST searches and the latest IWGSC annotation information as a putative heat shock protein 40 (HSP40) gene. AT findings were validated through a morphological analysis experiment of a *Triticum durum* mutant (Kronos4610) purported to have a stop codon mutation within a TraesCS4B01G319100.1 homolog. Future work, utilising genetic markers identified by the AT analysis, is suggested. Associative Transcriptomics is here demonstrated as a novel method for the investigation of plant abiotic stress responses such as hyperosmotic stress.

List of Contents

Abstract ……. (Page 1)

List of contents ……. (Page 2)

List of tables ……. (Page 6)

List of figures ……. (Page 7)

List of abbreviations ……. (Page 10)

List of accompanying material ……. (Page 11)

Declaration ……. (Page 11)

1. Introduction ……. (Page 12)

1.1 Changes to global climate conditions and their impact on food security ……. (Page 12)

1.2 Effects of drought and hyperosmotic stress on plants ……. (Page 13)

1.3 Plant responses to drought and hyperosmotic stress ……. (Page 17)

1.3.1 Hyperosmotic stress signaling ……. (Page 19)

1.3.2 Regulation of genes differentially expressed in response to hyperosmotic stress ……. (Page 23)

1.3.3 Hyperosmotic stress-induced reprogramming of cellular metabolism ……. (Page 24)

1.3.4 Hyperosmotic stress-induced production of reactive oxygen species ..... (Page 25)

1.3.5 Regulation of photosynthetic metabolism in response to hyperosmotic stress ……. (Page 26)

1.3.6 Regulation of respiration in response to hyperosmotic stress ……. (Page 28)

1.3.7 Repartitioning of photoassimilate ……. (Page 29)

1.3.8 Drought escape response …… (Page 30)

1.4 Predicted effects of drought on agricultural yields …….. (Page 32)

1.5 *Triticum aestivum* and the identification of drought resistance genetic markers …….. (Page 33)

1.5.1 Evolution, domestication and the *Triticum aestivum* genome ……..….. (Page 33)

1.5.2 Landraces: Genotypic and phenotypic variation ……..…..…. (Page 35)

1.5.3 Drought resistant markers identified in *Triticum aestivum* …….……..… (Page 37)

1.6 Genome Wide Association Studies and Associative Transcriptomics ……. (Page 39)

1.7 Specific project aims …… (Page 41)

2. Methods ……..... (Page 42)

2.1 Seed sterilization, germination and culturing ……..…… (Page 42)

2.1.1 *Triticum aestivum* sterilization and germination ……...… (Page 42)

2.1.2 *Arabidopsis thaliana* sterilization, germination and soil culture …..….. (Page 43)

2.2 Hydroponic culture and polyethylene glycol treatment …….……. (Page 44)

2.3 Landrace box position allocation …….. (Page 45)

2.4 Biomass trait measurement …….…….. (Page 46)

2.5 Associative transcriptomics ……... (Page 47)

2.6 Arabidopsis thaliana genotyping ……..……. (Page 48)

2.6.1 DNA extraction ……..…… (Page 48)

2.6.2 Polymerase chain reaction (PCR) conditions ……..……. (Page 49)

2.7 Root architecture measurements ……..…… (Page 50)

2.8 Statistical analysis ……...…… (Page 51)

3. Pilot experiments ……. (Page 52)

3.1 Introduction …….. (Page 52)

3.2 Methods …… (Page 54)

3.2.1 Pilot 1 …… (Page 54)

3.2.2 Pilot 2 …… (Page 55)

3.3 Results …….. (Page 56)

3.3.1 Pilot 1 …… (Page 56)

3.3.2 Pilot 2 …… (Page 59)

3.4 Discussion …… (Page 60)

4. Phenotyping and Associative Transcriptomics ……. (Page 64)

4.1 Introduction …….. (Page 64)

4.2 Methods …… (Page 65)

4.2.1 Osmotic stress tolerance phenotyping…….. (Page 65)

4.2.2 SNP analysis …… (Page 67)

4.2.3 GEM analysis …… (Page 68)

4.3 Results …….. (Page 69)

4.3.1 Osmotic stress tolerance phenotyping…….. (Page 70)

4.3.2 SNP marker analysis ……... (Page 74)

4.3.3 GEM marker analysis …… (Page 76)

4.4 Discussion …… (Page 79)

4.4.1 Osmotic stress tolerance phenotyping…….. (Page 79)

4.4.2 SNP marker analysis ……... (Page 80)

4.4.3 GEM marker analysis …… (Page 83)

5. Candidate gene analysis and validation …… (Page 87)

5.1 Introduction …… (Page 87)

5.2 Methods ……..…….. (Page 89)

5.3 Results ……..…… (Page 91)

5.4 Discussion ……..…….. (Page 93)

6. Discussion ……..……..… (Page 95)

References ……..…..……. (Page 99)

Appendix ……..……... (Page 122)

Acknowledgements ……..……. (Page 141)

List of tables:

Table 4.1: GAPIT results for the two SNP markers contained within the association peak region used to identify the aquaporin candidates …….…….. (Page 75)

Supplementary Table 1: Information for the 135 YoGI panel landraces that were phenotyped in this study …….……. (Page 123)

Supplementary Table 2: Top NCBI protein sequence identity matches for TraesCS6A01G405600.1 without taxa specification (Aegilops tauschii/ Tritcum urartu) and with Arabidopsis taxa specification (Arabidopsis thaliana) …….……... (Page 131)

Supplementary Table 3: Twenty-one gene expression markers from Regress_multi analysis of NCFRB data with FDR p-values lower than 0.05 …….……... (Page 131)

Supplementary Table 4: List of designed primer sequences used to amplify PIP1;5 DNA extracted from SALK lines for genotyping …….……..…… (Page 132)

Supplementary Table 5: List of designed primer sequences used to amplify a section of the Kronos4610 TraesCS4B01G319100.1 homolog …….……. (Page 132)

Supplementary Table 6: Showing results of Tukey multiple comparisons of means 95% family-wise confidence level test for ANOVA test completed, for which results are presented in Supplementary Figure 9 ….…….. (Page 138)

List of figures:

Figure 1.1: Schematic depicting the sequential cellular actions involved in hyperosmotic stress signal detection and the downstream coordination of responses…… (Page 18)

Figure 1.2: Schematic depicting the phylogenetic history of Triticum aestivum ….. (Page 34)

Figure 1.3: Global map of wheat mega-environments (Sonder and Kai, 2016) ..… (Page 37)

Figure 3.1: Pilot 1 Box plots ……. (Page 57)

Figure 3.2: Pilot 1; Normalised change in total dry biomass of landraces …….. (Page 58)

Figure 3.3: Pilot 2 plots …… (Page 59)

Figure 4.1. Box plots comparing mean total dry biomass of box groups ……. (Page 70)

Figure 4.2. Plots comparing mean dry root fraction of box groups ……... (Page 71)

Figure 4.3. Plots comparing mean total dry biomass and dry root fraction of control and PEG treated samples from phenotyping experiment ……. (Page 72)

Figure 4.4: Photograph images showing the cessation of primary root growth of YoGI 310 …… (Page 73)

Figure 4.5: Plot showing the normalised change in mean dry total biomass for 135 T. aestivum YoGI landrace lines grown in phenotyping experiment ……..… (Page 74)

Figure 4.6: Manhattan plots depicting SNP marker trait association test -log10 p-values for normalised change in fresh root biomass trait data ……. (Page 75)

Figure 4.7: Manhattan plot depicting GEM marker trait association test -log10 p-values for normalised change in fresh root biomass trait data …….…….. (Page 77)

Figure 4.8: Scatter plots showing the relationship between RPKM and normalised change in fresh root biomass for TraesCS4B01G319100.1 for the 135 landraces measured in the panel screen experiment ……...…….. (Page 78)

Figure 5.1: Plot comparing normalised change in fresh root biomass (NCFRB) of Kronos parental, Cadenza parental and Kronos4610 ……..…… (Page 87)

Figure 5.2: Plot comparing root volume (cm3) of Kronos parental, Cadenza parental and Kronos4610 …….……..…….. (Page 89)

Supplementary Figure 1: Schematic depicting the landrace position allocations of hydroponic samples in the Latin square design used in the PEG6000 and PEG8000 pilot experiments …….…… (Page 122)

Supplementary Figure 2: Schematic depicting the position of Cadenza and Kronos lines in hydroponics in the morphological analysis experiments ……..…… (Page 122)

Supplementary Figure 3: Image depicting the microbial growth typical of the polyethylene glycol 8000 (PEG8000) treatment boxes …….. (Page 122)

Supplementary Figure 4: Manhattan plots depicting SNP marker trait association test -log10 p-values for normalised change in fresh shoot biomass (A), fresh total biomass (B), dry root biomass (C), dry shoot biomass (D), total dry biomass (E), fresh root fraction (F) and dry root fraction (G) trait data …… (Page 127)

Supplementary Figure 5: Manhattan plots depicting GEM marker trait association

test -log10 p-values for normalised change in fresh shoot biomass (A), fresh total biomass (B), dry root biomass (C), dry shoot biomass (D), total dry biomass (E), fresh root fraction (F) and dry root fraction (G) trait data …….…….. (Page 129)

Supplementary Figure 6: Scatter plots showing the relationship between RPKM and NCFRB for the eight candidate genes investigated for the 134 landraces measured in the panel screen experiment ……..…… (Page 133)

Supplementary Figure 7: Scatter plots showing the relationship between RPKM and NCFRB for the eight candidate genes investigated for the 134 landraces measured in the panel screen experiment (YoGI 28 removed) …… (Page 135)

Supplementary Figure 8: Plot comparing root surface area (cm2) of Kronos parental, Cadenza parental and Kronos4610 ……..… (Page 137)

Supplementary Figure 9: Plot comparing average root diameter (mm) of Kronos parental, Cadenza parental and Kronos4610 …….…… (Page 138)

R Script 1: Script used to calculate landrace mean trait values for PEG and control treatment ……..…… (Page 140)

List of abbreviations:

Abscisic acid (ABA)

Adenosine triphosphate (ATP)

Associative Transcriptomics (AT)

Cyclin dependent protein kinase (CDPK9)

DNA extraction buffer (DEB)

Drought escape (DE)

False discovery rate (FDR)

Flowering locus T (FT)

Gene expression marker (GEM)

Genome wide association study (GWAS)

Heading date 3A (HD3A)

Heat shock protein 40 (HSP40)

Intergovernmental Panel on Climate Change (IPCC)

International Plant Genetic Resources Institute (IPGRI)

International wheat genome sequencing consortium (IWGSC)

Murashige and Skoog (MS)

Non-structural carbohydrates (NSCs)

Plasma membrane intrinsic protein (PIP)

Polyethylene glycol (PEG)

Quantitative trait loci (QTL)

Reactive oxygen species (ROS)

Root apical meristem (RAM)

Root fraction (RF)

Single nucleotide polymorphism (SNP)

Soil plant atmosphere continuum (SPAC)

Systemic acquired acclimation (SAA)

Targeted induced local lesions in genome (TILLING)

United Nations (UN)

Vapour pressure deficit (VPD)

List of accompanying material:

Pilot experiment 1 fresh biomass results

Pilot experiment 1 dry biomass results

Pilot experiment 2 fresh biomass results

Pilot experiment 2 dry biomass results

Phenotyping experiment raw data

Phenotyping experiment calculated landrace trait data

Phenotyping experiment control sample trait data

Phenotyping experiment PEG sample trait data

Random Allocations of Sample Positions for AT Experiment

Associative Transcriptomics SNP analysis results files

Associative Transcriptomics GEM analysis results files

Validation experiment biomass results

Validation experiment root imaging results

'HSP40' Kronos Genotyping Sequencing Results

Declaration:

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

1. Introduction

1.1 Changes to global climate conditions and their impact on food security

Constructing and maintaining efficient food systems that succeed in both producing and distributing food stuffs effectively, so as to prevent hunger and malnutrition, at present remains an ongoing challenge for all countries. In 2018, a consortium of United Nations working groups published their annual State of Food Security and Nutrition in the World report (FAO, IFAD, UNICEF, WFP and WHO, 2018). The report details the current state of food intake globally, emphasizing the nature of existing problems and their estimated causes based on a working contemporary global population data set compiled by combining available country level data (FAO, 2017). Importantly, the report highlights estimations of increases in global hunger for a third consecutive year and draws much attention to numerous reported risks posed to global food security, underscoring climate change, population increases and economic and military conflicts as major contributors (UN, 2017; FAO, 2016).

Climate change particularly is widely recognised as a serious threat to global agriculture (Li et al., 2009; Schmidhuber and Tubiello, 2007; Wheeler and von Braun, 2013). The Intergovernmental Panel on Climate Change (IPCC) has described ‘robust evidence’ for an increase in the proportion of populations expected to be affected by climactic extremes such as drought, heat stress, storms, extreme precipitation and flooding, as a consequence of changes in global climate in the 21st century (IPCC, 2014), concurrent with other independent findings (Fedoroff et al., 2010). Currently, research efforts in the climate science community are shifting a focus from attributing changes in individual meteorological variables to human activity, to producing general circulation models,

capable of attributing changes in the overall risk of extreme climactic events to anthropogenic influences, with some success (Otto, 2016; National Academies of Sciences, Engineering and Medicine, 2016; Shepard, 2016).

The outcome of annual crop yields represents a crucial factor in mitigating global malnutrition, requiring a continuation of research efforts aimed at improving the efficacy of end-use quality and crop yield parameters. Drought particularly, severely reduces agricultural productivity and yield (Qaseem et al., 2019).

1.2 Effects of drought and hyperosmotic stress on plants

Drought is a meteorological term defined as when a below-average quantity of precipitation causes an environment to experience a protracted period of water shortage – be it atmospheric, surface water or groundwater. Droughts are exceedingly detrimental to plant growth, development, reproduction and yield (Gallagher et al., 1976). When studying drought, it serves as useful to consider it to produce a number of effects, each of which contribute to plant health in a distinct fashion.

A chief element in a plant's experience of drought is that of an environmental water deficit. Key to understanding how a plant manages this circumstance and a core concept underpinning plant water relations is that of water potential (Ψ, units Pa), a measure of the relative concentration of water molecules present in a solution. Water potential can differ between neighbouring regions, such as cellular compartments. Water molecules belonging to a region of higher Ψ have a greater level of free energy than molecules in

the less concentrated, lower Ψ region, and so have a tendency to flow to regions of lower Ψ potential. However, cellular structures that exist between these regions, such as phospholipid membranes, can interfere with and even prevent this flow.

Well established conceptual whole plant water relation models describe a water potential gradient spanning from a height, found in the soil area drawn upon by the plant, throughout the plant, from root to shoot, ending with the exit of water via the stomata as vapour (Williams et al., 1996). This pathway by which water continually moves is referred to as the Soil Plant Atmosphere Continuum (SPAC). Decreased groundwater levels produce a lowered water potential in the soil environment immediately surrounding plant roots, hence decreasing or even abolishing the gradient of the pre-existing water potential gradient spanning the SPAC. Plants respond by correspondingly lowering their internal water potential so as to maintain a gradient and an influx of water, which allows the plant to preserve turgidity and structure. This lowering of water potential also facilitates the continuing flow of water from soil to leaf, enabling the persistence of physiological and metabolic processes crucial for survival, growth and reproduction (Morgan, 1977). Sufficiently severe environmental water deficits reduce water influx to a level by which turgidity can no longer be maintained, resulting in wilting, and a plethora of consequences which are detrimental to metabolism and physiology.

The well-established ‘cohesion theory’ describes a continuous column of water that moves across the SPAC due to transpiration forces acting on the column (Hungate, 1934). Stomata, cellular pore complexes found on the adaxial and abaxial surfaces of leaves, are key players in the regulation of plant water loss. Transpiration, the loss of water from leaves via stomata as a result of differences in the water potential between the surrounding bulk air and the water vapour in substomatal cavities, termed the vapour

pressure deficit (VPD), represents the primary source of water loss in plants. Plants are continuously tasked with the challenge of balancing stomatal water loss with the uptake of gaseous carbon dioxide required for the fixation of carbon in photosynthesis.

Humidity, defined as the amount of atmospheric water vapour, is typically lowered in instances of drought, resulting in a lowered water potential of the bulk air surrounding a given plant. This produces an increased gradient between the bulk air and the air occupying the intracellular air spaces within the leaf, heightening the VPD. Consequentially, water evaporates at a greater rate from mesophyll cells, further increasing plant water losses. Typically, stomatal aperture is reduced in response to such conditions, reducing rates of gas exchange, resulting in a decrease in the concentration of mesophilic carbon dioxide and rates of photosynthesis. Reduced rates of photosynthesis and intracellular water potential produce a multitude of detrimental physiological and morphological consequences, at both the cellular and whole plant level. Therefore, a complex integrated response to water deficient conditions is required whereby water loss is minimized and water uptake levels maintained, combined with mechanisms to protect the functions of proteins and cellular structures (e.g. cell walls) and repair damages that are inherent to this abiotic stress.

Reductions in groundwater levels also result in a hard-baked soil, which poses a substantial mechanical impedance to the progression of root growth in the exploration of deeper soils in search of lower layers likely to hold water (Passioura, 2002). Mesophytic and xerophytic species typically exhibit deep, elaborate root systems (Fitter and Hay, 2002), likely owed to the effective drying of soils from the surface downwards, facilitating the exploitation of water saturated soil horizons and delaying the onset of the detrimental effects of a water deficit.

Collectively these factors contribute to an increased loss of water by the plant and thus a state of excessively increased intracellular solute concentrations, termed hyperosmotic stress. Hyperosmotic stress is commonly referred to simply as osmotic stress, due to its comparatively high research interest relative to hypoosmotic stress, encountered by plants experiencing an environmental excess of water.

Osmotic stress disrupts homeostasis and metabolism, thus requiring a variety of physiological and morphological responses throughout the plant in order to survive, function and even maintain growth and development. The exact manner in which these various response phenomena correspond and coordinate with one another remains unclear, therefore a picture of whole plant physiological responses to hyperosmotic stress remains piecemeal. Generally, osmotic stress inhibits cell division, expansion and elongation (Nonami, 1998), resulting in decreased growth traits at the whole plant level. As the plant also experiences a reduction in net carbon uptake, so the scope for growth and development is reduced further, requiring a tightly regulated whole plant response.

Whilst hyperosmotic stress can have stereotypic deleterious consequences for plant physiology and productivity, varying adaptive responses are produced depending upon exposure time, the intensity of the factor, the species or even the individual plant exposed to the stress. For example, leaves respond to a reduction in water potential in a stereotypically graded manner, such that a lowering of tissue water potential by less than 0.5 MPa – such as that of mild osmotic stress - is conducive to a minor deregulation of basic homeostatic functions such as protein function and the biosynthesis of cellular components such as cell wall materials (Fitter and Hay, 2002). Under moderate osmotic stress, leaf water potential is reduced between 0.5 and 1.0 MPa; this state is

accompanied by a reduction in rates of photosynthesis, a more pronounced reduction of turgor, and a decrease in stomatal aperture (Fitter and Hay, 2002). Severe osmotic stress reduces leaf water potential values to lower than -1.0 MPa, resulting in a halting of photosynthetic CO2 exchange, substantially heightened respiration rates and even xylem cavitation within herbaceous plants (Borghetti, 1993; Fitter and Hay, 2002).

It is to be expected that these stress levels elicit somewhat distinct responses from plants, highlighting once more the need for a clear understanding of the potency of osmotic stress applied to plants in this investigation. This overlapping graded effect also illustrates some of the difficulties associated with asserting which characteristics correspond to particular stress levels. It also follows that similar graded responses to osmotic stress are to be found in other tissues, namely roots and stem structures.

1.3 Plant responses to drought and hyperosmotic stress

As a result of the sessile nature of plants, they are prone to biotic and abiotic changes in their environment that can impede growth and development. Geographical distributions of plants are influenced by the combinatorial effects of abiotic factors such as drought, cold, heat stress, flooding, nutrient deficiencies, etc., thus serving as major evolutionary drivers of adaptation (Normand, 2009). Hence, the mechanisms by which plants detect and coordinate responses to such abiotic stresses represents a fundamental area of biological research.

Drought signals, such as hyperosmotic stress, are detected by plants and relayed to effectors that coordinate short term and long-term adaptive approaches that involve

alterations to physiology and metabolism, so as to direct morphological and phenological responses that help to mitigate adverse environmental conditions (Figure 1). An established inherent variation between genotypes, in the approaches employed and the efficacy of their action, forms the basis for this study's experimental rationale (Vassileva et al., 2011; Mwadzingeni et al., 2016).

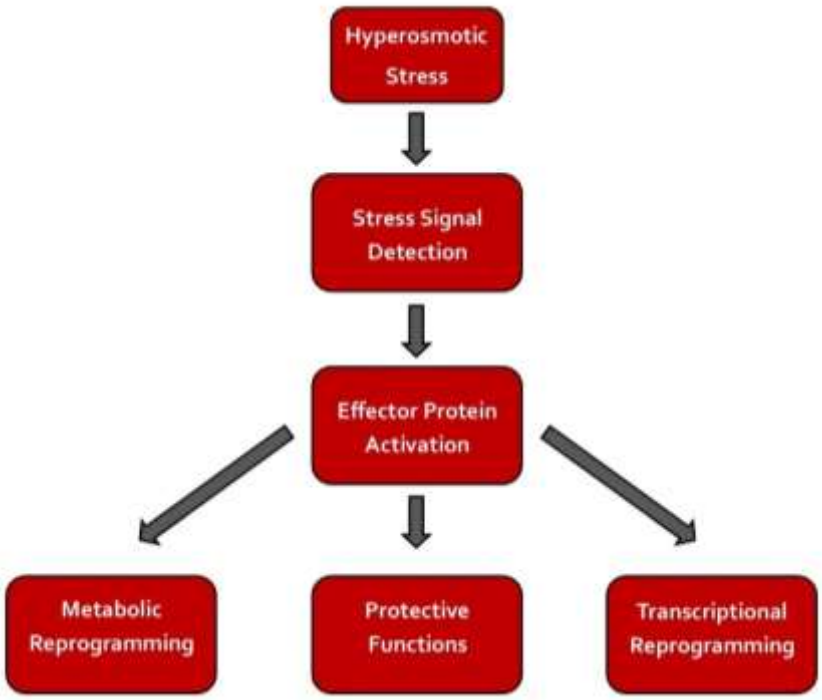


Figure 1.1: Schematic depicting the sequential cellular actions involved in hyperosmotic stress signal detection and the downstream coordination of responses.

Prevailing conceptual models postulate that plants experience decreases in productivity as a result of drought, and respond by adjusting their physiology and morphology to optimize root exploration of surrounding soils so as to maximize water uptake, whilst also decreasing the rate of photosynthesis in leaves (Grant et al., 2005; Nicotra et al., 2007).

1.3.1 Hyperosmotic stress signalling

Contemporary models of abiotic stress describe primary and secondary stress signals that are detected by the plant at a cellular level, and transduced in order to regulate proteins that coordinate transcriptional and metabolic reprogramming, and mediate homeostatic and protective functions (Figure 1.1; Zhu, 2016). Primary signals, initial cues to the cell that the plant is experiencing either local or systemic stress, can in theory originate from anywhere in the cell that is altered by the stress. In the case of drought, the primary signal is hyperosmotic stress. Secondary signals are defined as the complex secondary effects of the initial stress action, examples include metabolic dysfunction, oxidative stress and detected damages to cellular constituents such as proteins, lipids and nucleic acids (Zhu, 2016). Stress induced disruption of organelle function is also a recognized component of cellular stress responses; indications of change to organelles including the chloroplast, mitochondrion, cell wall, peroxisome and nucleus have been considered as signals that are integrated into the plant's response to abiotic stress, although these processes remain poorly understood (Estavillo et al., 2011; Ng et al., 2014; Tenhaken et al., 2015; Liu and Howell, 2016).

Progress has been made in elucidating components of some core response pathways and in identifying their downstream effectors. The identification of putative hyperosmotic stress sensors has proved difficult; current thinking points to environmentally-induced physical changes in phospholipid membranes (Sangwan et al. 2002), implicating integral membrane-bound proteins such as OSCA1, a hyperosmolality-gated calcium-permeable channel protein that has been identified in Arabidopsis (Yuan et al, 2014). Characterising OSCA1 has proved difficult due to a lack of drought or salt stress phenotypes in *osca1* mutants, suggesting a degree of functional redundancy amongst such putative hyperosmotic stress sensors.

The phytohormone abscisic acid (ABA) accumulates in the presence of hyperosmotic conditions generated by drought and salt stress, hence both stresses produce distinct and overlapping responses (Zhu et al., 2002). ABA is central to osmotic stress signalling and has been shown to elicit a variety of cellular responses (Zhu et al., 2002). Numerous ABA receptors have been identified, such as the PYL/PYR/RCAR family of membrane bound receptor proteins (Park et al., 2009; Ma et al., 2009), which upon ABA binding form a complex which activate SnRK2 proteins. The SnRK2 family of protein kinases have been shown to mediate transcriptional regulation via phosphorylation of bZIP transcription factors (Furihata et al., 2006), and control ABA-induced stomatal closure in response to drought, by phosphorylation of SLAC1, a plasma membrane anion channel protein (Geiger et al., 2009). Many other similar ABA-induced effector activations have been demonstrated. The ABA-PYL-SnRK2 signalling module has also been shown to activate MAP kinase (MAPK) cascades, proposed to similarly regulate ABA-induced effects (de Zelicourt et al., 2016).

Calcium signalling, reactive oxygen species (such as H_2O_2), nitric oxide (NO) and phospholipids (including phosphatidic acid) have been shown to be important second messengers in ABA-mediated stress response signalling (Hua et al., 2012; Brandt et al., 2015; Hou et al., 2016). Evidence has also been presented for an ABA-independent signalling pathway, responsible for the regulation of drought and osmotic stress-related genes (Jonak et al., 2002). Drought stress applied to a localised region produces systemic signalling in plants whereby responses are invoked in tissues distal to the site of the initial stress application, in a process named systemic acquired acclimation (SAA). Much remains unclear about the details of this process but it is known that hydraulic and electrical signals are used over long distances, and that waves of reactive oxygen species (ROS) and calcium ions elicit transcriptional responses in distal tissues (Miller et al., 2009; Choi et al., 2014). Importantly, yeast SNF1 and mammalian AMPK are related

to SnRK proteins (Hardie et al., 2016). Since SNF1 and AMPK are key regulators of energy metabolism, it has been proposed that SnRK proteins evolved to regulate the metabolic response to the reduced energy production observed in plants exposed to abiotic stresses (Hardie et al., 2016).

Roots are considered to be the primary tissue site responsible for the initial perception of osmotic stress, whereby initially molecular signals are produced and sent through the transpirational stream, via the xylem, towards leaves. Root-to-shoot signals include cytokinins, ABA, malate and ethylene (Zhang and Davies, 1987; Else and Jackson, 1998; Patonnier et al., 1999; Sakakibara, 2006). One of the more apparent structures to respond first to this signal are stomata. Stomata represent a crucial mediator in the trade-off between the uptake of carbon dioxide required for photosynthesis, and the corresponding loss of water vapour that occurs as a result. Stomatal conductance can respond to signals of dry soil or increased midday temperatures in a tightly regulated process controlled by a reduced root-to-shoot transport of stomatal-opening cytokinins (Davies, Kudyarova & Hartung, 2005) and increased production of stomatal-closing ABA (Geiger et al., 2009). The penetration of ABA signals is thought to be supported by the generation of ethylene in shoots (Sobeih et al., 2004; Wilkinson and Davies, 2008). Malate has been shown to function synergistically with ABA in the closing of stomata (Patonnier et al., 1999). It is proposed that subsequent to the initial presentation of water deficient soil and chemical root-to-shoot signalling, hydraulic signalling becomes primarily responsible for the regulation of shoot structure responses and the shoot-localised synthesis of ABA (Christmann et al., 2007),

Experiments with soybean (*Glycine max*) have demonstrated that independent variation in stomatal conductance between 11 genotypes contributed to water-use efficiency before and during exposure to mild drought stress (Gilbert et al., 2011). Over longer

periods, environmental conditions such as the CO2 concentration of surrounding bulk air and the intensity of incident light, influence not only stomatal aperture, but also stomatal development (Berryman, 1994; Woodward, 1987; Beerling, 1998, Ticha, 1982). It has been shown in *Arabidopsis thaliana* that mature leaves exposed to increased concentrations of CO2 respond by subsequently producing leaves with reduced stomatal density and index (Lake, 2001). Interestingly, drought induced conditions were found to reduce stomatal indices in two genotypes of *Populus balsamifera*; homologues of *STOMAGEN*, *ERECTA (ER)*, *FAMA* and *STOMATA DENSITY DISTRIBUTION 1 (SDD1)* displayed variable transcript abundances, supporting their purported role in drought-mediated adaptive stomatal development responses (Hamanishi et al., 2012). Reducing the number of stomata developed following the onset of a period of protracted water shortage reduces the capacity for gaseous exchange to a range that, given persistent root-to-shoot signalling, the plant would not exceed anyway. Hence, this could be perceived as an energy saving mechanism, preventing the wastage of biomaterials lost developing ‘unused’ stomata. Such regulation of the allocation of carbon-based resources is a primary theme of whole plant physiology responses to water shortage.

Despite these understandings, much is still unclear as to the molecular mechanisms by which plant cells detect and signal the presence of hyperosmotic stress and coordinate downstream responses. Major signalling components such as protein kinases, receptor proteins and transcription factors, of pathways involved in the coordination of adaptive morphological responses to hyperosmotic stress conditions, are likely to contribute to measurable biomass accumulation trait values, in a variable manner. It may also be expected that plant populations belonging to environments often challenged by hyperosmotic stress conditions will possess genetic profiles, containing sequence variants and expression patterns, for these said pathway constituents, that allow them to persist during exposure to hyperosmotic stress. These resistant profiles could be

expected to also reliably contribute to variation in measured morphological traits; facilitating the identification of some of these core signalling components.

1.3.2 Regulation of genes differentially expressed in response to hyperosmotic stress

As mentioned previously, ABA-dependent and ABA-independent signalling modules have been described (e.g. ABA-PYL-SnRK2) that regulate the activity of transcription factors, such as bZIPs, WRKYs, bHLHs, ERFs and MYBs, which themselves either directly regulate the expression of genes associated with drought and osmotic stress responses or indirectly regulate such genes by first controlling the expression of other transcription factors (Khan, 2012; Sun et al., 2012; Lee et al., 2014; Sun and Yu, 2015; Liu et al., 2015).

In addition, water deficient conditions have been shown to affect the synthesis, accumulation and redistribution of hormones other than ABA, namely auxin, cytokinin, ethylene, brassinosteroids, jasmonates and salicylic acid – which have been shown to regulate gene expression in a manner concurrent with or antagonistic to ABA action (Sun et al., 2014; Ma et al., 2015; Ahmad et al., 2016; Shakirova et al., 2016; Yang et al., 2017; Noreen et al., 2017). Supplementary to hormone regulated gene expression in response to stress, sugars play a key role in activating specific or hormone crosstalk transduction pathways and controlling gene expression (Ramel et al., 2009). Sugars have also been shown to be among the primary signallers in the regulation of 25 – 50% of transcripts under circadian control (Blasing et al., 2005), highlighting the link between sugars, photosynthesis-related transcript regulation and circadian rhythm. The way in

which these hormone and sugar signals are integrated in order to coordinate plant responses to water deficits remains unclear.

Some of the genes differentially regulated in response to osmotic stress have been characterised and their physiological role at least partially defined. This has allowed for the emergence of a working, albeit piecemeal, generalised model for cellular responses to hyperosmotic stress. Cellular responses to osmotic stresses vary depending on the developmental stage and location of the tissue to which a given cell belongs (Skirycz et al., 2010); however, certain common physiological processes and characteristics have been associated with plant cells exposed to osmotic stress. Downstream of hyperosmotic stress signalling pathways are effector proteins which coordinate metabolic reprogramming, protect proteins and cellular structures, and remove harmful molecular products associated with hyperosmotic stress such as reactive oxygen species (ROS).

1.3.3 Hyperosmotic stress-induced reprogramming of cellular metabolism

The regulation of nitrogen and carbon metabolism at the cellular level is key to controlling growth and development. Under normal stress-free conditions, carbohydrate quantities are continually altering due to the carbon and energy requirements of the whole organism, and the cell starch-sucrose partition, both of which are under the control of drought (Chaves et al., 1991; Pinheiro and Chaves, 2011). Canonical growth and development also affect rates of protein expression, altering the existing quantities of nitrogenous compounds such as amino acids and peptides, requiring facilitatory changes in the expression and activity of nitrogen metabolism-associated enzymes. Hyperosmotic stress is defined by a decrease in rates of photosynthesis and total productivity, meaning

also a decrease in energy resources and growth potential for the whole plant. In response to this shift in resource availability, a plethora of alterations to nitrogen and carbon metabolism are invoked in the various plant organs, the fine details of which remain unresolved (Pinheiro et al., 2001; António et al., 2008). Aside from demonstrated instances of transcription factor-mediated regulation of metabolism associated genes, sugars particularly have been shown to regulate the expression of genes associated with nitrogen, carbon and lipid metabolism (Koch, 1996; Rolland, Baena-Gonzalez and Sheen, 2006), in addition to modulating the activity of carbon-related pathway enzymes and serving as substrates (Gibson, 2000).

1.3.4 Hyperosmotic stress-induced production of reactive oxygen species

Large rapid decreases in cellular water potential, such as that of osmotic stress, result in an increased production of reactive oxygen species (ROS; Mitler et al., 2006; Miao et al., 2006). This is due to an uncoupling of numerous metabolic pathways, contributing to an energy imbalance and the subsequent transfer of high-energy state electrons to oxygen molecules, producing ROS. Also, the act of decreasing stomatal aperture in response to water deficient conditions and thus reducing stomatal conductance, produces a decreased intercellular CO2 concentration in leaves, which when coupled with sustained high irradiance can produce a relative excess of incident energy (Pinheiro and Chaves, 2011). This excess of incident light can generate an over reduction of the NADP pool and produce ROS. Typical examples of the species produced include singlet oxygen (O_2), superoxide (O_2^-) and hydrogen peroxide (H_2O_2). As well as their previously described role as signal transduction molecules, each of these ROS function as toxic by-products, damaging proteins, lipids and DNA (Apel and Hirt, 2004). The rate of ROS production is

dependent upon the duration and intensity of a given stress, as well as the capability of the cell to acclimatise to the conditions (Miller et al., 2010). In order for efficient ROS detoxification, sufficient levels of ROS scavenging enzymes such as catalase, ascorbate peroxidase and superoxide dismutase, and antioxidants including ascorbic acid and glutathionine, are required (Takahashi and Asada, 1988; Apel and Hirst, 2004; Mitler, 2004; Dietz, 2006).

1.3.5 Regulation of photosynthetic metabolism in response to hyperosmotic stress

Photosynthesis is a primary metabolic process responsible for the light-dependent production of adenosine triphosphate (ATP) and the reduction of coenzyme (e.g NADP), facilitating the assimilation of carbon dioxide in the Kreb's cycle to form carbohydrates vital for cellular homeostasis, growth and development. Photosynthetic rates are reduced in response to osmotic stress, primarily due to a decreased supply of diffusible carbon dioxide and various metabolic constraints imposed by water deficient intracellular conditions (Galmes et al., 2007).

Sugars have been highlighted as key signalling molecules in the regulation of metabolism. Low intracellular sugar concentrations are understood to up-regulate source activities like photosynthesis and nutrient export whilst high intracellular sugar concentrations are seen to do the opposite, down-regulating such activities by gene de-repression mechanisms (Pego et al., 2000; Pinheiro and Chaves, 2011). This can be considered a partial explanation for the down regulation of starch synthesis seen in response to water deficient conditions (Chaves, 1991).

Interestingly, the relative concentration of non-structural carbohydrates (NSCs) are seen to increase immediately following water stress onset, an observation attributed to a greater immediate decline in growth rates compared to decreases in photosynthetic rate (Pinheiro and Chaves, 2011; McDowell, 2011). Numerous crop species, including wheat, also exhibit stem structures used as temporary reserve stores (Chaves et al., 2002), which are accessed to support growth and development upon exposure to water deficits.

It has also been shown that C3 plants respond to the above mentioned excess of reducing power by down regulating photosynthesis via the use of regulated thermal dissipation achieved by light harvesting complexes such as those which involve the xanthophyll and lutein cycles (Demmig-Adams and Adams, 1996; Demmig-Adams et al., 2006; Pinheiro and Chaves, 2011). Reducing the production of ROS by the photosynthetic machinery in response to osmotic stress is key, as these ROS that have been purported to damage photosynthetic machinery, namely ATP synthase (Lawlor and Tezara, 2009). Reduced production of ATP in response to osmotic stress, demonstrated by Tezara et al. (1999), inhibits the regeneration of ribulose bisphosphate by ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco), an enzyme vital in carbon fixation. Rubisco was shown to exhibit species-specific patterns of reduced activity in the presence of water deficit-induced decreases in CO2 concentrations in eleven Mediterranean plant species (Galmés et al., 2011). Patterns were found to be dependent upon leaf characteristics, with species ordinarily accustomed to lower CO2 concentrations displaying lower thresholds of rubisco de-activation, hence maintaining activity under more severe drought conditions (Galmés et al., 2011). In addition, it has been suggested that species belonging to environments which regularly expose the plant to low intercellular CO2 concentrations are adapted to such conditions in part thanks to an ability to maintain rubisco activity under cases of severe drought (Pinheiro and Chaves, 2011).

Hence, the substantial decreases in intercellular CO_2 concentrations brought about by restrictive water concentrations evoking decreases in stomatal conductance, have wide reaching effects on photosynthetic output. This challenge for plant energy metabolism requires tight regulation in response, reflected in the protein homology of SnRKs, SNF1 and AMPK. As would be expected, water deficient conditions also elicit regulatory alterations to respiratory metabolism.

1.3.6 Regulation of respiration in response to hyperosmotic stress

Rates of growth and development depend largely upon the relative balance of photosynthetic carbon fixation and the release of carbon via respiration. Well-watered plants have been shown to release 30 – 70% of their daily fixed carbon back into the atmosphere by respiration (Atkin and Macherel, 2009). Regulated alterations to respiratory metabolism are made in response to water deficits. Although the complex nature of these changes is yet unresolved, generally it appears as though rates of respiration can either increase, maintain or decrease (Gimeno et al., 2010). Prevailing explanations for this attribute decreases in respiratory rate to a decrease in the photosynthetic production of mitochondrial substrate, coupled with a decrease in leaf growth rates (Pinheiro and Chaves, 2011; Gimeno et al., 2010). Increases in respiratory rates exhibited by plants exposed to water shortages have been explained as evidence for respiration processes compensating for the decrease in photosynthetic ATP production observed in water stressed plants (Tezara et al., 1999), and to provide the energy-rich molecules required to support adaptive drought responses (Gratani et al., 2007; Slot et al., 2008). Ultimately, relative to net photosynthetic carbon fixation,

respiration rates proportionally increase, resulting in an increase in the respiratory contribution to intercellular CO2 levels (Lawler and Tezara, 2009).

1.3.7 Repartitioning of photoassimilate

Substantial metabolic reprogramming occurs throughout the plant in response to hyperosmotic stress. In addition to the transcriptional regulation of metabolic enzymes discussed previously, post-translational modifications also represent a predominant factor in the regulation of these enzymes (Stitt et al., 2010). Subsequently, the plant implements a reorganisation of photoassimilate partitioning; shoots experience a greater decrease in growth than roots, wherein growth rates are better maintained, resulting in a stereotypical increase in root fraction (RF, Sharp et al., 2002). Current research has yet to produce a clear consensus on the growth responses of roots exposed to osmotic stress, beyond the accepted increase in RF; many point out that *Catharanthus roseus*, for example, produces an increased rate of root growth (Jaleel et al., 2008), whilst crop species such as maize do not demonstrate a dramatic decrease in root growth (Anjum et al., 2011). In wheat, root growth rates are typically seen to substantially decrease following the introduction of an osmotic stress (Fang et al, 2017).

Water deficient conditions also produce substantial reductions in plant canopy size, whereby the plant exhibits a reduction in total leaf numbers per plant and the size of individual leaves (Anjum et al., 2011), a fact which has been attributed to decreased quantities of photoassimilate. This decrease in leaf area is considered as a means by which to reduce the capacity for gaseous exchange-related water loss and to reduce leaf

exposure to incident light and the associated negative temperature effects that arise when coupled with osmotic stress. Numerous crop plants including bread wheat have been shown to reduce leaf angle to produce similar effects on leaf temperature (Lonbani and Arzani, 2011).

Water shortages also increase rates of leaf senescence. Leaf senescence, the breakdown of membranes and chlorophyll associated with a deterioration in function in water deficient leaves, is a key indicator of reduced photosynthetic activity (Chandler, 2001). The process appears genetically controlled, with NAC-domain and WRKY transcription factors implicated as key players, in wheat and numerous other species, indicating a genetic senescence program common to monocots and dicots (Gregersen and Holm, 2007).

1.3.8 Drought escape response

Aside from the immediate and short-term drought response strategies described previously, many species including cereal crops such as wheat, employ drought escape (DE) responses, a well-defined adaptive response to terminal drought conditions that are of a high likelihood of occurring late in the growing season (Shavrukov et al., 2017). DE species exhibit a rapid completion of their life cycle so as to avoid the negative consequences of a severe water deficit during the sensitive periods of anthesis and grain-filling. Employing a high metabolic rate and open stomata, which provide a high rate of gas exchange, DE plants produce high rates of photosynthesis, enabling rapid plant development in a manner of low water-use efficiency; counterposing the strategies of drought tolerance (Bodner et al., 2015; Kooyers et al., 2015). Therefore, plants that are

unable to switch from a physiological mode state characteristic of DE responses, are acutely vulnerable to a sustained mild osmotic stress, suffering relatively substantial negative growth effects (Loss & Siddique, 1994; Schmalenbach, 2014).

Drought escape is a term classically only assigned to ephemeral native plants; however, it has been shown that wheat displays a highly similar response typically referred to as ‘early flowering’ or ‘early maturing wheat’ (Araus et al., 2002). Such short cycle plants are influenced by a strong genetic background of gene groups described as photoperiod, vernalisation and earliness genes (Zheng et al., 2013). Early flowering responses in wheat growing in Mediterranean-type climates that frequently experience terminal drought, such as those grown in Northern and Southern regions of Africa, Southern Europe and Middle Asia (Bentley et al., 2013), display early flowering responses. Landraces with countries of origin within these regions such as South Africa, Spain and Morocco are included within the panel subset focused upon in this study. It has been demonstrated that in such environments wheat plants that employ DE mechanisms are capable of producing higher yields compared with plants which do not (Stapper and Fischer, 1990). Therefore, identifying efficient genetic variants belonging to the well-defined genetic pathways responsible for DE mechanisms represents an attractive prospect to breeders attempting to generate cultivars to be grown in Mediterranean-type climates.

1.4 Predicted effects of drought on agricultural yields

Triticum aestivum (bread wheat) contributes approximately a fifth of global daily protein and calorie intake, and annually is cultivated on an area of 220 million Ha, being at

present the most extensively grown stable food crop worldwide (Shiferaw, 2013; FAO, 2017). Abiotic stresses such as drought and heat stress can dramatically reduce bread wheat productivity by impeding photosynthetic and metabolic processes necessary for optimal growth and development, and correspondingly robust yields (Gregorová, 2015; Zhang, 2017). Wheat yields have already been shown to be suffering greater decreases as a result of drought and heat stress incidents over the last decade, compared to several of the last previous decades (Lesk, 2016).

Understanding the mechanisms by which plants respond to such stresses, and the adaptations that underpin their tolerance, is key to improving agricultural productivity and meeting the existing demands for higher yields. Advancements in next-generation sequencing technologies, coupled with the recent genome publication by the International Wheat Genome Sequencing Consortium (IWGSC), of the *Triticum aestivum* cultivar Chinese Spring, have facilitated new approaches to identifying and understanding genetic markers associated with osmotic stress (Appels et al., 2018).

1.5 *Triticum aestivum* and the identification of drought resistance genetic markers

1.5.1 Evolution, domestication and the *Triticum aestivum* genome

Triticum aestivum is considered to have originated in tandem with the advent of modern agriculture, in the Fertile Crescent, some 10,000 years ago (Salamini et al., 2002). The hexaploid genome of bread wheat (AABBDD) is thought to have derived from multiple allopolyploidization and hybridization events by which diploid progenitor species donated

their respective genomes. Current models describe hybridization between an unknown relative of the goatgrass *Aegilops tauschii* contributing the B genome, and *Triticum urartu* (AA), sometime within the last few thousand years, producing the extant tetraploid *Triticum turgidum* (AABB; emmer wheat; Peterson et al., 2006). Subsequently, it is hypothesized that emmer wheat hybridized with *Aegilops tauschii*, supplying the D genome, to form hexaploid bread wheat (Peterson et al., 2006; Salamini et al., 2002). The timing of this second hybridization event is placed at some time in the last 10,000 years, solely on the grounds of fossil evidence and the absence of *T. aestivum* growing outside of a domesticated setting (Lev-Yadun, Gopher & Abbo, 2000; Riehl, Zeidi & Conard; 2013)

Attempts have been made in estimating the phylogenetic history of the A, B and D genomes based on available genome sequences using gene tree topology analysis of 275 gene trees taken from the *T. aestivum* subgenomes and the related species *T. monococcum, T. urartu, Ae. speltoides, Ae. sharonensis* and *Ae. tauschii* - whilst taking *Hordeum vulgare, Brachypodium distachyon* and *Oryza sativa* as outgroup species (Figure 1.2; Marcussen et al., 2014). Time estimations were also made for the divergence and hybridisation events using Bayesian hierarchical models. Marcussen et al. also present evidence for the D genome being a consequence of a hybridisation of the A and B genomes (Figure 1.2). It is clear that the D subgenome within *T. aestivum* contains considerably reduced genetic diversity compared with the A and B subgenomes – supporting the notion of a far more recent hybridization (Peterson et al., 2006). This is further strengthened by the evidence of higher levels of homologues within the D genome, coupled with an apparent reduced number of pseudogenes (Appels et al., 2018). Hence, genetic variation within the YoGI panel is localised primarily to the A and B subgenomes; thus, genetic markers from these subgenomes are considerably more likely to be associated with phenotypic variation in genome wide association studies (GWAS).

Collectively, what is clear from existing research is that the ancestry of the *T. aestivum* genome is complex and the D subgenome is a relatively new addition to the hexaploid crop, resulting in decreased genetic diversity as compared with its counterparts.

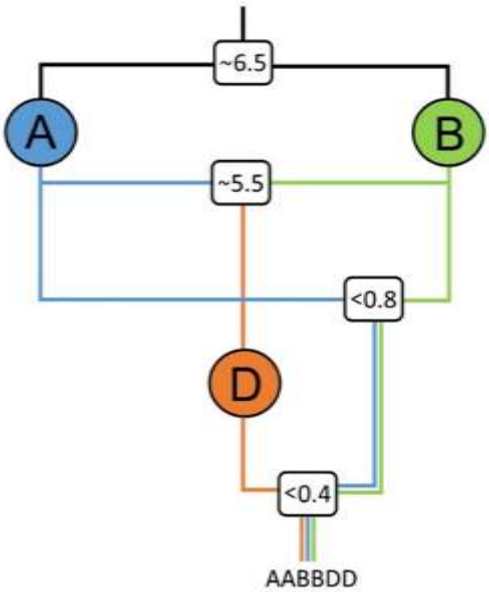


Figure 1.2: Schematic depicting the phylogenetic history of *Triticum aestivum*: Approximate dates of the divergence of A and B genomes, a D genome hybridisation event and allopolyploidization events. Time estimate units are in Ma. Figure is adapted from Marcussen et al., (2014).

1.5.2 Landraces: Genotypic and phenotypic variation

Interspecies variation in form and function when faced with water deficient conditions can be drastic, yet given the millions of years since these species diverged from a common ancestor, such distinct genetic profiles as are required to produce these contrasting phenotypes may be expected. When aiming to identify loci that contribute to agronomically relevant drought resistant phenotypes, breeders seldom rely on the introgression of genetic material from distinct species and instead rely upon naturally

existent intraspecies variation. Modern cereal cultivars, defined by a mono-genotypic profile and their adaptation to producing reliably robust yields under high input environmental conditions, have become widely adopted since the development of formal breeding methodologies (Harlan, 1975). Modern cultivars have undergone breeding programmes to optimize agronomically relevant traits such as grain yield and disease resistance. Whilst this has so far proved successful in bolstering agronomically relevant traits, it has resulted in the negative consequence of a reduced allelic diversity coupled with the observation of a relative stagnation of such agronomic advancements in low input environments (Annicchiarico & Pecetti, 1998; Tilman, 1998). Therefore, the identification of genetic elements in non-cultivar genotypes that contribute to drought resistant phenotypes, without negatively affecting yield traits, and their subsequent introgression with elite cultivars, may serve to better optimize wheat agriculture.

Previous to the global largescale adoption of modern cultivars, farmers were reliant upon native landraces, defined by Camacho Villa (2005) as being a ‘dynamic population of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as being often genetically diverse, locally adapted and associated with traditional farming systems. Increasingly, landraces are being recognized for their abundance of genetic diversity, that represents a potential source of novel abiotic stress and disease resistance genes that can be bred into existing superior modern cultivar lines. Already examples such as the Barley gene *mlo* which confers resistance to mildew, having been identified in Ethiopian landraces, has since been incorporated into modern cultivar lines (Piffanelli et al., 2004).

Bread wheat landraces are now rarely cultivated on a large scale, instead they are typically grown by subsistence farmers, or preserved in various germplasm collections.

Characterised by their abundance of variety and heterogeneity, bread wheat landrace genetic diversity still requires much description and understanding. Typically, such diversity is explained by morphological and agronomic markers defined by the International Plant Genetic Resources Institute (IPGRI, 1985), such as flowering time (Motzo and Guinta, 2007), grain yield (Moragues, 2006a) and harvest index (De Vita et al., 2007; Newton et al., 2007). Recently other morpho-physiological traits have been employed, such as biomass accumulation (Moragues, 2006b; De Vita et al., 2007), which will be adopted for use in this study. Such morphological descriptors are opted for on the basis of their heritability and their ubiquitous expression regardless of environment.

A key resource in this investigation is the YoGI diversity panel, which contains 346 *Triticum aestivum* accessions, with transcriptomic gene expression marker (GEM) and single nucleotide polymorphism (SNP) information derived from mRNA-seq data taken from the first true leaf of 6-day old seedlings. Transcriptomic data available for the panel represents a reflection of the normal leaf gene expression profile (GEMs), at that stage of development, as well as providing information on sequence variations (SNPs) found in transcribed regions of the genome. The YoGI panel contains landraces taken from 65 countries, including drought prone regions such as Kenya, Morocco and Sudan (Rojas et al., 2011), as well as regions of high precipitation such as Ecuador, Bolivia and Brazil (Figure 1.3); hence it is expected that landraces belonging to these localities will have adaptations reflective of the demands of these environments, and correspondingly possess underlying genetic variation. YoGI panel transcriptomic data is supported by gene functional annotation information recently provided by the IWGSC upon completion of the sequencing of the genome of the hexaploid *T. aestivum* cultivar Chinese Spring (CS; Appels et al., 2018).

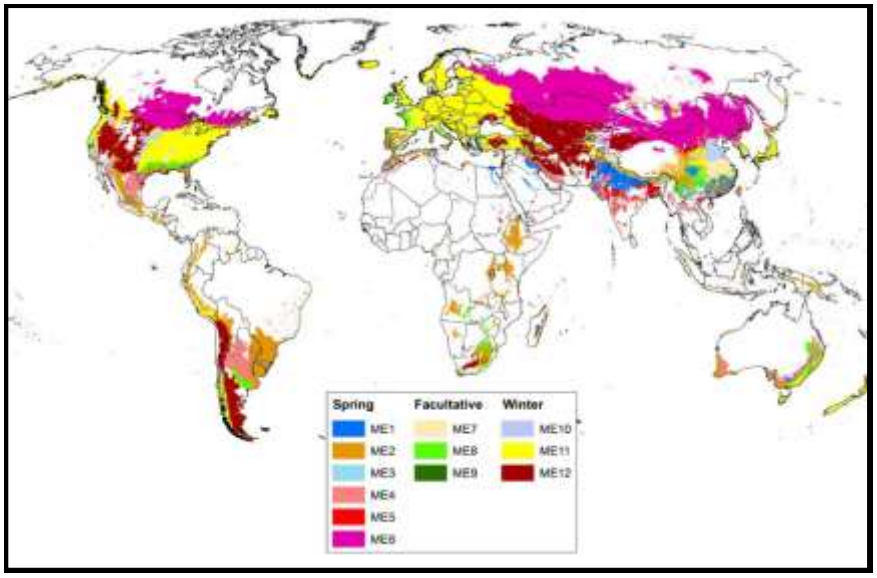


Figure 1.3: Global map of wheat mega-environments (Sonder and Kai, 2016).

1.5.3 Drought resistant markers identified in *Triticum aestivum*

Combining the use of metabolomic and proteomic techniques with the recently published Chinese Spring genome has provided a powerful insight into *T. aestivum* drought responses and the identities of genetic markers associated with drought resistant phenotypic profiles.

Metabolomic studies comparing the levels of 58 key metabolites between *T. aestivum* seedlings with drought resistant or drought tolerant accessions, found that drought elicited ‘system alterations in widespread metabolic networks involving transamination, tricarboxylic acid cycle, glycolysis, glutamate-mediated proline biosynthesis, shikimate-mediated secondary metabolisms and γ-aminobutyric acid metabolisms’ (Guo et al., 2018). Drought resistant wheat accession JD17, when exposed to a prolonged drought stress demonstrated an additive accumulation of metabolically compatible osmolytes,

namely proline, fructose, malic acid, mannose and sucrose (Guo et al., 2018). This process, termed osmotic adjustment (OA), is a well-defined facet of drought resistant responses in crop species such as wheat, rice and maize (Zhang et al., 1999). Involving the generation, transport and maintenance of solutes, OA lowers intracellular osmotic potential and so acts to rebalance osmotic potential gradients, reducing cellular water loss and maintaining turgor. Wheat accessions with a higher OA capacity, were shown to be capable of extracting a 1.3-fold higher quantity of water from their environment than those with lower OA capacities (Morgan and Condon, 1986).

Michaletti et al. (2018) compared different drought-susceptible (Bahar) and drought-tolerant (Kavir) *T. aestivum* spring cultivar lines by applying both metabolomics and proteomics, in an attempt to better understand the biochemical networks governing leaf responses to drought. Drought stress induced more severe effects on the Bahar proteome, producing a large decrease in total protein content, particularly in levels of photosynthesis, and carbon and nitrogen metabolism-associated genes. Metabolome analysis findings were concurrent with Guo et al. (2018), showing that amino acids, sugars and organic acids were primarily affected by drought stress. Proline particularly was shown to have increased levels in Bahar whilst Kavir presented a metabolism less affected by the drought stress (Michaletti et al., 2018). Starch synthesis appears to be repressed by water deficits (Chaves, 1991). Overall, it is clear that large scale metabolic reprogramming is initiated in response to water deficient environmental conditions.

It is upon the aforementioned predating research, focused upon drought, and more specifically, hyperosmotic stress responses, that this investigation hopes to build further understanding. Instead of utilising contemporary methods such as proteomics and metabolomics, this study instead employs an association genetics methodology.

1.6 Genome Wide Association Studies and Associative Transcriptomics

Hyperosmotic stress resistance is an example of a quantitative trait, meaning it is relatively very complex, with a multitude of genes contributing to a given individual's trait profile. Genes involved in osmotic stress resistance and responses have diverse functional roles such as cell signalling, metabolism, transcription and transport etc. (Kreps, 2002). Tissues in different developmental stages will also likely require specific responses to osmotic stress based on their individual biochemistries and physiologies characteristic of their growth requirements and morphologies. It is also to be expected that any given plant possesses a substantial amount of functional redundancy, as regards genetic elements governing hyperosmotic stress responses – such as that seen with OSCA1 (Yuan et al., 2014). In addition, the hexaploid nature of the bread wheat genome provides an additional complicating aspect of functional redundancy amongst contributing factors. In combination, this renders classical forward genetics approaches less appropriate.

Recent approaches to understanding such quantitative traits include Genome Wide Association Studies (GWAS), which attempt to use statistical methodologies to associate large genetic variation data sets with measured phenotypic trait variation data in order to identify some of the more prominent contributing factors (Huang and Han, 2014). So far, GWAS has been applied with success to crop plants including rice, maize and foxtail millet (Huang et al., 2012; Jia et al., 2013; Kump et al., 2011).

This study utilises a related approach, termed Associative Transcriptomics (AT). AT is a utilizes both gene expression level and sequence variation information derived from Illumina transcriptome data, here derived from the first true leaf of six day old *T. aestivum*

seedlings, and scores these markers according to the degree to which they correlate with measured trait variation (Harper et al., 2012). This enables the delineation of genomic regions responsible for controlling a particular trait of interest. A marked benefit of AT is that by utilizing only transcriptomic data, much of the issues relating to gene copy number, such as gene silencing and differential expression (Adams, 2003; Pires et al., 2004), and large sequence repeat regions, that arise when applying GWAS techniques to complex polyploid crops, are circumvented. AT has already proved successful in identifying major underlying genetic components of important traits in the tetraploid crop *Brassica napus*, namely two quantitative trait loci (QTL) encoding transcription factors that control glucosinolate biosynthesis (Harper et al., 2012).

1.7 Specific project aims

This study aimed to generate phenotypic trait data that is descriptive of a subset of the YoGI panel's landrace's osmotic stress resistance, and associate this variation with transcriptomic information available for the panel. Genetic markers identified by the experimental procedure represent key elements in transcriptional configurations prior to stress application, that significantly associate with measured trait variation – making interpretation of such results as one of gene candidacy, not of definitive involvement. It is likely that landraces will exhibit subtle variations in their hyperosmotic stress response strategies, by which a combination of additive advantageous alleles will contribute to profiles with efficient biochemistries and physiologies that better dispose an individual landrace to the challenges of a water deficient environment.

Hence, this study aimed to elucidate genetic markers that are associated with chronic moderate hyperosmotic stress responses in *Triticum aestivum*, with the mind that these identified candidates may serve to better our understanding of the molecular events involved. Methods involved the static hydroponic culturing of a subset of the YoGI diversity panel's spring lines, and the application of polyethylene glycol to a treatment group, allowing the calculation of normalised change in biomass (NCB) values. These NCB trait values were considered as a quantified phenotypic reflection of a landrace's tolerance and adaptive morphological response to hyperosmotic stress.

Subsequent use of these trait values in an Associative Transcriptomics (AT) analysis aimed to utilize RNA-seq transcriptomic data available for the panel, to reveal gene expression and sequence variation markers that significantly associate with the observed trait data. Genetic markers identified by the AT analysis were to be assessed for their plausibility as a candidate, initially through cross referencing the marker with previously published literature. In order to further validate these candidate markers identified through the AT analysis, subsequent experiments were planned and performed using Wheat Target Induced Local Lesions In Genome (TILLING) lines. Finally, the usefulness of the collection of candidate genes identified in this study, for further focused investigations into hyperosmotic stress responses in *Triticum aestivum*, is discussed and appropriate further work suggested.

2. Methods:

Experiments were conducted in the Department of Biology, at The University of York between October 2018 and September 2019. All *Triticum aestivum* seeds used within the hyperosmotic stress screen are taken from the YoGI diversity panel, stored in a temperature-controlled room at 4ºC. *Triticum aestivum* cv 'Cadenza' and *Triticum durum* cv 'Kronos' Target Induced Local Lesions In Genome (TILLING) lines were ordered from the John Innes Centre. *Arabidopsis thaliana* plasma membrane intrinsic protein 1;5 (PIP1;5) T-DNA insertion lines SALK056898 and SALK056041 were ordered from the Nottingham Arabidopsis Stock Centre (NASC) and Arabidopsis Biological Research Centre (ABRC).

2.1 Seed sterilisation, germination and culturing:

Triticum aestivum and *Arabidopsis thaliana* seeds were sterilised prior to germination to remove any microbial contaminants such as bacteria, fungi and other pathogens so as to reduce detrimental confounding effects on plant health, and to increase germination rates.

2.1.1 *Triticum aestivum* sterilization and germination:

Sixteen seeds per landrace were sterilised by being placed in 5ml Eppendorf tubes (one per landrace), soaked in 3ml ddH_2O, vortexed and then left for 1.5 hours to soften seed casings. Subsequently, seeds were soaked in 3ml triton-bleach solution treatment (33% bleach, 67%

ddH_2O, 200 µL triton-X), vortexed and then left for 10 minutes. Triton-bleach solution was then removed and seeds rinsed with 3ml ddH_2O. This was repeated 4 times. Finally seeds were soaked in 3ml ddH_2O for 1.5 hours and then transferred to square petri dishes (10 x 13 cm) containing Whatman filter paper (PEG6000 pilot) or tissue paper (PEG8000 pilot) that was dampened with ddH_2O. Petri dish covers were sealed with tape to prevent emerging coleoptiles removing petri dish covers and thus allowing the dampened substrate to dry. Seeds were then transferred to a growth room with a light-dark period of 16h/8h with a light intensity of 90 µmol $m^{-2}s^{-1}$ and left to germinate for 6 days. After three days within the growth room the paper substrate was re-dampened with 2mL ddH_2O.

2.1.2 *Arabidopsis thaliana* sterilisation, germination and soil culture:

In a laminar flow hood SALK056898, SALK056041 and Columbia (Col-0) seeds were sterilised by being placed in 5ml Eppendorf tubes (one per line), soaked in 500 µL 100% ethanol and microcentrifuged. Ethanol was removed and seeds were then soaked in 500 µL triton-bleach solution (33% bleach, 67% ddH_2O, 200 µL triton-X) and microcentrifuged for 2 mins at 16000g. Triton-bleach solution was removed and seeds were soaked in 500 µL ddH_2O and microcentrifuged again as above, then ddH_2O was removed and seeds suspended in 0.01% agar water. Sterilized seeds were then streaked in agar water on agar gel set in square petri dishes (10 x 13 cm) containing Murashige and Skoog (MS) media (4.4g/L MS basal salt, 0.5g/L MES, 1.5% Phyto Agar). Seed plates were moved to a 4ºC fridge for 72 hours for stratification. Plates were then moved to a growth room set to a light-dark period of 12h/12h with a light intensity of 90 µmol $m^{-2}s^{-1}$. Seeds were kept in a growth room for 10 days and then transferred to F2 compost (Levingtons, UK) where they were cultured in glasshouse conditions. Seedlings were initially covered by plastic domes to

maintain humidity levels; after 4 days domes were removed and plants watered regularly until seeds were ready for collection. SALK lines were cultured in GM facilities, all waste biomaterial was autoclaved prior to disposal.

2.2 Hydroponic culture and polyethylene glycol treatment

Following germination, *T. aestivum* seedlings were transferred to 9L black plastic boxes containing 9L of growth media (55% Hoagland's Solution, 45% ddH_2O). Complete Hoagland solution contained: 1.1mM·$MgSO_4 7H_2O$ (Sigma-Aldrich, Japan), 6mM·$Ca(NO_3)_2 \cdot 4H_2O$ (Alfa Aesar, USA), 4.1mM·KNO3 (Acros Organics, Spain), 1.9mM·KH2PO4 (Acros Organics, Spain), 3.6µM·MnSO·$4H_2O$ (Fisher Chemical, USA), 48µM·H3BO3 (VWR Chemicals, USA), 0.5µM·Na2MoO4·$2H_2O$ (Sigma-Aldrich, USA), 2µM·ZnSO4·$5H_2O$ (Duchefa Biochemie, Netherlands), 0.4µM·CuSO4·$5H_2O$ (Duchefa Biochemie, Netherlands), 65µM·C10H12N2NaFeO8 (Sigma-Aldrich, USA). Once formulated, media pH was measured using an electronic pH meter and 5% hydrogen chloride (HCl) used to regulate media pH to 6.5.

Control boxes contained only 9L of growth media (Hoagland's and ddH_2O) whilst PEG treated boxes contained 9L of growth media with varying concentrations of PEG added according to the experiment. PEG treatments were weighed using a top pan balance and stirred into hydroponics boxes with previously added growth media, using a sterile serological pipette, until dissolved. Plants transferred to hydroponics were given 24 hours to acclimatise prior to the addition of PEG. Nutrient solution was replaced after 7 days. Plants were grown in hydroponics for a total of 14 days. After 14 days plants were removed from hydroponics and trait measurements taken.

Plants cultured in controlled growth rooms, at room temperature, using cool white fluorescent lamps set to a 16h/8h day/night cycle, set to an intensity of 90 µmol $m^{-2}s^{-1}$. Glasshouse facilities maintained a 16h/8h day/night cycle using supplementary lighting, which only operate when light level measurements fall below 190 W m^{-2}. Temperature regulated windows open and close so as to maintain a 16°C/14°C day/night temperature.

2.3 Landrace box position allocation:

Seedlings of similar phenological stage were selected and placed into hydroponics in a randomized configuration, according to landrace identity, so as to reduce the effects of any nutrient gradients or localisation effects that may have occurred. Sample position allocation methods varied according to the experiment and size of box used, as follows:

- PEG6000 and PEG8000 pilot experiments: Four 9L hydroponics boxes were set up, with 16 seedlings in each (4 landraces, 4 biological repeats), placed in a Latin square configuration (Supplementary Figure 1).
- Panel screen experiment: Eighteen 9L hydroponics boxes were set up with 60 seedlings in each (A-R). Due to throughput limitations, boxes were separated into three technical groups, each containing three no treatment control boxes and three boxes treated with 7.5% PEG8000 (w/v; six box groups). Four biological repeats were used for each landrace in each condition (1080 total samples). Positions were allocated by assigning a random number between 0 and 1 to each biological repeat (e.g. YoGI_002, YoGI_004 etc.) for the 135 landraces used, using the =RAND() function in Microsoft Excel. Biological repeats are then ordered according to the generated number. 135 randomly ordered biological repeats, representing each landrace, were then placed into the first 135 positions of each of the six box groups.

135 randomly ordered biological repeats were then similarly generated and divided into groups of 45 and placed into the remaining positions of box groups 1, 2 and 3, in ascending order. This was then repeated for box groups 4, 5 and 6. This ensured a biological repeat for each landrace was placed in each box in a randomized fashion, with the fourth biological repeat also being assigned to a box group in a randomized manner.

- HSP40 TILLING experiment: The three lines used (Cadenza parent, Kronos parent and Kronos4610) were placed into hydroponics according to the same repeating Latin square design (Supplementary Figure 2).

2.4 Biomass trait measurements

Following hydroponic culturing, samples were removed from their respective box, and delicately dried, using tissue paper to remove any remaining surface water remaining on the plant that would otherwise offset fresh biomass measurements, and placed in paper bags labelled with the corresponding position reference. Samples would be separated at the boundary between root and shoot using surgical scissors. Root and shoot fresh biomass would then be measured for each sample as quickly as possible following their removal from hydroponics, so as to reduce the effect of drying. Weight measurements were taken using a top pan balance, to an accuracy of three decimal places. Samples were then dried in an oven set to 60ºC for 72 hours, removed and similarly weighed to measure dry root and shoot biomass. Total biomass, root fraction and normalised change in biomass trait values were calculated as follows (Equation 1, Equation 2, Equation 3 and Equation 4):

Total biomass (g) = Root biomass (g) + Shoot biomass (g) [1]

Root fraction (%) = (Root biomass (g)/Shoot biomass (g)) × 100 [2]

Normalised change in biomass = (Landrace mean control biomass (g) – Landrace treatment biomass (g)) / Landrace mean control biomass (g) [3]

Normalised change in root fraction = (Landrace treatment root fraction (%) – Landrace mean control root fraction (%)) / Landrace mean control root fraction (%) [4]

2.5 Associative Transcriptomics (AT)

Genome sequence scaffolds were orientated and ordered according to the published chromosome models developed by the International Wheat Genome Sequencing Consortium (IWGSC) with the completion of the spring *T. aestivum* cultivar Chinese Spring (Appels et al., 2018). Initial genetic marker functional information was also taken from IWGSC annotations. Expression data representing transcript abundance (RPKM) and sequence variation data taken from the panel for use in the analysis was generated using next-generation Illumina RNA-seq.

Landrace mean trait values for the adaptive morphological trait data gathered from the panel screen experiment, were calculated for each landrace using a script written in R Studio Version 1.1.453. Biomass data from samples that died during the course of the experiment was not included within the final data set. However, data for at least 2 biological repeats in each condition was collected for each of the 135 landraces tested. Subsequently, the

generated datasheet was saved in MS Excel, formatted so as to be compatible with AT R scripts, and then saved as a text delimited file.

2.6 Genotyping:

Each individual Kronos plant utilised in Wheat TILLING line experiments (Kronos4610 and the Kronos parental line), was genotyped in order to confirm the presence of the mutation originally predicted by the exome resequencing and the alignment of raw data to the IWGSC Chinese Spring chromosome arm survey sequence. Amplicons were sequenced by Sanger sequencing (GATC Biotech LIGHTrun).

Segregating T-DNA insertion SALK mutant lines utilized in this study were genotyped in order to establish genotype identity, facilitating isolation from other genotypes prior to flowering so as to prevent cross pollination when bulking seed.

2.6.1 DNA extraction

Seven days following their transfer to hydroponics, a small sample was taken from a true leaf end from each Kronos plant and stored in a 1.5ml Eppendorf tube. 2 weeks after the transfer of Arabidopsis seedlings to the glasshouse facilities, a single rosette leaf was taken from each SALK mutant and stored in a 1.5 ml Eppendorf. DNA extractions for all samples were performed according to the same CTAB protocol, as follows:

100µL of 1M DNA Extraction Buffer (DEB), containing 25 mM EDTA, 240 mM NaCl, 200 mM Tris pH 8.0 and 1% (w/v) SDS, was added to tissue leaf samples. Samples were then homogenized using an IKA RW16 basic overhead stirrer, and an additional 400µL of DEB added to each sample. 75µL of Chloroform was then added to each sample. Samples were vortexed for 5 minutes and tubes centrifuged using a countertop centrifuge at 16K for 10 minutes. 350µL of supernatant was removed using a pipette and added to new 1.5ml Eppendorf tubes containing 350µL of Isopropanol and inverted gently. Supernatant-Isopropanol tubes were centrifuged at 14.2K RCF for 10 minutes. Supernatant was decanted and 500µL of 70% Ethanol (DNA grade) added to tubes. Tubes were then inverted and centrifuged at 14.2K RCF for 5 minutes. Supernatant was decanted and any residual supernatant carefully pipetted and allowed to air dry for 30 minutes. 100µL of 1x TE was then added to dried Eppendorfs; tubes were then vortexed and frozen at -20ºC until later required.

2.6.2 Polymerase chain reaction (PCR) conditions:

Touchdown PCRs were used in all instances. Each reaction mix contained 1µL of extracted DNA solution as a template, 11.5µL of PCR Master Mix, 0.5µL of 10µmol forward primer and 0.5µL of 10µmol reverse primer.

TILLING primers were designed by aligning HSP40 homologues using BioEdit v7.2.5 (Hall, 1999), and interrogating the candidacy of appropriately placed potential primers using the Sequence Manipulation Suite's PCR Primer Stats tool (Stothard, 2000), assessing predicted melting temperature, percent GC content, and PCR suitability. Salk Institute Genomic Analysis Laboratory (SIGNAL) T-DNA Primer Design Tool was used to generate left and right border primers (LP and RP) for each SALK line (Supplementary

Table 1). LBb1.3 T-DNA insert left border primer (LB) was used.

PCR conditions were as such: 5 min at 94ºC, 15 cycles of 94ºC for 30s, 63ºC for 30s (decreasing by -1C per cycle), 72ºC for 1 min, followed by 30 cycles of 94ºC for 30s, 53ºC for 30s and 72ºC for 1 min, ending with a final elongation step at 72ºC for 15 min. Amplicons were resolved in 1.0% agarose gels that were stained with ethidium bromide.

2.7 Root architecture imaging and analysis

Root morphology was studied in Cadenza (Cadenza parent) and Kronos lines (Kronos parent and Kronos4610), germinated according to section 3.1.1 and grown in hydroponics for 14 days, half of which received a 7.5% PEG8000 (w/v) treatment according to section 3.2.

The roots of three biological repeats were imaged for each line, in each condition (no treatment and PEG treatment). Roots were captured using an Epson V1.0.3.7 to produce 600 ppi 8-bit greyscale positive film type Tagged Image File Format (.TIFF) images. WinRHIZO's standard morphological analysis was applied to these images, measuring surface area (cm^2), total volume (cm^3) and average diameter (mm).

2.8 Statistical analysis

Statistical tests performed throughout the course of this investigation, namely Kruskal-Wallis H test, t-test and analysis of variance (ANOVA), were carried out using R Studio Version 1.1.453. Shapiro Wilk test, and Bartlett's homogeneity of variance test were conducted to assert whether data sets could be treated as parametric or non-parametric. In instances of non-parametric data sets Kruskal-Wallis test were employed. Plots generated to represent the results of statistical tests were also completed in and exported from R Studio as JPEG file images. R Studio plots were produced using the ggplot2 package. Packages used for data handling in R include dplyr, plyr and the standard R maths packages.

3. Pilot Experiments

3.1 Introduction to pilot experiments

Nutrient solution supplemented with polyethylene glycol (PEG) displays a lowered water potential, simulating a water deficit (Lagerwerff, 1961). Hydroponically grown *T. aestivum*, treated with PEG8000 over a period of 72 h to simulate prolonged mild osmotic stress (5% w/v), were shown to display a stereotyped adaptive morphological response (Ji et al., 2014). Productivity was reduced; the number of lateral roots were increased; swelling of root tips and cessation of primary root growth was observed and premature differentiation of root apical meristems (RAM) demonstrated by histological observations of vasculature and Tetrazolium Violet staining of active cells (Kurzbaum et al., 2010; Ji et al., 2014). This defined adaptive morphological response was seen to be conserved in *Brachypodium distachyon*, maize, wheat, *Arabidopsis*, soybean and rice grown in water deficient soil systems (Ji et al., 2014). It was found that hydroponically cultured, known drought resistant wheat lines (SYN604 and Xiaoyan-54) treated with PEG8000 (5% w/v) exhibited a faster and more extensive response, with a greater number of lateral roots and swollen root tips, than non-drought resistant lines (Sundor and Jing-411; Ji et al., 2010; Sečenji M, et al., 2010; Ji et al., 2014). A paper by the same authors demonstrated, via loss of function mutations in ABA2 and ABA receptors, that this process is mediated by the drought response phytohormone ABA (Ji and Li, 2014). The manner in which this demonstrated phenomenon fits into the prevailing paradigm of drought resistant root traits minimizing the metabolic costs of root exploration of deeper soils is unclear. However, this investigation demonstrates that PEG8000 hydroponic culture experiments can produce evolutionarily conserved differential adaptive morphological responses from *T. aestivum* lines that relate to drought sensitivity. A similar hydroponics technique was adopted for use in this study, for which pilot experiments were carried out to test the efficacy of the setup.

Pilot experiments were performed to ensure that the planned hydroponic culture and PEG treatment setup produced morphological responses similar to those seen in other investigations, such as reduced productivity, cessation of primary root growth and increased root fraction. It was also of interest to see whether the selected landraces exhibited differences, in the measured normalised change in biomass (NCB) traits, which were statistically significant. This would provide some evidence for genotypically driven variation for these traits among landraces in the YoGI panel; therefore, demonstrating the suitability of the setup for hyperosmotic stress tolerance phenotyping and subsequent Associative Transcriptomics analysis. Hence, four landraces were selected on the basis of their country of origin, two belonging to relatively drought prone areas (X1 and X2) and two to low drought-risk countries (Y1 and Y2):

- YoGI 205, Afghanistan (X1)
- YoGI 207, Afghanistan (X2)
- YoGI 005, Bolivia (Y1)
- YoGI 333, Colombia (Y2)

Initially, polyethylene glycol 6000 was first considered, since it produces highly similar morphological responses to those mentioned above (reduced productivity, increased root fraction and cessation of primary root growth), and so was adopted for use in the first pilot. Following the results of the first pilot, and after a costing assessment for materials, PEG8000 was later trialled in a second pilot.

3.2 Methods

3.2.1 Pilot 1 – PEG 6000

Different PEG6000 concentration treatments (2.5% w/v, 5.0% w/v and 7.5% w/v) were tested in order to assert the most effective concentration for the Associative Transcriptomics experiment. The concentrations of 2.5%, 5.0% and 7.5% approximately correspond to a negative water potential of -0.025 MPa, -0.05 MPa and -0.1 MPa, which span the spectrum of mild to the threshold of moderate hyperosmotic stress (Michel and Kaufmann, 1973). Here, 2.5% and 5.0% are referred to as mild osmotic stress whilst 7.5% is termed moderate osmotic stress.

Seeds were sterilised as outlined in Chapter 2.1. Plants were cultured in black opaque plastic boxes (30 x 15 cm), in an attempt to reduce the growth of microorganisms, such as bacteria, fungi and algae, in the Hoagland's based growth media used. Media was maintained at pH6.5. Seedlings of similar phenological stage were selected and placed into hydroponics in a Latin square configuration, according to landrace identity, so as to reduce the effects of any nutrient gradients that may have developed. Four hydroponics boxes were set up, with 16 seedlings in each (4 landraces, each 4 repeats) and polyethylene glycol (PEG6000, Sigma-Aldrich, Germany) added as follows: Box A (no treatment control), Box B (2.5% w/v), Box C (5.0% w/v), Box D (7.5% w/v). Nutrient solution was replaced after 7 days. Plants were cultured in controlled growth rooms at room temperature (approx. 17-22 °C), using cool white fluorescent lamps set to 16/8 h day/night cycles, set to an intensity of 90µmol $m^{-2}s^{-1}$.

Plants were cultured according to Chapter 2.3. Fresh root and shoot weight was measured using a countertop balance. Whole plants were then dried in an air draft oven for 48 h at 60ºC. Dry root and shoot weights were similarly recorded, and root fraction and total biomass measurements calculated. Normalised change in biomass measurements were calculated according to Chapter 2.4.

3.2.2 Pilot 2 – PEG 8000

A second pilot experiment was conducted using PEG8000 to ensure that the treatment satisfied the same morphological effect criteria as previously required (cessation of primary root growth, increased root fraction and reduced productivity). Germinated seedlings were still cultured in an identical manner to that of the PEG6000 pilot experiment, except for variation on the concentrations of PEG, which were as follows: Box A (no treatment control), Box B (7.5% PEG6000 w/v), Box C (7.5% PEG8000 w/v) and Box D (15% PEG8000 w/v).

These PEG concentrations were selected on the basis of the first pilot experiments results which demonstrated that a minimum of a 7.5% PEG treatment was required to produce the aforementioned hyperosmotic stress-induced morphological effects. A 15% PEG (w/v) treatment was also trialled in order to assess whether a higher PEG8000 concentration treatment should satisfy the required morphological criteria, should the 7.5% treatment have failed. The same landraces were used, biomass trait values were similarly measured, and plants were cultured in growth rooms set to PEG6000 experimental conditions, regulation of pH was similarly conducted.

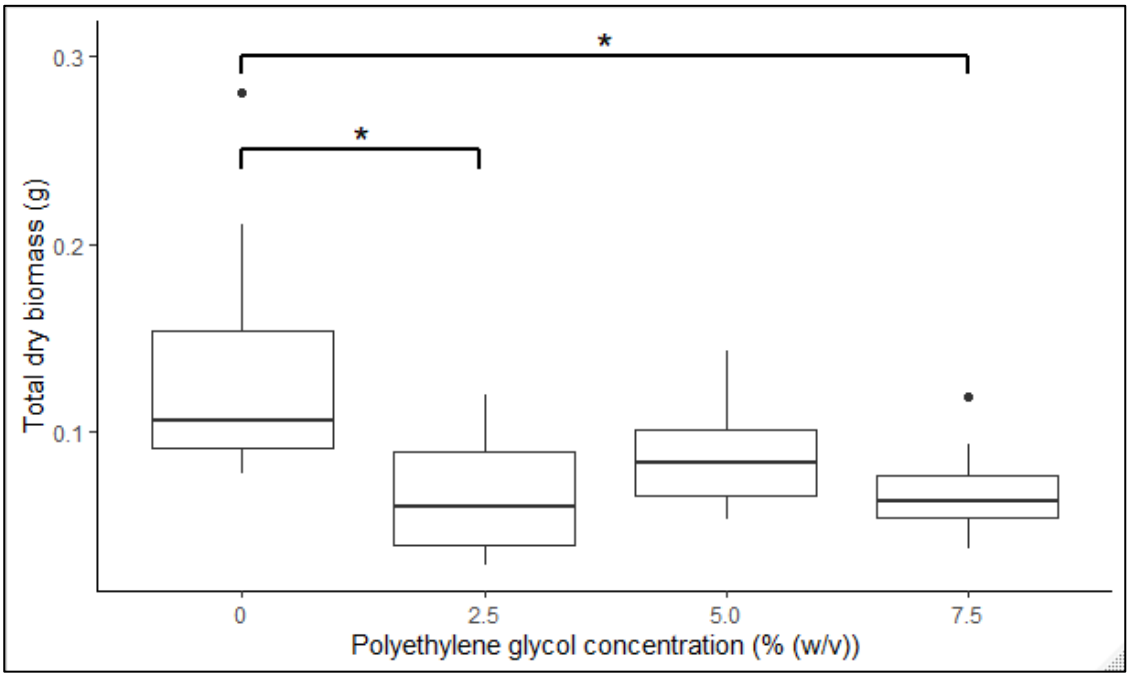
However, an alternative germination methodology was also trialled. Here seeds were sterilised as done previously and sowed, 1cm deep, in pots containing a mixture of sand and compost (20% sand and 80% F2 compost), and kept under glasshouse conditions for 6 days. The alternative germination method was adopted in an effort to seek a solution to the problems encountered in the PEG6000 experiment, in the transferral of seedlings to hydroponics.

3.3 Results

3.3.1 Pilot 1 – PEG 6000

Total dry biomass was significantly reduced at 2.5% and 7.5% PEG6000 (Figure 3.1, Kruskal-Wallis: $\chi = 22.377$, d.f. = 3, $p = 5.55e^{-05}$). Cessation of primary root growth was observed. Root fraction increased significantly at a concentration of 7.5% PEG6000 (Figure 3.1, Kruskal-Wallis: $\chi = 13.034$; d.f. = 3; p-value = 0.00456). Plants cultured under the 5.0% PEG6000 treatment did not demonstrate a significantly reduced total dry biomass, although they did demonstrate a decrease in the median total dry biomass.

A



B

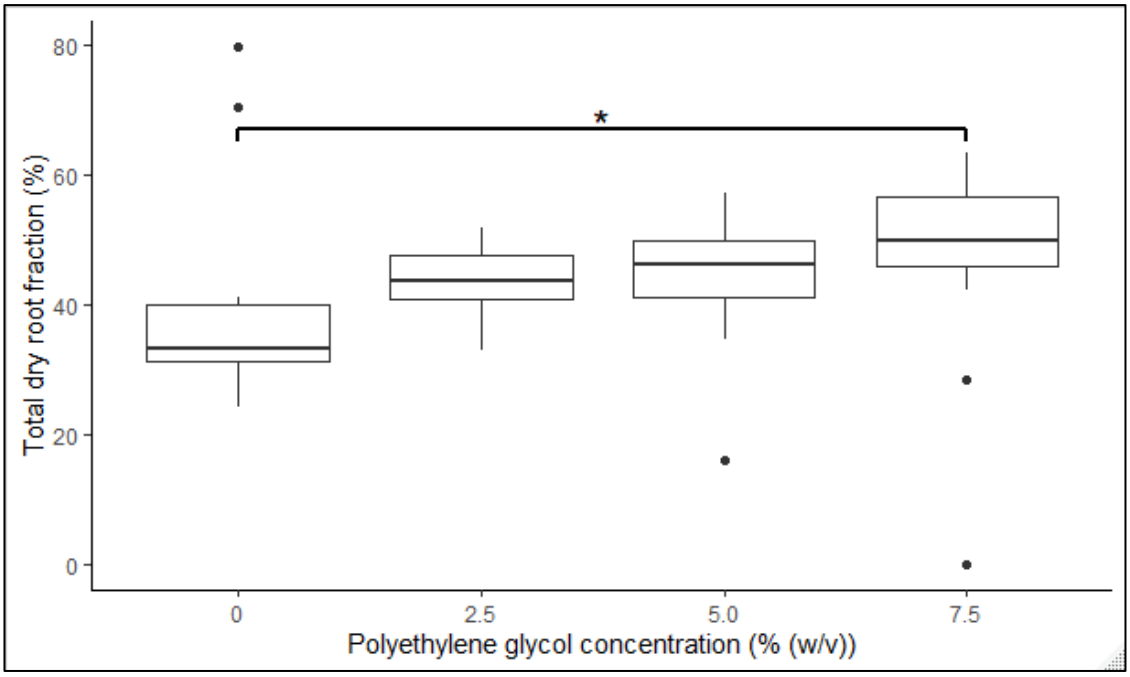


Figure 3.1: Pilot 1 Box plots. (A) Total dry biomass of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water and varying polyethylene glycol 6000 concentrations (0%, 2.5%, 5.0%, 7.5%). B) Total dry root fraction of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water and varying polyethylene glycol 6000 concentrations (0%, 2.5%, 5.0%, 7.5%). 0%; N = 15, 2.5%; N = 16; 5.0%, N = 16; 7.5%, N = 15, $p < 0.05$ (*).

Normalised changes in total dry biomass differed significantly between landraces at all concentrations of PEG, as demonstrated in Figure 3.2 for the PEG6000 7.5% treatment (ANOVA, F = 7.165, p = 0.00616). Generally, patterns of difference such as Y1 exhibiting a statistically lower normalised change in dry root biomass than X1 and Y2, were replicated in other PEG concentration treatments or no difference was found (data not shown). Normalised changes in dry root biomass and root fraction were similarly observed across PEG treatments (data not shown).

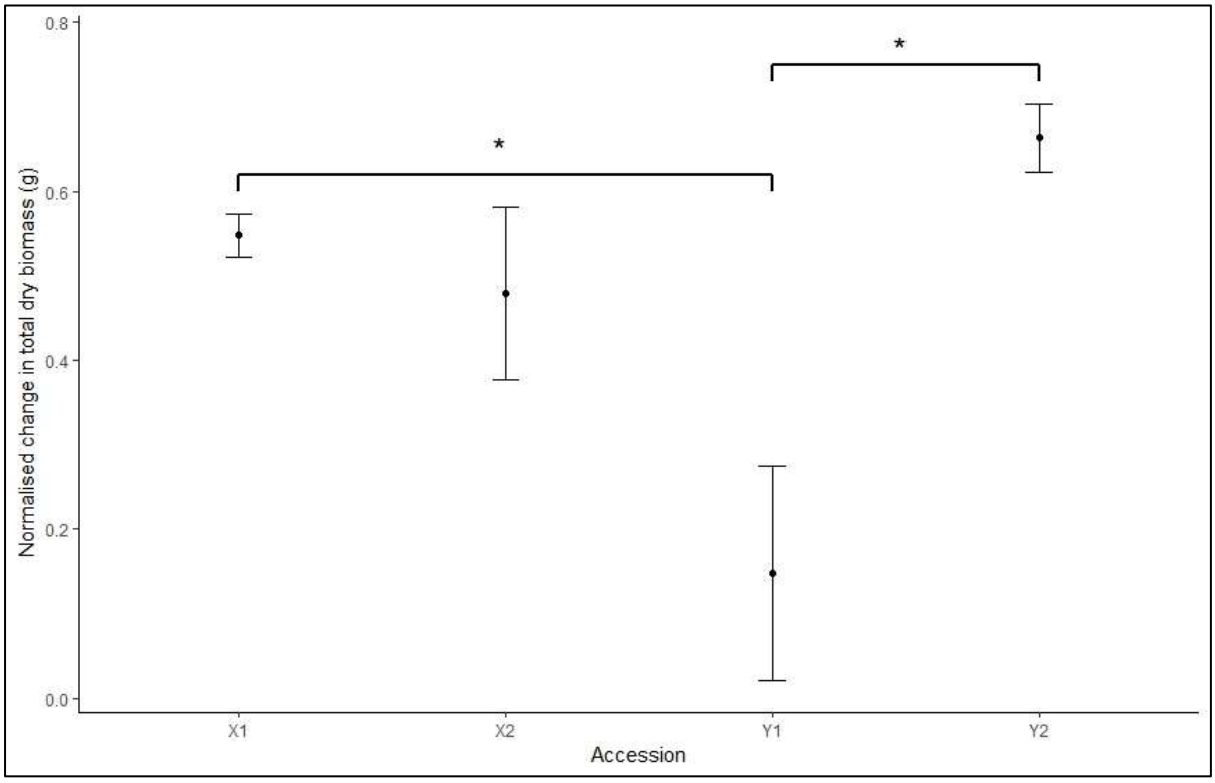


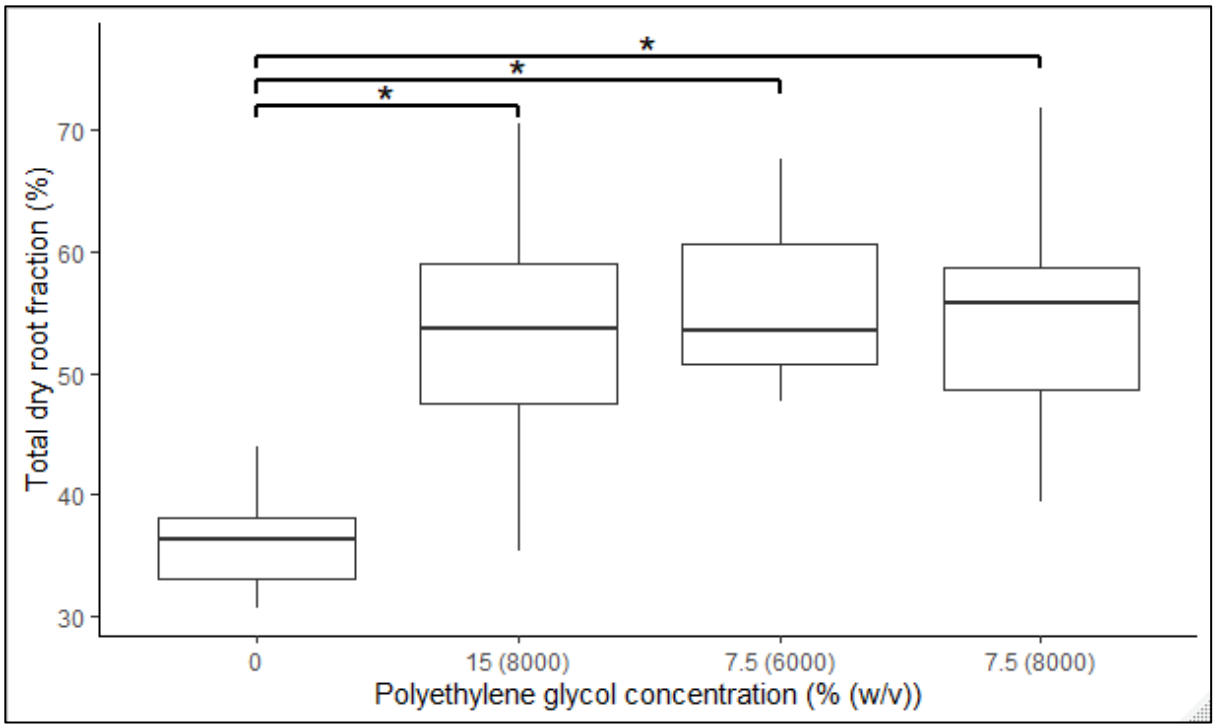
Figure 3.2: Pilot 1; Normalised change in total dry biomass of landraces of *T. aestivum* grown in static hydroponic systems containing 55% Hoagland's solution, 45% water and 7.5% (w/v) polyethylene glycol 6000. X1, N=4; X2, N=4; Y1, N= 3; Y2, N=4, $p<0.05$ (*).

3.3.2 Pilot 2 – PEG 8000 Experiment

Relative to the control group, total dry biomass was significantly reduced at 7.5% PEG6000 and 15% PEG8000 (Figure 3.3B, Kruskal-Wallis: $\chi^2 = 22.712$, d.f. = 3, $p = 4.638e^{-05}$), however the 7.5% PEG8000 treatment failed to demonstrate a statistically significant decrease in total dry biomass. Relative to the control group, total dry root fraction was significantly increased at all concentrations of PEG (Figure 3.3A, Kruskal-Wallis: $\chi^2 = 33.5$, d.f. = 3, $p = 2.526e^{-07}$).

Also, of note was that no significant differences were found between landraces within a given treatment box, for any of the measured traits – except for dry total biomass in the 7.5% PEG8000 treatment (ANOVA: F = 4.081, d.f. = 3, p = 0.0356). It was observed that the sand-compost germination method produced seedlings of substantially further developed form, appearing taller and with more elaborate and extensive root systems.

A



B

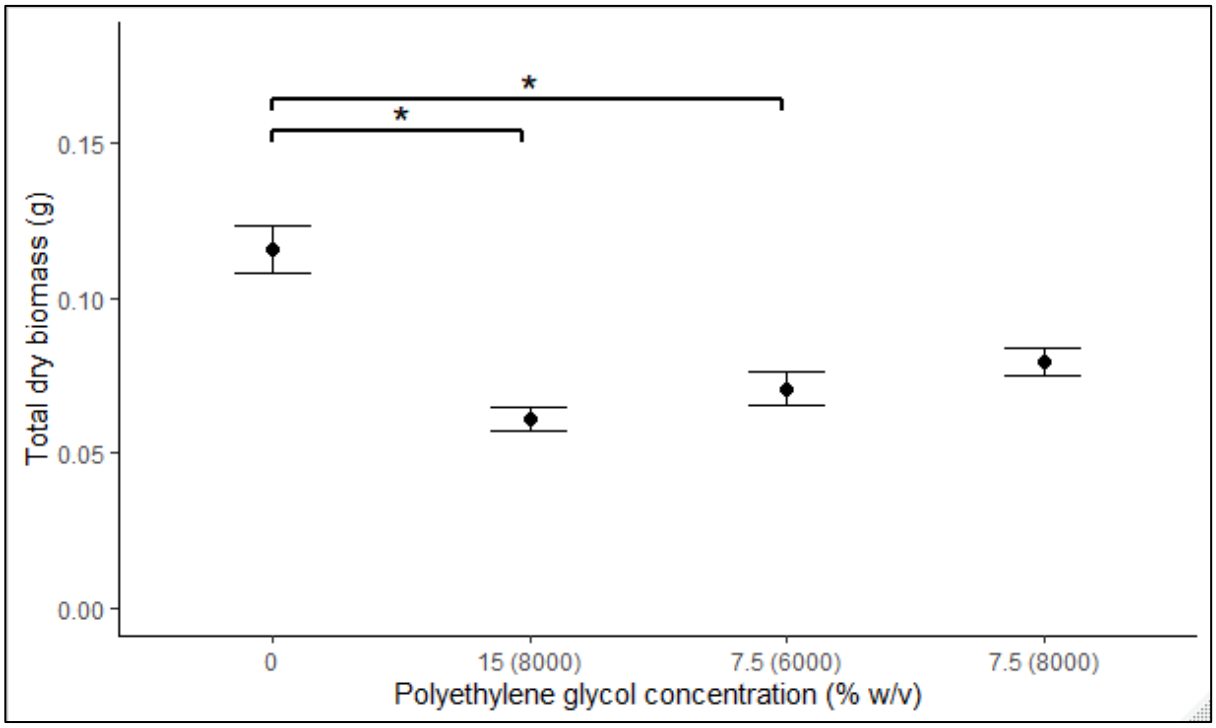


Figure 3.3: Pilot 2 plots. (A) Total dry biomass of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water and varying polyethylene glycol concentrations (0%, 15% PEG8000, 7.5% PEG8000, 7.5% PEG6000). (B) Total dry root fraction of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water and varying polyethylene glycol concentrations (0%, 15% PEG8000, 7.5% PEG8000, 7.5% PEG6000). 0%; N = 15, 15% PEG8000; N = 14; 7.5% PEG6000, N = 15; 7.5% PEG8000, N = 15. $p < 0.05$ (*).

3.4 Discussion

Since treating the plants at a concentration of 7.5% PEG6000 (w/v) appeared to meet the outlined morphological response criteria (reduced total dry biomass, increased root fraction and cessation of primary root growth), and returned differential trait values from

the landraces used, it was deemed an appropriate concentration for use in the subsequent Associative Transcriptomics experiment.

It was observed that a 5.0% PEG6000 treatment did not significantly reduce total dry biomass, whilst a 2.5% PEG6000 treatment did significantly reduce total dry biomass. This can perhaps be explained as unrepresentative variability in phenotypic plasticity, and a repeat of this experiment with increased sample sizes, for example 10 samples per treatment, would demonstrate that it is in fact more likely for these landraces to exhibit a statistically significant reduction in total dry biomass. An alternative explanation would be that stochastic growth of microorganisms in the hydroponic growth media, as growth was observed (Supplementary Figure 3), may have been responsible for, or attributed to, this lack of statistical difference.

The second pilot experiment produced results indicating that a 7.5% PEG8000 treatment was not sufficient to reduce total dry biomass of samples compared to control samples (Figure 3.3B). Given the literature surrounding the purported lack of difference between similar concentrations of PEG6000 and PEG8000 as a hyperosmotic stress agent, and the manner in which the PEG8000 treatment did appear to cause some decrease in total dry biomass (Figure 3.3), this lack of decrease was again attributed to unrepresentative variability in phenotypic plasticity. This was supposed with the expectation that the use of PEG8000 at a concentration of 7.5% (w/v), in the later AT experiment, would demonstrate that the treatment does in fact produce a statistically significant decrease in total dry biomass relative to the control group. Indeed, this turned out to be the case.

It is worthy of note, that studies which utilise PEG as a hyperosmotic stress agent to simulate severe stress in hydroponically cultured plants, use aquatic air pumps. These air

pumps act to oxygenate the growth media, since the addition of PEG in such high quantities (e.g. 15%, 20% or 25% (w/v)) increases the viscosity of the media, reducing the diffusion capacity for oxygen and so reduces the oxygen status of the media. Since this experiment utilised such low concentrations of PEG, aquatic pumps were not utilised.

Of mention was the difficulty encountered when using Whatman paper as a substrate for germinating seedlings, chiefly due to the penetration of the paper by emerging radicles. Subsequent development of the roots, which anchored the seedlings to the substrate, made the eventual transferal of seedlings to hydroponics, without substantial damage, a difficult and time-consuming activity, inappropriate for the upscaling required for the planned Associative Transcriptomics experiment. Thus, a revision of this portion of the methodology was required prior to the Associative Transcriptomics experiment. Unfortunately, many of the issues encountered with the initial germination method involving Whatman paper, such as time-cost and root damage, were similarly encountered with the sand-compost method that was employed in the second pilot. In order to mitigate these issues, a third germination technique was devised in which, following sterilisation, seeds are laid to water dampened 2 ply blue roll tissue paper. The paper is folded four times prior to insertion in the petri dish, a characteristic that along with the small grooves of its surface, act to encourage root growth without the penetration of the tissue substrate. This method was attempted and found to be successful before beginning the AT experiment.

Landraces used in these experiments were selected on the basis of their countries of origin receiving either exceptionally high (Bolivia and Colombia, Y1 and Y2) or exceptionally low amounts (Afghanistan, X1 and X2) of annual precipitation. This was done so as to assess the efficacy of the country of origin in predicting a landraces productivity in a hyperosmotic stress background. However, it was found that Y1, a

landrace predicted to be less well adapted to hyperosmotic stress conditions, in fact demonstrated a significantly lesser decrease in productivity on exposure to 7.5% PEG than X1 and Y2 (Figure 3.2). This could serve as evidence that country of origin is a poor indicator hyperosmotic stress tolerance. It could also be considered that this points to the fact that the methods used to simulate hyperosmotic stress conditions in this study, such as the use of hydroponics, are so removed from natural soil-based conditions that landraces are not performing as they would in wild-type conditions. It is key, when considering the findings of this study, that little information was available to corroborate the performances of landraces in vitro, with their performance in wild-type stressed conditions (e.g. soil).

Hence, taken together the two pilot experiments provided evidence that the devised static hydroponics system was capable of supporting *T. aestivum* canonical growth, and that a 7.5% PEG treatment is sufficient to produce landrace-specific differences in measured biomass traits. Thus, a 7.5% PEG8000 treatment was adopted, coupled with the above-mentioned germination method, to conduct a phenotypic screen of hyperosmotic stress responses for 135 *T. aestivum* YoGI panel spring landraces.

4. Phenotyping and Associative Transcriptomics analysis

4.1 Introduction

In this study Associative Transcriptomics was applied in the analysis of *T. aestivum* resistance to PEG induced osmotic stress conditions by sampling a subset of the YoGI panel (135 spring lines), and using relative normalised changes in biomass as a measure of a plants ability to cope with hyperosmotic stress conditions. Plants were hydroponically cultured and measured for their adaptive morphological responses to PEG8000 (7.5% w/v) induced osmotic stress, namely anatomical biomass accumulation. Prolonged application of PEG was used to simulate chronic moderate hyperosmotic stress - this treatment was deemed appropriate following the two pilot experiments completed as part of this investigation.

Biomass accumulation was expected to vary between landraces, with relatively hyperosmotic stress tolerant lines capable of better mitigating the effects of hyperosmotic stress and so maintaining higher rates of productivity and thus accumulating more biomass. As mentioned previously, YoGI transcriptome data is derived from leaf samples grown under normal conditions (see section 2.3). Hence, genetic markers identified within this study by AT, are not necessarily those responsible for high productivity trait values, but instead significantly associate with normalised changes in biomass trait variability and so are likely to contribute towards that variability. Since these genetic markers belong to transcriptome profiles that are not exposed to hyperosmotic stress, they are here considered as important elements of transcriptional profiles that contribute to phenotypic profiles which then feed in to variation in adaptive morphological responses to hyperosmotic stress.

Associative Transcriptomics single nucleotide polymorphism (SNP) and gene expression marker (GEM) analysis produces Manhattan plots which facilitate the isolation of genomic regions of interest for further interrogation; yet this requires suitable candidate genes from these regions to be identified. Associative Transcriptomics acts to identify genetic markers which significantly associate with phenotypic variation so as to explain the distribution of observed phenotypic variance. Therefore, it cannot be definitively stated how the genetic markers identified by AT analysis of the data derived from the hyperosmotic stress panel screen conducted in this study, contribute to phenotypic variation. Thus, in order to effectively discern appropriate candidate markers, particularly from association analyses, a sufficient understanding of the emerging models surrounding plant responses to osmotic stress and more generally abiotic stress, proves particularly useful.

4.2 Methods

4.2.1 Osmotic stress tolerance phenotyping

135 of the YoGI diversity panel's spring wheat lines were hydroponically cultured and measured for their adaptive morphological responses to polyethylene glycol (PEG) induced hyperosmotic stress (normalised change in biomass: fresh root, fresh shoot, fresh total, dry root, dry shoot, dry total, fresh root fraction, dry root fraction; see section 2.4). For each landrace, four biological repeats were cultured in each condition – meaning a total of 1080 samples were transferred to hydroponics, requiring eighteen hydroponics boxes.

Seeds were sterilised using a triton-bleach treatment, laid to folded H_2O dampened 2 ply blue roll tissue paper, stratified for 24 hours at 4°C, germinated in a controlled growth room at room temperature under a 90 $\mu mol\ m^{-2}s^{-1}$ 16h/8h light/dark regimen for 6 days. Subsequently, seedlings of similar phenological stage were transferred to hydroponics.

As a matter of workload, these eighteen hydroponics boxes were split into three groups of six; it was in these groups that the experiment was staggered, with each of the three groups placed into glasshouse conditions five days apart. Each staggered group contained six boxes, three of which served as no treatment controls, and three which received a 7.5% PEG8000 treatment. Each staggered set of six boxes contained at least one replicate of 135 landraces in each condition (see Chapter 2.3 for details). The remaining sample for each landrace, for each condition, was randomly allocated to one of the three staggered sets of boxes. The randomised block design employed in this experiment acts to spread the samples across all box groups, and so acts to prevent these potential differences incurring misrepresentative trait measurements for certain landraces that would produce false associations in the AT analysis.

Seedlings were cultured for 14 days in hydroponics prior to the measurement of biomass trait values. Samples which received a PEG treatment were cultured in hydroponics for 24 hours prior to the introduction of the stress agent. This acclimation step was introduced in an attempt to reduce the rate of sample death as a result of the inherent stresses associated with such a transition of environmental conditions. Hence, plants received a hyperosmotic stress treatment for 13 days. Growth media was changed every 7 days and managed as outlined in Chapter 2.3.

4.2.2 Single Nucleotide Polymorphism (SNP) marker analysis

Biomass data taken from the panel screen experiment was used to calculate normalised change in biomass trait values for each landrace, as described in section 2.4. Each of these sets of trait values was utilised in an association analysis to identify single nucleotide polymorphism markers that significantly associated with the variation within a given set of trait values (Harper et al., 2012).

This analysis used a mixed-liner model, which incorporated population structure inference using kernel-PCA and optimization (PSIKO) (Popescu et al., 2014) as a fixed effect and relatedness via a kinship matrix generated by GAPIT (Zhang et al., 2010) as random effect, so as to reduce the likelihood of generating false positive results. Marker-trait association test scores were generated in R version 3.5.1 using GAPIT version 3.0 (Zhang et al., 2010) and results for markers with minor allele frequencies greater than 0.05 were visualised using Manhattan plots produced using Grapher, in R version 3.5.1 (R script courtesy of A. Harper).

Grapher plots for each trait analysed were qualitatively assessed for the presence of tall well-defined peaks produced by the distribution of data points. Manhattan plots for each trait analysed were qualitatively assessed for the presence of tall well-defined peaks produced by the distribution of data points. Peak regions were arbitrarily defined as those which contained as at least 10 data points with $-\log_{10}P$ values equal to or greater than 2. Boundaries of peak regions were defined on an individual basis as the most peripheral Grapher data points on the left and right side of a given peak. Peaks which satisfied qualitative assessment and also contained at least one marker with a false discovery

rate-adjusted p-value lower than 0.05 were investigated for candidate genes using IWGSC functional annotation information.

4.2.3 Gene expression marker (GEM) analysis

The normalised biomass trait data utilised in the SNP analysis, was utilised in a simple linear model which incorporated population structure inference using kernel-PCA and optimization (PSIKO) (Popescu et al., 2014) as a fixed effect, to identify gene expression markers that significantly associated with the variation within a given set of trait values. Marker-trait association test scores were generated and plotted in R version 3.5.1 using R scripts (Regress Multi, Regress Plotter) courtesy of A. Harper. Markers with false discovery rate marker-trait association test p-values less than 0.05 were taken forward as potential candidates for further investigation.

4.3 Results

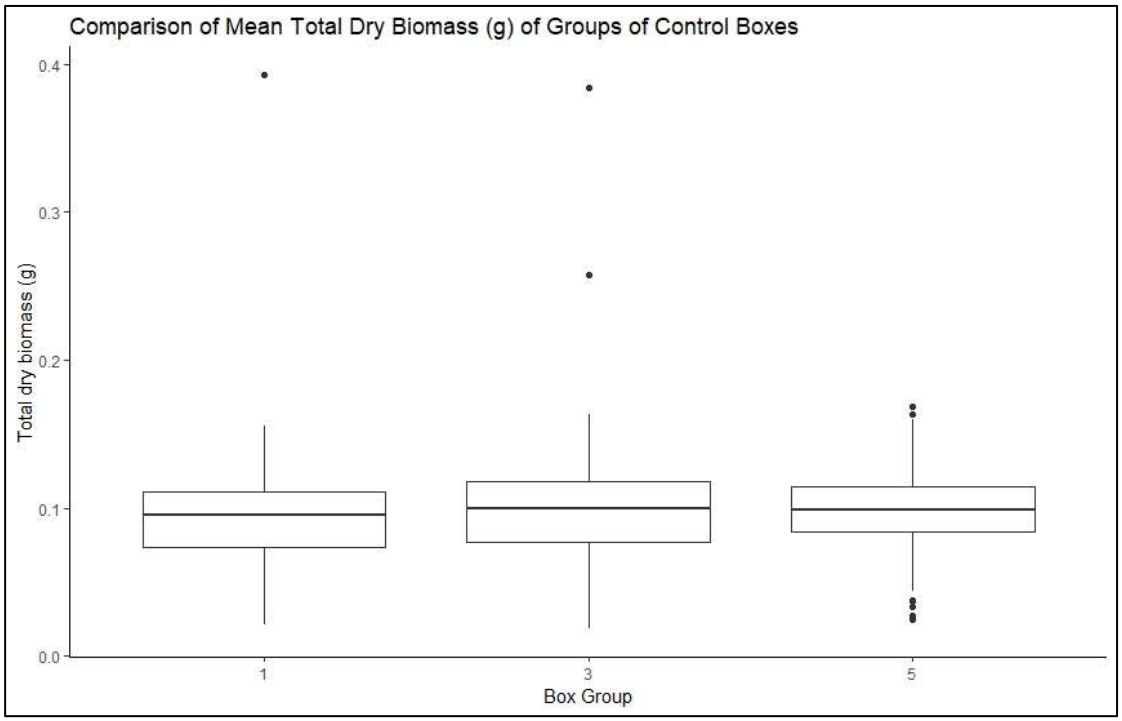
4.3.1 Osmotic stress tolerance phenotyping results

Thirty samples died over the course of the experiment and were therefore excluded from the biomass measurements. All landraces had at least two samples that survived in each condition.

Of interest was the relative performances of samples according to their box group (e.g. 1 and 2 versus 3 and 4 etc.), as despite the partial control of environmental conditions such as light and dark periodicity, temperature and humidity were capable of fluctuating over time. Box groups were also placed at three separate locations within the same glasshouse, thus receiving varying amounts of light. There was no significant difference found between the control boxes for total dry biomass (Kruskal-Wallis, χ^2= 5.0306, d.f. = 2, p-value = 0.08084, Figure 4.1A). Total dry biomass was found to significantly differ between treatment boxes, with box group 4 displaying significantly higher total dry biomass and the other two box groups displaying no significant difference from each other (Kruskal-Wallis, χ^2 = 22.099, d.f. = 2, p-value = $1.589e^{-05}$, Figure 4.1B). Similarly, there was no significant difference found between the control boxes for dry root fraction (ANOVA, F = 0.333, d.f. = 2, p-value = 0.717, Figure 4.2A), but dry root fraction was found to significantly differ between treatment boxes, with all groups displaying significant differences from all other groups (ANOVA, F = 14.17; d.f. = 2; p-value = $1.02e^{-06}$, Figure 4.2B).

However, when comparing control samples and PEG treated samples across the whole experiment, total dry biomass was found to be significantly lower in PEG treated samples than in control samples (T-test, t = 22.783, d.f. = 1048, p-value < $2.2e^{-16}$, Figure 4.3A). Also, dry root fraction was found to be significantly higher in PEG treated samples than in control samples (Kruskal-Wallis, χ^2 = 700.89, d.f. = 1, p-value < $2.2e^{-16}$, Figure 4.3B).

A



B

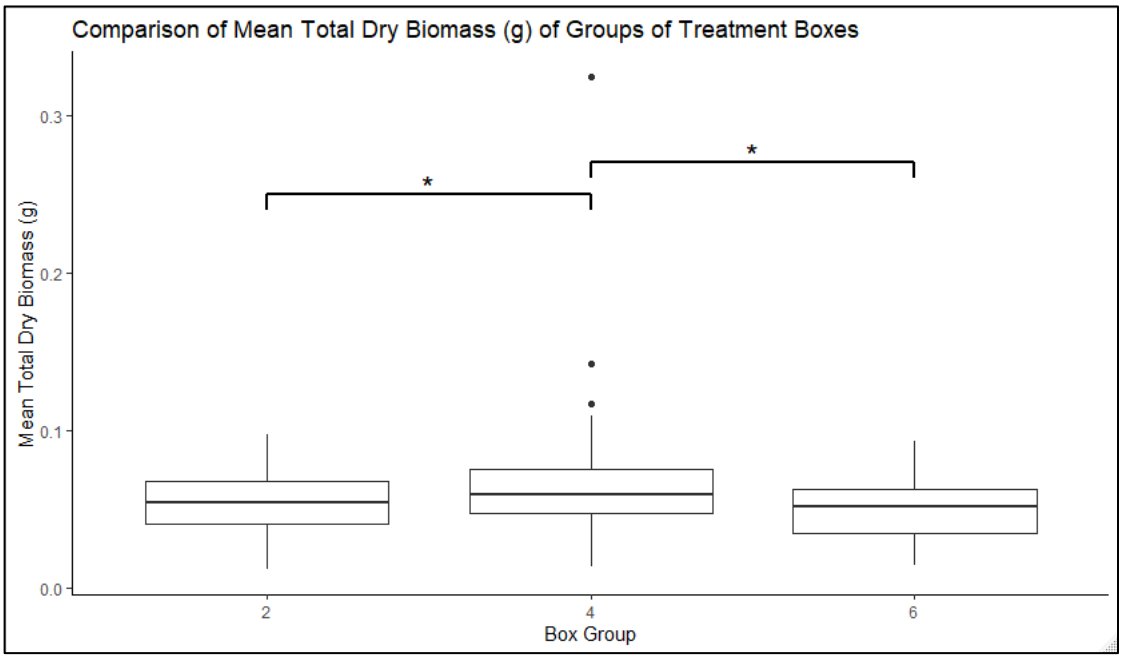
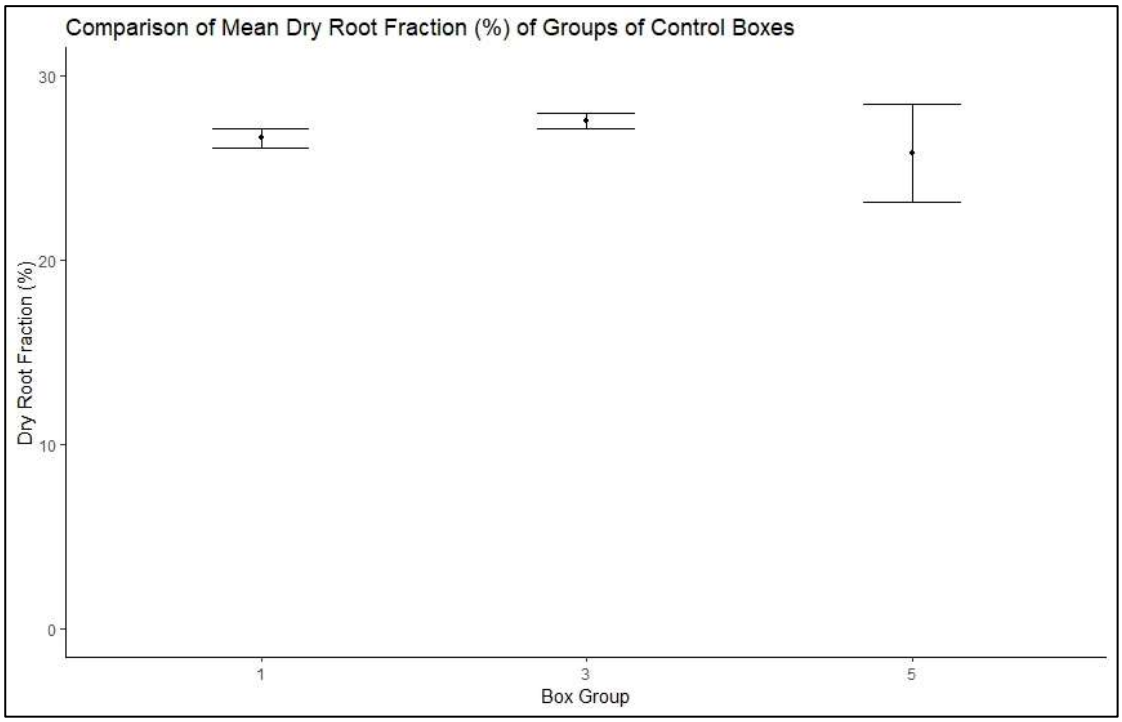


Figure 4.1. Box plots comparing mean total dry biomass of box groups of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (A) and supplemented with 7.5% (w/v) polyethylene glycol 8000 (B). A: Group 1, N = 180; Group 3, N = 179; Group 5, N = 174. B: Group 2, N = 175; Group 4, N = 178; Group 6, N = 164 (Kruskal-Wallis Test, $p < 0.05$ (*)).

A



B

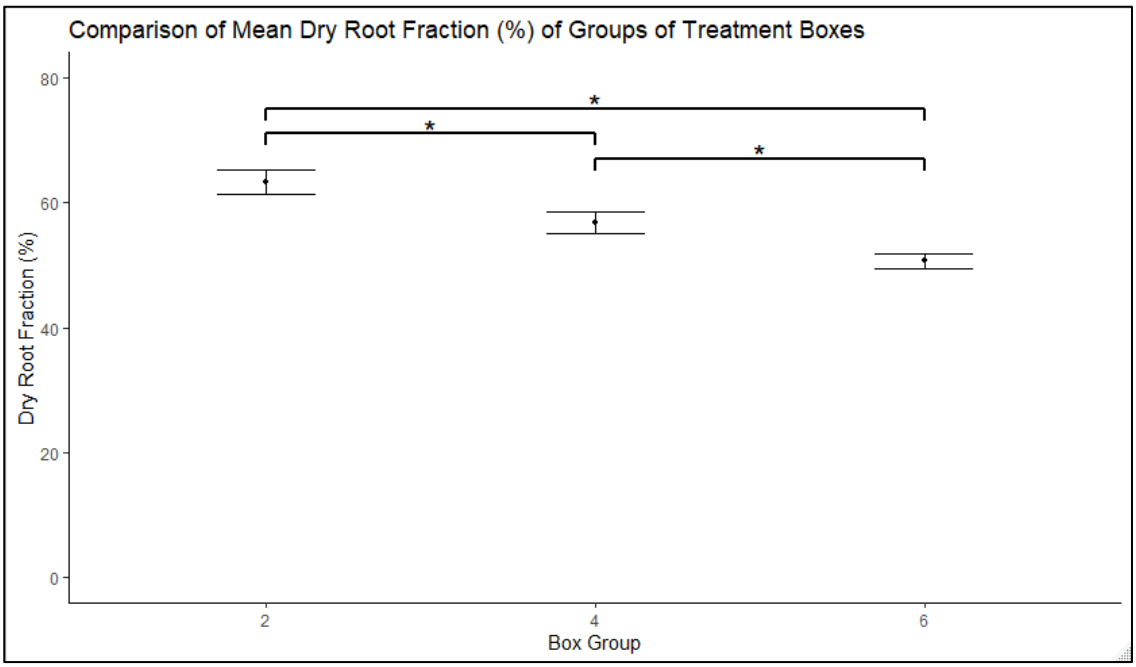
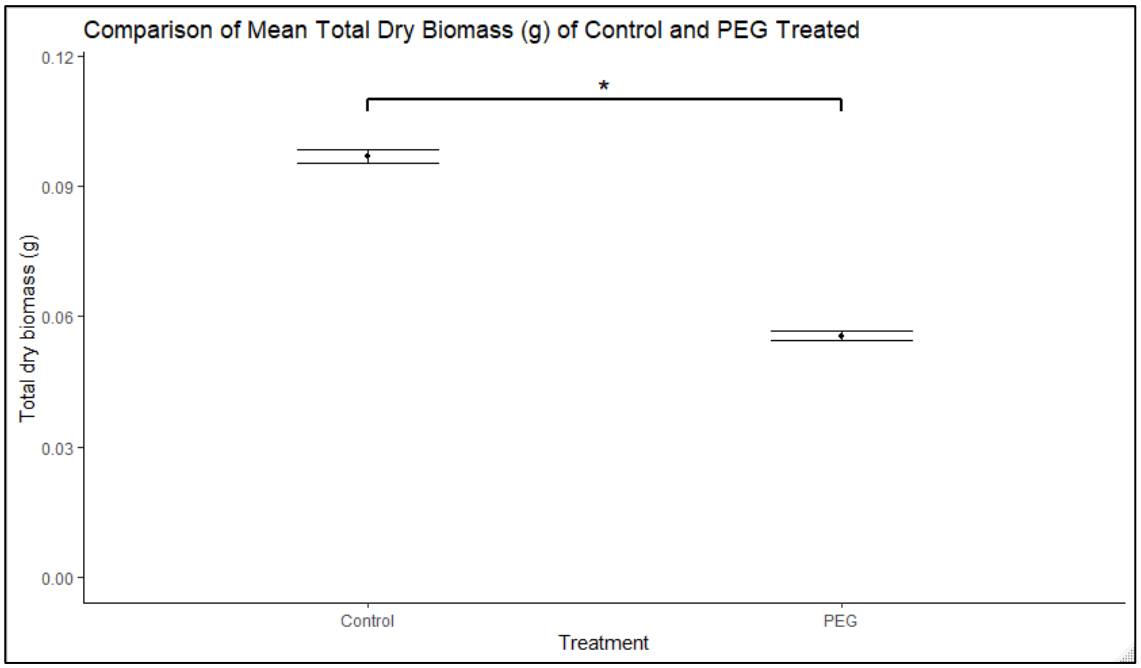


Figure 4.2. Plots comparing mean dry root fraction of box groups of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (A) and supplemented with 7.5% (w/v) polyethylene glycol 8000 (B). A: Group 1, N = 180; Group 3, N = 179; Group 5, N = 174. B: Group 2, N = 175; Group 4, N = 178; Group 6, N = 164 (ANOVA, $p < 0.05$ (*)).

A



B

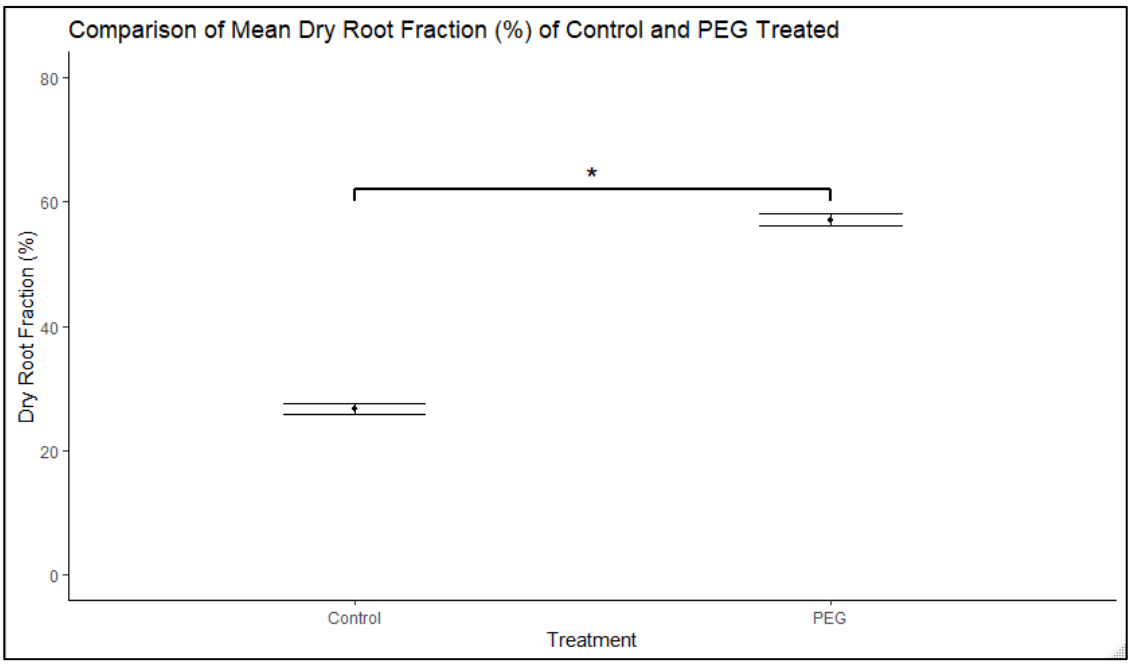


Figure 4.3. Plots comparing mean total dry biomass (A) and dry root fraction (B) of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (Control) and supplemented with 7.5% (w/v) polyethylene glycol 8000 (PEG). Control, N = 533; PEG, N = 517 (A, T-test, $p < 0.05$ (*); B, Kruskal-Wallis, $p < 0.05$ (*)).

It was again observed that the PEG treatment produced a cessation of primary root growth, concurrent with previous reports (Figure 4.4). Microbial growth was again observed in the PEG treatment boxes (Supplementary Figure 3).

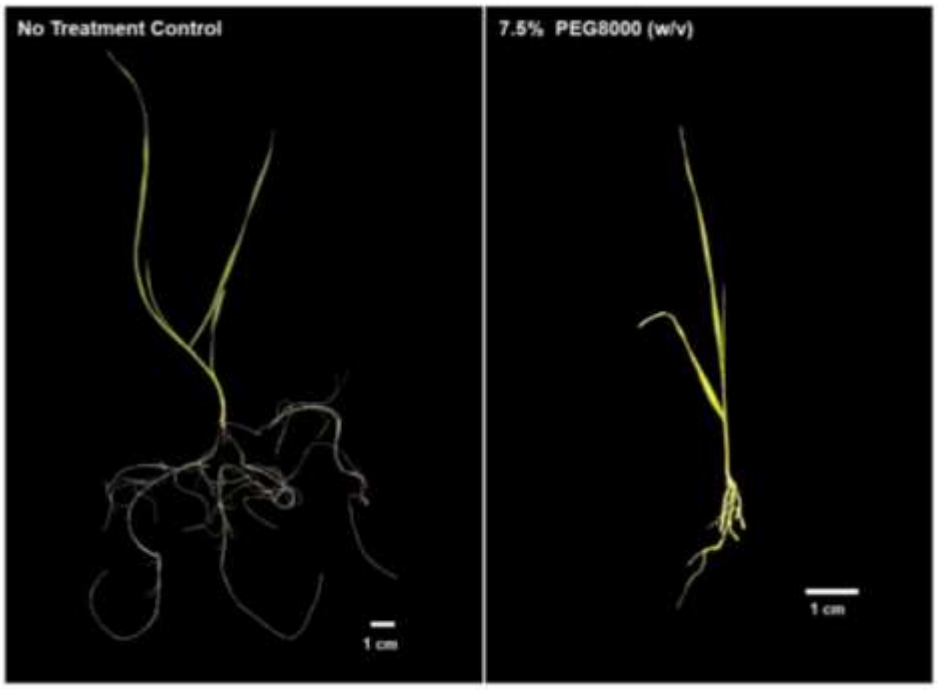


Figure 4.4: Photograph images showing the cessation of primary root growth of YoGI 310, grown in a static hydroponic system containing 55% Hoagland's solution, 45% water (left) and supplemented with 7.5% PEG8000 (right).

Taking normalised change in total dry biomass as a measure of a plant's productivity and efficacy in mitigating the effects of the hyperosmotic stress, variation was seen across the panel subset investigated (Figure 4.5).

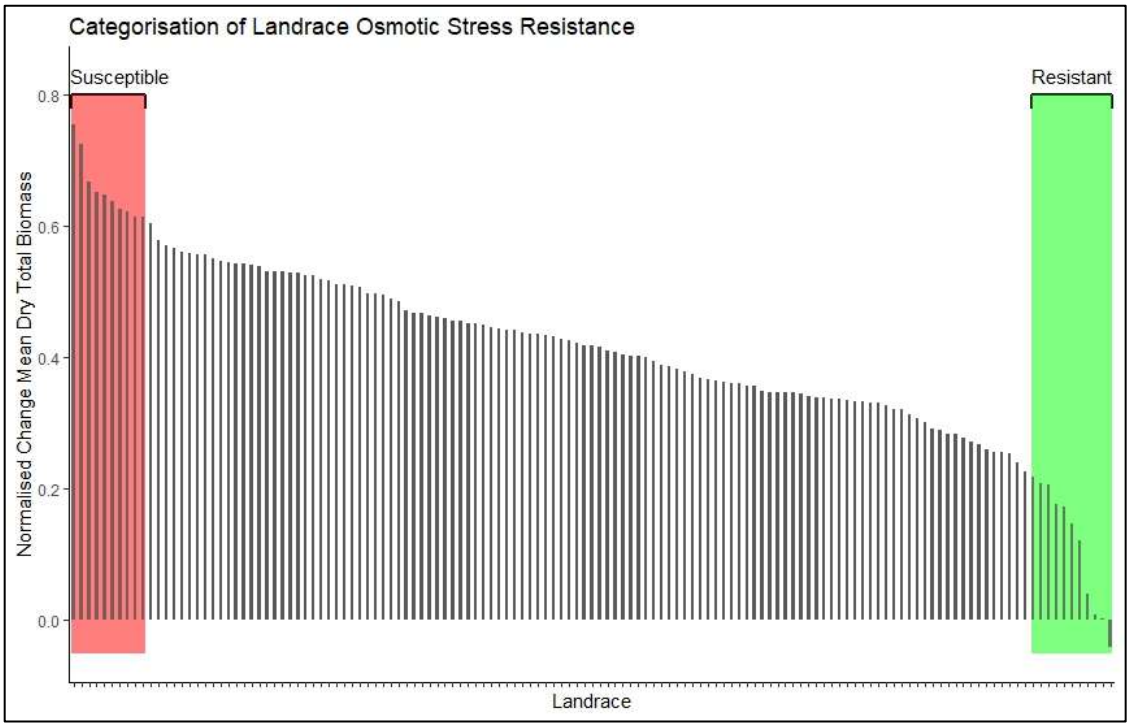


Figure 4.5: Plot showing the normalised change in mean dry total biomass for 135 *T. aestivum* YoGI landrace lines grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (control) and 7.5% polyethylene glycol 8000 (w/v; treatment). Boxes represent the ten landraces most susceptible (red) and most resistant (green) to decrease in total dry biomass as a result of hyperosmotic stress.

4.3.2 SNP marker analysis results

Normalised change in fresh shoot biomass (NCFRB), total fresh, dry root and total dry biomass trait data did not return any markers with false discovery rate adjusted (FDR) p-values lower than 0.05 and so were disregarded in further analysis (Supplementary Figure 2). Trait data for normalised change in fresh root and dry shoot biomass and fresh and dry root fraction returned 157, 71, 87 and 33 markers respectively with FDR p-values lower than 0.05 (Figure 2).

Only one peak met with the stringent criteria for selection of putative candidate markers outlined above (Figure 4.6). The peak, located on chromosome 6A, was generated from fresh root data and contained two SNP markers with marker-trait association test scores lower than 0.05 (Table 4.1).

Table 4.1: GAPIT results for the two SNP markers contained within the association peak region used to identify the aquaporin candidates. Results displayed include false discovery rate adjusted p-values and calculated effect size.

SNP	FDR Adjusted P-values	Effect
TraesCS6A01G406600.3:1420:G	2.87E-08	-0.033214125
TraesCS6A01G406600.3:519:C	0.001124035	-0.113839987

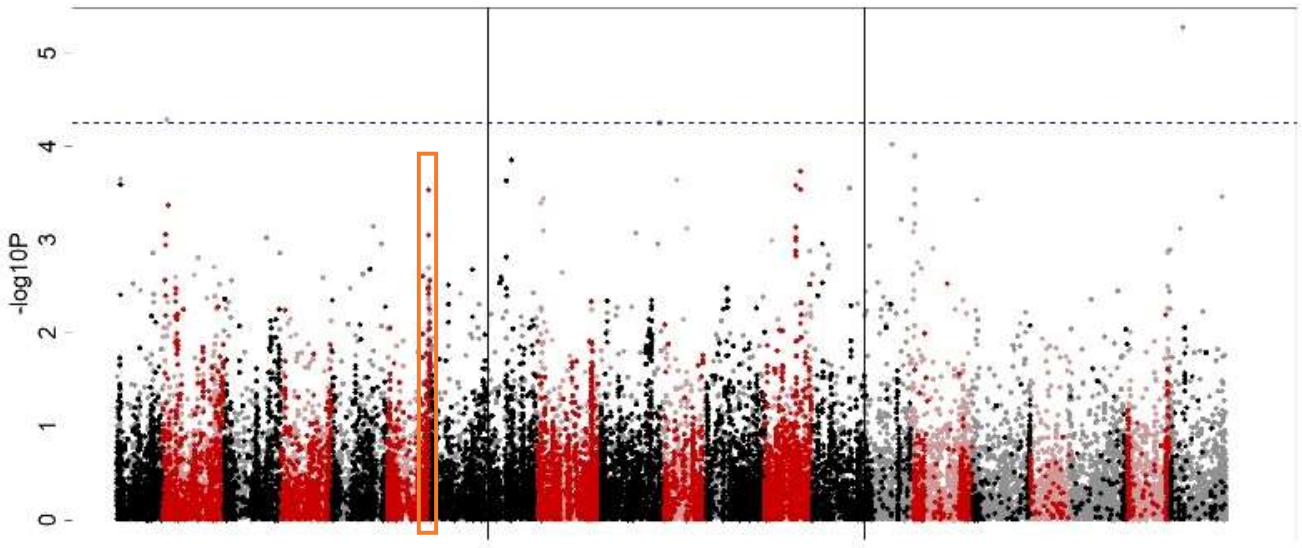


Figure 4.6: Manhattan plot depicting SNP marker trait association test -log10 p-values for normalised change in fresh root biomass trait data. Analysis was performed and results generated in R 3.5.1 using GAPIT and subsequently plotted using Grapher. Orange box indicates ‘chromosome 6A peak’.

Measured fresh root biomass spanned a range of 0.884g (3 s.f.; 0.0057g – 0.89g). Of this variation, the two SNPs found within the peak highlighted in Figure 4.6 and shown in Table 4.1, were found to have a combined effect size of 0.147 (3 s.f.). Therefore, these two SNPs were found to be explain 0.129g of the variation in fresh root biomass.

The peak region spanned a distance including 22 genes (TraesCS6A01G404700.1-TraesCS6A01G406600.1). This region was examined using the IWGSC provided functional gene annotation information in order to establish whether the region contained any putative genes with functions that could have contributed to the variation in the normalised change in fresh root biomass trait data. Of the annotations within this region, two putative genes annotated as 'plasma membrane intrinsic proteins/ aquaporins' were taken as the most probable contributors primarily responsible for this association peak (TraesCS6A01G405600.1_1 and TraesCS6A01G405700.1_1).

4.3.3 Gene expression marker (GEM) analysis

GEM marker analysis found that normalised change in fresh shoot, total fresh, dry root, dry shoot and total dry biomass trait data, as well as dry root fraction and fresh root fraction trait data, did not return any markers with FDR-adjusted p-values lower than 0.05 (Supplementary Figure 5). However, analysis of GEM markers for normalised change in fresh root biomass found 21 markers to have FDR-adjusted p-values lower than 0.05 (Figure 4.7, Supplementary table 2).

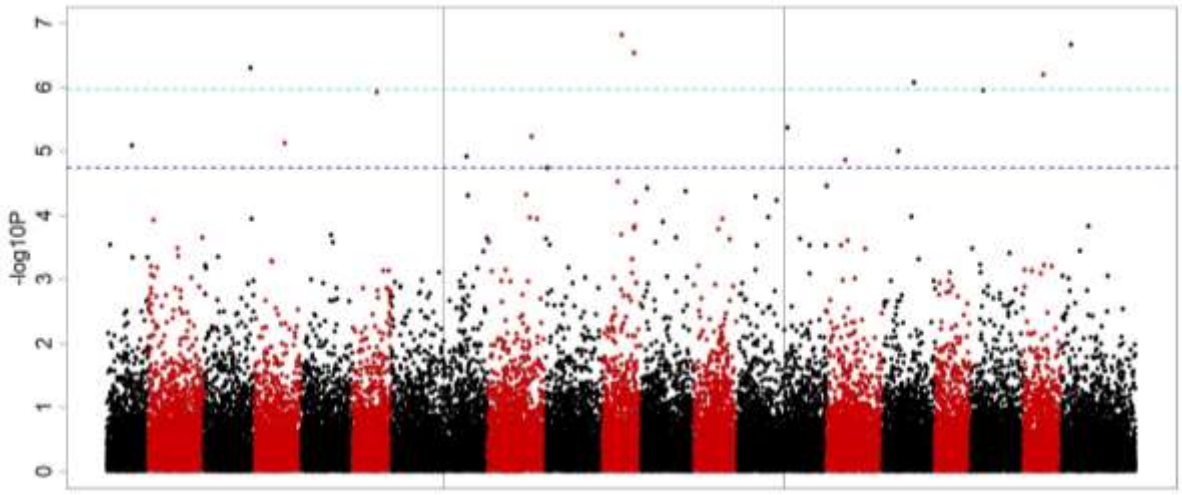


Figure 4.7: Manhattan plot depicting GEM marker trait association test -log10 p-values for normalised change in fresh root biomass trait data. Analysis was performed and results generated in R 3.5.1 using Regress Multi and subsequently plotted using Regress Plotter. Blue line indicates the FDR threshold, cyan line indicates Bonferroni threshold.

Taking existing IWGSC functional annotation information for these markers, it can be seen that the majority of markers (18) have previously been linked to drought or osmotic stress responses. The remaining three markers (TraesCS2B01G205800.1, TraesCS1D01G037500.1 and TraesCS3D01G176100.1) are linked to disease resistance.

Eight candidate genes were selected from the twenty-one ‘FDR makers’ on the basis of the stronger amount of evidence linking their function to drought and hyperosmotic stress responses. Marker sequences were compared with the current NCBI protein sequence database in an effort to better clarify the accuracy of the IWGSC functional annotation information.

Scatter plot visualisation of RPKM versus NCFRB data for the eight candidates for the subset of 135 landraces found that for many of the genes, the relationship between a given gene's expression and NCFRB variation was dependent upon YoGI 28, a landrace with an exceptionally low NCFRB trait value (Supplementary Figure 6 and Supplementary Figure 7). TraesCS4B01G319100.1 did not appear to exhibit this effect so strongly (Figure 4.8) and is annotated as a Heat Shock Protein 40 family member. A linear regression analysis revealed this relationship to be statistically significant ($F = 16.14$; d.f. = 1, 132; $p = 9.805e^{-05}$) although the association was not significant when YoGI 28 was removed from the analysis ($F = 1.521$; d.f. = 1, 131; $p = 0.2196$).

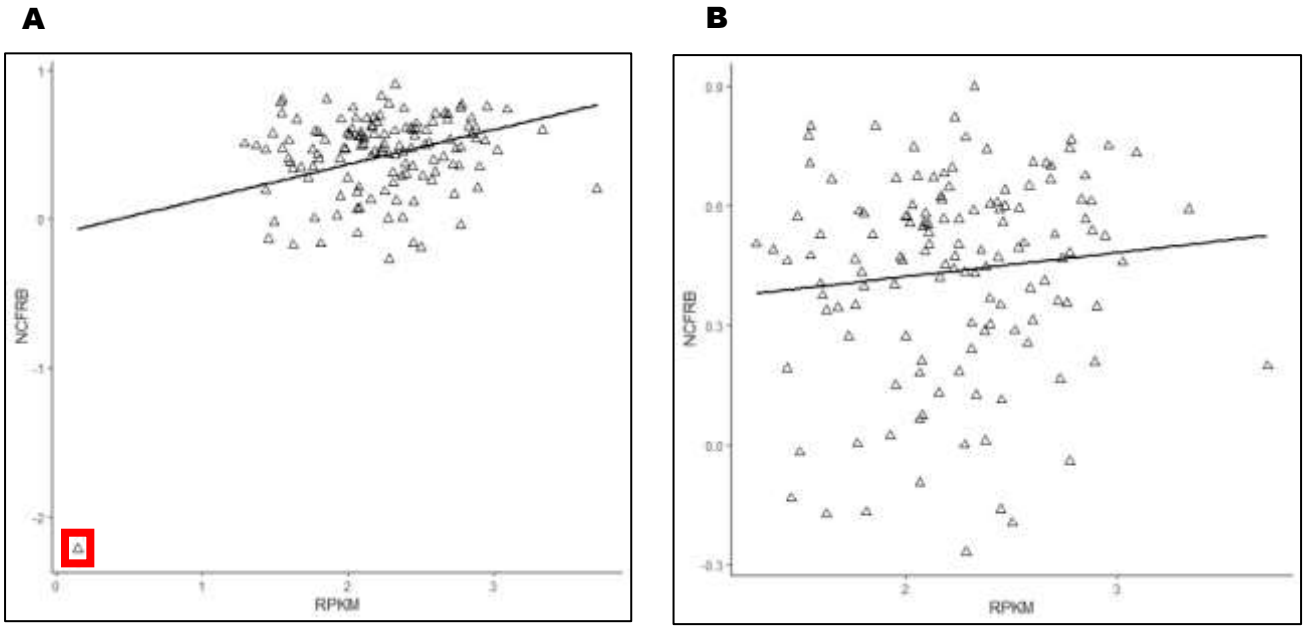


Figure 4.8: Scatter plots showing the relationship between RPKM and normalised change in fresh root biomass for TraesCS4B01G319100.1 for the 135 landraces measured in the panel screen experiment. YoGI 28 data highlighted in red

4.4 Discussion

4.4.1 Hyperosmotic stress tolerance phenotyping

Despite the identified differences discovered between box groups (Figure 4.1 and Figure 4.2), these differences were not so substantial as to affect the expected morphological responses of PEG treated plants relative to control plants (decreased productivity, increased root fraction etc.). Differences in total dry biomass seen between treatment box groups could have been partially owed to a variation in glasshouse conditions in the time periods that each group was cultured in hydroponics. It is recommended that growth cabinets be instead used, providing a greater degree of control over environmental conditions such as temperature and atmosphere conditions. Alternatively, these measured differences in productivity could be owed to the observed differential growth of microorganisms in the growth media (Supplementary Figure 3). However, the phenotyping experiment successfully produced the expected morphological responses from landraces, that were outlined in the pilot experiment (Figure 3.4).

The panel screen experiment was not repeated and so the replicability of these results is, so far, difficult to assert. Associative Transcriptomics requires phenotypic variation within the trait data used for its analysis; therefore, variation such as that seen in Figure 4.5, supported the use of the gathered trait data in an AT analysis to identify genetic markers associated with this identified variation.

4.4.2 Associative Transcriptomics SNP analysis

Genomic regions containing SNPs or combinations of SNPs that produce advantageous phenotypes are expected to be in linkage disequilibrium with surrounding markers, represented on the Grapher Manhattan plots as distinct peaks. 346 of the 348 SNP markers identified by this investigation as having marker-trait association test FDR-adjusted p-values lower than 0.05 did not fall within genomic regions that contained distinct Manhattan plot peaks. Since they did not belong to a peak region, they were considered to likely be statistical anomalies, unrepresentative of underlying biological phenomena. Therefore, pursuing these markers in further analysis was deemed an unprofitable expenditure of resources.

Only one genomic region was identified as having sequence variation associated with the calculated biomass variation measured for the 135 landraces utilised in the panel screen experiment (Figure 4.6). Genome wide association techniques such as AT are more powerful in identifying the underlying genetic variation responsible for trait variation when there are fewer factors responsible for controlling the trait in question. A trait such as the normalised change in biomass for an anatomical region in response to hyperosmotic stress, is expected to be under the control of a vast suite of genes. Hence, it is to be also expected that few genetic markers are likely to be identified when applying such a technique to the data such as that gather in this study; the results of this analysis give credence to that notion.

The aquaporins were selected as likely candidates on the basis of the literature purporting a functional role for aquaporins in hyperosmotic stress responses. Aquaporins, integral membrane channel proteins that facilitate the transport of water across

phospholipid membranes have become of great interest in the study of plant water relations, particularly in response to water deficient conditions. Aquaporins have well established functional roles in maintaining cellular water homeostasis (Verkman, 1992), are conserved in plants, animals, yeast and bacteria (Finn et al., 2014), and also transport carbon dioxide, boron, glycerol and other small uncharged molecules (Durbak et al., 2014; Groszmann et al., 2017; Luu and Maurel, 2005; Zeuthen et al., 1999). Contemporary research postulates a role for aquaporins in drought response (Li et al., 2015), supported by the tight regulation of their expression and activity in response to abiotic stresses such as osmotic and salinity stress (Alexandersson et al., 2005; Horie et al., 2011). Functionally, aquaporins are considered to contribute to whole plant hydraulic conductance and expression studies purport a tissue specific pattern of activity (Matre et al., 2002; Alexandersson et al., 2005). As a result, aquaporins are considered key players in the mitigation of hyperosmotic stress conditions.

Protein sequence similarity was compared between the two aquaporin sequences using EBI's Clustal Omega identity matrix tool, finding the sequences to share 98.63% sequence identity. NCBI protein BLAST results found the two aquaporin protein sequences to have identities closest to that of plasma membrane intrinsic protein 1;5 (PIP1;5) in closely related species (*Triticum urartu* and *Aegilops tauschii*) and in *Arabidopsis thaliana* (Supplementary Table 1). Due to the exceptionally high sequence similarity of TraesCS6A01G405600.1 and TraesCS6A01G405700.1, the order of the top NCBI results did not differ depending on which sequence was selected.

Q-RT-PCR data from *Arabidopsis thaliana* has revealed *PIP1;5* expression in roots, leaves and flowers indicating a physiological role across the plant (Alexandersson et al., 2005). Q-RT-qPCR data from *Zea mays* found *PIP1;5* expression in leaves to be

preferentially localised to the stomatal guard cells, suggesting a functional role in regulating stomatal aperture (Heinen et al., 2014). Similar to many other PIPs, *AtPIP1;5* is substantially downregulated in response to drought (Alexandersson et al., 2005). There is yet to be a published characterisation of *Atpip1;5* morphology in response to osmotic stress.

A major challenge that faces this study lies in asserting that the methods employed in simulating hyperosmotic stress are in fact a close enough representation of the conditions that the results of the AT may be considered valid and worthy of further investigation. In an effort to accomplish this, mutant lines were analysed with the aim to demonstrate that individuals containing mutations which disrupted function in the aquaporin(s), identified by the aforementioned SNP analysis, would demonstrate a differential sensitivity to hyperosmotic stress relative to an individual lacking the mutation. This would demonstrate that the AT analysis conducted in this study is at least partially capable of identifying genetic markers that play a role in plant hyperosmotic stress responses. Therefore, these aquaporins were selected for mutational analysis.

Target Induced Local Lesions In Genome (TILLING) lines were considered as a source for mutant germplasm. Both sequences were BLAST searched against the TILLING database sequences and the putative gene with the closest sequence identity match, that belonged to the appropriate genome(6A), was considered the corresponding homologue. For both sequences there were no stop codon mutations listed and so no mutation could be readily accessed with a confidence that it would produce sufficient disruption in structure and function.

Instead, the decision was taken to work with *Arabidopsis thaliana*. Despite having a lack of agricultural significance, *A. thaliana* is a rigorously studied genetic model organism, with a relatively small (approx. 125Mb), fully sequenced genome, as well as a rapid life cycle and a prolific seed production rate. Segregating SALK mutants with T-DNA insertions in the single listed copy of PIP1;5 were available for ordering (SALK056898 and SALK056041). These mutants were ordered, sterilised using an ethanol and triton-bleach treatment, germinated and grown in compost in glasshouse conditions until they produced seed. Plants were genotyped according to Chapter 2.7. Unfortunately, due to time constraints, no further work was able to be accomplished with these lines. However, genotyping work did confirm that the F1 generation contained individuals which were homozygous for the T-DNA insertion, from which the seeds have now been collected for potential future work. Experiments similar to those conducted with the TILLING lines, discussed later in Chapter 5, could demonstrate that PIP1;5 plays a substantial role in mediating Arabidopsis morphological responses to hyperosmotic stress. If this were indeed the case then PIP1;5 would warrant further investigation.

4.4.3 Associative Transcriptomics GEM analysis

AT GEM analysis of the normalised change in biomass data for traits other than fresh root biomass, found no GEM markers with marker-trait association test scores lower than 0.05. Similar to the SNP analysis, this may be evidence of the limitations of using a technique such as AT to study such an exceptionally complicated trait as biomass accumulation in response to hyperosmotic stress. It is also here suggested that a greater degree of accuracy in biomass measurements, such as measuring weight in grams to

four decimal places rather than three, would allow for the distinguishing of numerical variation for traits such as dry root biomass - which display particularly low values.

Of the 21 markers identified in the AT GEM marker analysis as having FDR-adjusted marker-trait association test scores lower than 0.05 for fresh root biomass, three markers (TraesCS2B01G205800.1, TraesCS1D01G037500.1 and TraesCS3D01G176100.1) were linked to disease resistance using IWGSC gene annotation information and existing literature. Decreases in carbon assimilation, coupled with the partitioning of photoassimilate towards processes such as osmotic adjustment and the repair and prevention of oxidative stress, constrains the carbon resources available for investment in relatively high cost defence compounds (Herms and Mattson, 1992; Poorter and Villar, 1997; Norton et al., 2008). This creates a situation whereby plants are made more susceptible to biotic factors such as disease, placing an emphasis on the plant defence system during periods of exposure to osmotic stress. It is by such thinking that we can begin to understand how the physiological state of the whole plant, that is imposed by an abiotic stress, such as osmotic stress, can render the plant prone to other stresses, requiring a complex integration of the responses to each of these environmental conditions. It could therefore be expected that genes with functions related to disease resistance would influence biomass accumulation during a period of hyperosmotic stress by contributing to a more resilient phenotypic profile.

Further investigation of the eight strongest candidate genes produced by the AT GEM analysis, using NCBI's pBLAST search tool to produce better annotations for the candidates, revealed strong evidence for their validity:

TraesCS4D01G027600.1 showed a 100% coverage, 99.18% identity NCBI pBLAST match with cyclin dependent protein kinase 9 (CDPK9) in *Aegilops tauschii* (XM_020329660.1). rt-PCR results found *Triticum aestivum* CDPK9 expression to be responsive to drought and ABA treatments (Li et al., 2008), whilst exhibiting limited regulation in response to other stresses. CDPKs play a key role in the perception of abiotic stress, the regulation of signalling cascades and the subsequent prolongment of S-phase and a delaying of a cell's entry into mitosis (Kitsios and Doonan, 2011). Hence, the putative CDPK identified as having gene expression values that significantly associate with measured fresh root biomass data, has an exceptionally high sequence similarity with CDPK9, a growth regulating signalling molecule with a drought perceptive function.

TraesCS7D01G111600.1 showed 100% coverage, 99.63% identity NCBI pBLAST match with Heading Date 3A (HD3A) in *Aegilops tauschii* (XM_020344564.1). Encoded by the gene commonly referred to as Flowering Locus T (*FT)*, HD3A is a well characterised flowering stimulating hormone, or florigen. In *Arabidopsis thaliana*, *FT* is as a key constituent in the activation of the drought escape (DE) response, being activated by ABA under drought conditions to promote early flowering so as to escape the yield-related consequences of a water deficient environment (Riboni et al., 2013).

TraesCS3B01G013100.1 showed a 100% coverage, 92.26% identity match with UDP-glycosyltransferase in *Aegilops tauschii* (XM_020297789.1). In *Arabidopsis* expression of UDP-glycosyltransferases UGT79B2 and UGT79B3 enhanced drought stress tolerance via modulating anthocyanin accumulation (Li et al*.,* 2017). Double mutants (*ugt79b2/b3*), generated by both RNA interference and CRISPR-Cas9 strategies were more susceptible to drought stress (Li et al., 2017).

Collectively, these examples demonstrate that putative genes identified by the AT analysis have a highest protein sequence similarity with genes characterised as having functions which confer drought resistance. In an effort to validate the results of the AT analysis, it was decided that mutants of a putative gene, identified by AT, should be characterised, in a similar fashion to the aquaporins identified in the SNP analysis. Due to time constraints only one putative gene could be selected for further analysis (Chapter 5).

.

5. Validation of candidate gene identified by Associative Transcriptomics

5.1 Introduction

IWGSC gene annotation information lists TraesCS4B01G319100.1 as a 'Heat shock protein 40 (HSP40)'. HSP40 isoforms are chaperone proteins that control protein folding for many native polypeptides in plants (Barghetti et al., 2017). HSP40s organise into a conserved complex in which a HSP40 dimer binds an unfolded target protein, activating HSP70 hydrolysis, resulting in the organisation of a HSP70-ADP-target complex (Barghetti et al., 2017; Misselwitz et al., 1998; Hernandez et al., 2002). Subsequently, the target protein is transferred to HSP90 and numerous other chaperone proteins for final conformational maturation (Barghetti et al., 2017; Johnson et al., 1998; Pratt et al., 2008). Proteins crucial to plant development, such as auxin and jasmonate receptors, have been shown to have their conformation regulated by this assembly line, termed the HSP90 chaperone pathway (Zhang et al., 2015; Wang et al., 2016).

Protein farnesylation, a lipid-based post-translational protein modification in which an isoprenyl group is attached to a C-terminal cysteine residue, is an important regulator of protein-protein and protein-membrane interactions (Wang and Casey, 2016). Protein farnesyl transferase enzymes are heterodimeric proteins, consisting of an α and β subunit, that exhibit deeply conserved structures (Barghetti et al., 2017). In *Arabidopsis thaliana*, the α subunit is encoded by the PLURIPETALA gene (*PLP*; AT3G59380; Running et al., 2004) whilst the β subunit is encoded by ENHANCED RESPONSE TO ABA1 (*ERA1*; AT5G40280; Cutler et al., 1996).

Loss of function mutants era1 and plp, exhibit a substantially increased sensitivity to abscisic acid, and are hence considered to be involved in ABA signalling (Cutler et al., 1996; Running et al., 2004). It remains unclear as to how farnesylation plays a role in ABA signalling at the molecular level. Despite this limitation in understanding, Wang et al. (2005) demonstrated that 'transgenic *Brassica napus* carrying an *ERA1* antisense construct driven by a drought-inducible rd29A promoter' exhibited an 'enhanced ABA sensitivity, as well as [a] significant reduction in stomatal conductance and water transpiration under drought stress conditions', relative to non-transgenic plants.

Farnesyl transferase mutants also exhibit late flowering, decreased heat tolerance and defects in blue light-induced regulation of stomatal aperture (Wu et al., 2017; Jalakas et al., 2017). The molecular events underlying the phenotypes of these mutations also remain poorly understood, chiefly due to the identity of many farnesylation target proteins remaining unknown. Hence, protein farnesylation is central to many key physiological processes.

Barghetti et al. (2017) demonstrated that dysregulated farnesylation of a single HSP40 is sufficient to confer hypersensitivity to ABA and produce a drought resistant phenotype. It was suggested from this study that the post-translational modification of this HSP40 altered its function, and consequently altered the conformation of the proteins which the HSP40 is involved in folding – one or more of which are likely involved in ABA signalling. Whilst the HSP40 identified by this investigation (TraesCS4B01G319100.1) may not possess the CaaX motif (a, aliphatic residues; X, any residue) required to undergo farnesylation; this study demonstrated evidence for a potential role for other HSP40s and the HSP90 chaperone pathway in regulating ABA signalling and in producing drought resistant phenotypes.

The rationale for this section of the study was based upon a potential role for TraesCS4B01G319100.1 in regulating hyperosmotic stress responses, after having considered the role of HSP40s demonstrated by Barghetti et al. (2017). Since a linear regression analysis found there to be a significant relationship betweenTraesCS4B01G319100.1 RPKM and NCFRB, the effect of TraesCS4B01G319100.1 on NCFRB was investigated further. It was therefore hypothesised that a loss-of-function mutation in TraesCS4B01G319100.1 would result in a lowered normalised change in fresh root biomass (NCFRB) trait value compared to a control plant possessing the wild-type form of the gene. The nature of these changes to root morphology was also of interest, hence, WinRHIZO was utilised to characterise root morphology.

5.2 Methods

In order to retrieve the materials required to test this hypothesis, the TraesCS4B01G319100.1 nucleotide sequence was BLAST searched against the John Innes Centre wheat TILLING line database. The putative gene with the closest sequence identity match, that belonged to the appropriate genome (4B), was considered the corresponding homologue (IWGSC_CSS_4BL_scaff_7041092). Again, only mutations predicted to incur stop codon mutations were considered sufficiently disruptive to biological function to be appropriate for this study. A single homozygous mutation was listed at position 2003. Using NCBI's ORF finder tool, and the IWGSC_CSS_4BL_scaff_7041092 nucleotide sequence, 23 open reading frames were predicted. The largest open reading frame predicted was placed from position 1634 – 2593, constituting 960 nucleotides and 319 amino acids. The next largest ORF contained

only 148 nucleotides. When this predicted 319aa ORF is pBLAST searched against the NCBI protein database, the top result shows a 100% coverage, 97.49% identity match with an *Aegilops tauschii* DnaJ homolog subfamily B member 12 protein (EMS54548.1). This is a HSP40 protein. HSP40 proteins are typically between 300 – 400 aa is length. Provided the stop codon mutation listed is correct, then the resulting transcript would no longer utilise the information of 196 codons, thus producing a protein with a structure so drastically altered that is would be highly unlikely to retain its initial function(s) The mutation belonged to Kronos4610, a *Triticum durum* line. Kronos4610, the Kronos parent line and the Cadenza (*Triticum aestivum*) parent line, were ordered from the John Innes Centre.

KRONOS4610, the KRONOS parental line and the hexaploid Cadenza parental line were germinated and cultured hydroponically in the same manner as the panel screen experiment. Four biological repeats were cultured in each condition (no treatment control and 7.5% PEG8000 (w/v)), for each line. Samples were placed in hydroponics according to a repeating Latin square design. Fresh and dry biomass measurements were taken for root and shoot structures. Minimal excised leaf samples were taken from each plant, CTAB DNA extractions performed and Kronos samples genotyped according to section 2.7. The predicted C:T mutation has listed codon information, purports that a tgG:tgA sequence change is responsible for a stop codon. Primers were designed using an EnsemblPlants *Triticum dicoccoides* gene sequence listed as a homolog to IWGSC_CSS_4BL_scaff_7041092. Due to sequence variation between the IWGSC_CSS_4BL_scaff_7041092 and the *Triticum dicoccoides* sequence, primers were designed to allow for at least 100 nucleotides either side of the 2003rd position. The Kronos parental line was selected as a negative control, with the expectation that it would contain a wild-type copy of TraesCS4B01G319100.1 that did not contain a C:T mutation at position 2003. The Cadenza parental line was selected to allow comparison with

hexaploid bread wheat morphology. Roots of all three lines, from both conditions were subsequently imaged using an Epson V1.0.3.7, and measured for surface area (cm^2), total volume (cm^3) and average diameter (mm) using WinRHIZO's default morphology analysis (Arsenault et al., 1995).

5.3 Results

Normalised change in fresh root biomass (NCFRB) was seen to be significantly lower in Kronos4610 compared to the Cadenza and Kronos parent lines (Figure 5.1, ANOVA: F = 8.656, d.f. = 5, p = 0.000254). NCFRB was not seen to differ between the Kronos and Cadenza parent lines.

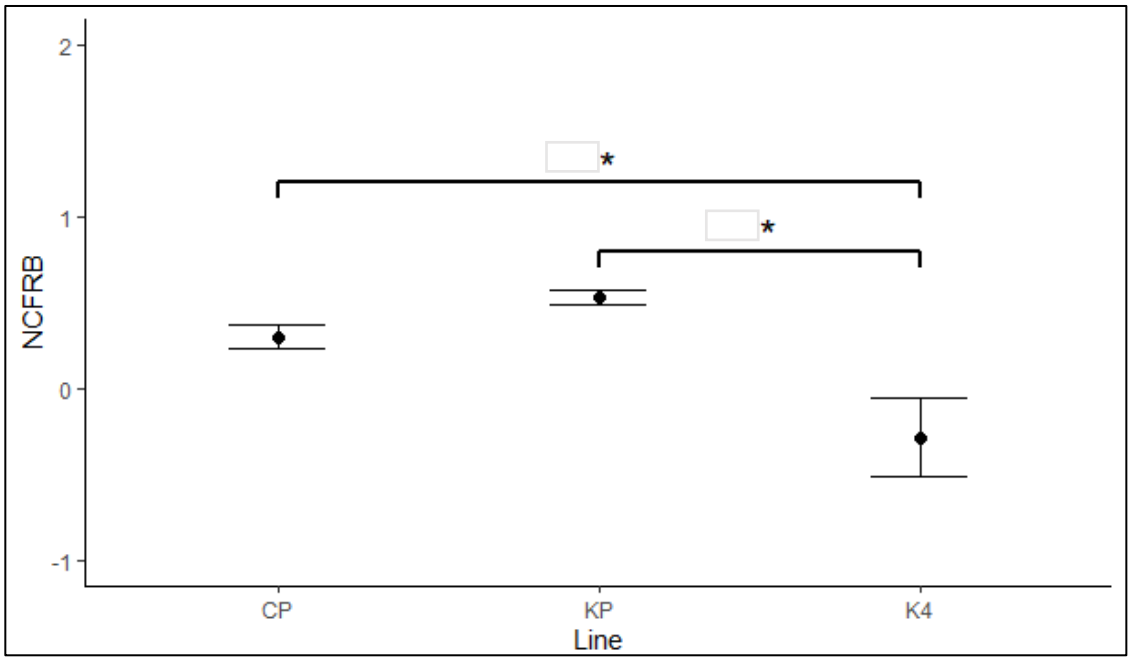


Figure 5.1: Plot comparing normalised change in fresh root biomass (NCFRB) of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water and supplemented with 7.5% (w/v) polyethylene glycol 8000.

Cadenza parent (CP), Kronos parent (KP) and Kronos4610 (K4). All lines, N=4, $p < 0.05$ (*).

Genotyping of Kronos4610 and Kronos parental line samples produced results that cast doubt as to whether Kronos4610 contains the stop codon mutation listed by the John Innes Centre (see Additional Information). PCR products, amplified using the L1 and R1 primers (Supplementary Table 4), were sent for Sanger sequencing (GATC Biotech LIGHTrun). Two sets of samples were sent, one including the L1 primer, another containing the R1 primer. Both sets of samples returned high quality sequence results (>30) for at least four samples of each line investigated. Both sets of results failed to demonstrate that Kronos4610 contained the C:T mutation, listed by the John Innes Centre as being responsible for a stop codon mutation. Evidence of the purported tgG:tgA sequence change could not be observed.

Analysis of WinRHIZO root imaging measurements showed Kronos4610 treatment plants better maintained fresh root biomass by not displaying a significant decrease in root volume or surface area, relative to control plants of the same line (Figure 5.2 and Supplementary Figures 5, ANOVA: $F = 8.656$, d.f. = 5, $p = 0.000254$). Treatment plants for Kronos and Cadenza parental lines displayed greater differences in median values relative to control plants of the same line. For all traits measured, Kronos lines displayed differences that were significant at $p<0.05$ level. However, post-hoc analysis found the Cadenza parental line to only display a difference at the typically disregarded $p<0.1$ level or at no regarded significance level (Figure 5.2 and Supplementary Figure 5). Average root diameter of treatment plants was found to significantly increase relative to control plants of the same line (ANOVA: $F = 24.28$, d.f. = 5, $p = 2.05e^{-07}$).

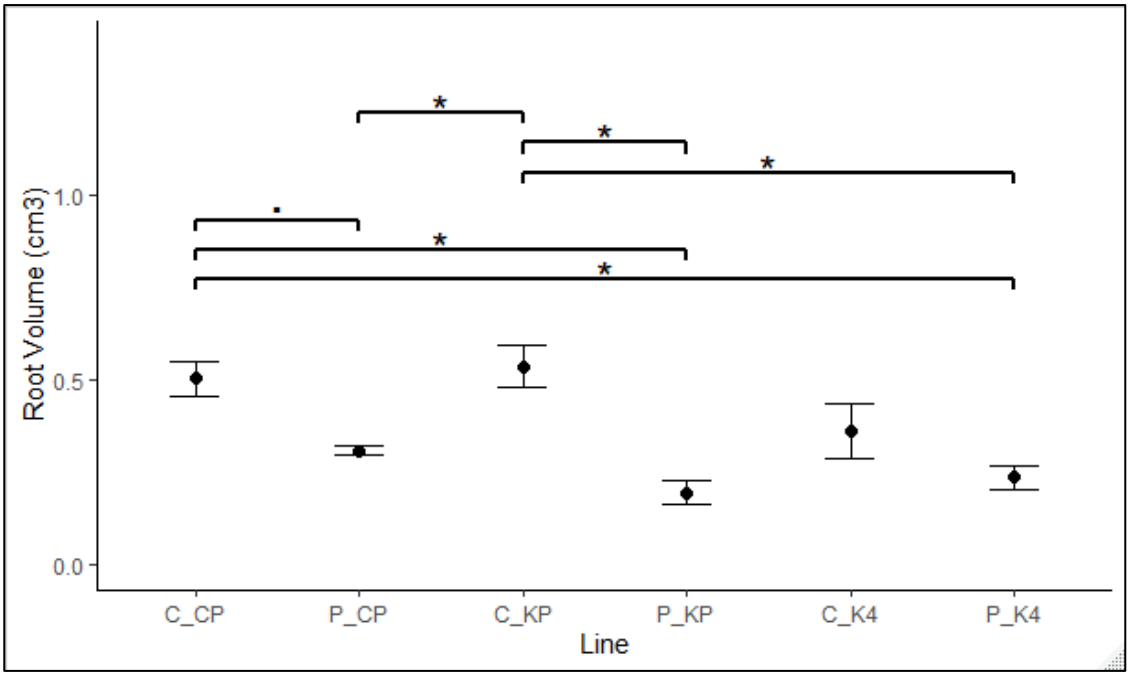


Figure 5.2: Plot comparing root volume (cm3) of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (C) and supplemented with 7.5% (w/v) polyethylene glycol 8000 (P). Cadenza parent (_CP), Kronos4610 (_K4) and Kronos parent (_KP). All lines, N=4 (ANOVA, $p < 0.05$ (*), $p < 0.1$ (·)). Root volume data was generated using WinRHIZO.

5.4 Discussion

Kronos4610 exhibited a significantly lower NCFRB value, compared with the Kronos and Cadenza parental lines (Figure 5.1), as predicted by the relationship between RPKM for and NCFRB values for the 135 landraces employed in the Associative Transcriptomics analysis (Figure 4.8).

Sequencing results failed to demonstrate the predicted C:T SNP mutation listed as causing a stop codon. Potential explanations for the results seen include: an amplification of a region in fact not containing the SNP but which belongs to a different section of the

target putative gene sequence in question, an amplification of a highly similar homolog found to exist on chromosome 5A, or the amplification of a section of DNA which does not belong to the *Triticum emmer* TraesCS4B01G319100.1 homolog. Primers, listed as semi-specific, that were initially used to genotype the Kronos4610 mutant are available in supplement to other information pertaining the stop codon mutation here in question. It is here suggested that a repeat of this analysis, coupled with a genotyping analysis, utilising these primers, is the best course of action to providing to stronger evidence that a TraesCS4B01G319100.1 homolog is indeed being truncated as expected. A qPCR analysis, alongside this work, would also serve as useful in demonstrating a lack of TraesCS4B01G319100.1 homolog expression in Kronos4610.

Root imaging results demonstrated that the observed increase in NCFRB in Kronos4610, relative to the Kronos parental line, was attributed to a maintenance of surface area, volume and average root diameter – rather than a drastic change in root architecture. Confirmation of this mutation and a repeated finding of this increase are here considered sufficient evidence to encourage further investigation of TraesCS4B01G319100.1 and its effects on fresh root biomass accumulation in response to hyperosmotic stress conditions. Especially interesting, would be a bioinformatics analysis, searching for the presence of known post-translational modification motifs similar to, or differ from those identified by Barghetti et al (2017), as a first step in then trying to establish the cause of this mutation’s effects.

6. Discussion

This study aimed to identify genetic markers significantly associated with hyperosmotic stress-responsive biomass accumulation traits, in the hexaploid crop *Triticum aestivum*. This phenotypic data was gathered for 135 hydroponically cultured YoGI panel landraces, exposed to both a no treatment control condition, and a polyethylene glycol, hyperosmotically stressed condition. Utilising Associative Transcriptomics (AT) to look for statistically significant associations between the transcriptomic data's genetic markers and the measured phenotype variation, twenty-two genetic markers were identified.

Twenty-one of those genetic markers were identified through the GEM AT analysis. Contemporary IWGSC gene annotation information and NCBI pBLAST search results described the functions of these putative genes as including cell signalling (CDPK9), flowering promotion (FT), and transcriptional regulation (MYB transcription factor) – all of which have previously demonstrated roles in hyperosmotic stress and drought responses (Li et al., 2008; Riboni et al., 2013; Lee et al., 2014). Hence, this serves as evidence supporting the notion, that genetic markers identified by this study, do in fact contribute to biomass accumulation in response to hyperosmotic stress.

Among those twenty-one genetic markers was TraesCS4B01G319100.1, a putative gene here described as a member of the heat shock protein 40 (HSP40) family. A linear regression analysis found TraesCS4B01G319100.1 RPKM expression data, for the 135 landraces used, to display a statistically significant relationship with the measured normalised change in fresh root biomass (NCFRB) trait data. The negative correlation revealed by this analysis suggested that *T. aestivum* landraces with a lower TraesCS4B01G319100.1 RPKM expression value, are more likely to also exhibit a

reduced NCFRB value. In an experiment replicating the conditions of those previously used to phenotype the landraces, *Triticum durum* TILLING line mutant Kronos4610 NCFRB data was compared to that of the Kronos parental line and the *T. aestivum* Cadenza parental line. Kronos4610 was purported to have a stop codon mutation at a nucleotide position that was here expected to cause a substantial disruption to TraesCS4B01G319100.1 function, thus representing the effects of ceasing TraesCS4B01G319100.1 expression. It was found that Kronos4610 NCFRB was significantly lower than that of either the Kronos parental line or the Cadenza parental line. Despite the failure of this study's efforts to validate the stop codon mutation listed by the John Innes Centre, these results are here presented as strong evidence in support of the validity of the asserted roles of the genetic markers, identified by this study's AT analysis, in hyperosmotic stress responses.

It is therefore suggested that further work be completed in the investigation of TraesCS4B01G319100.1's role in wheat hyperosmotic stress responses. Further molecular validation experiments are suggested, including an alternative approach to a sequencing-based validation of the alleged stop codon mutation, as well as a supplementary qPCR analysis. Since, the original phenotyping was conducted using *Triticum aestivum* lines, it would also be of major importance to demonstrate that a loss-of-function mutation in TraesCS4B01G319100.1 in bread wheat is capable of similarly producing a decrease in NCFRB. Such evidence would strengthen the case, put forward here for a role for TraesCS4B01G319100.1 in regulating NCFRB, and support a move for a continuation of work surrounding TraesCS4B01G319100.1.

The AT SNP analysis performed in this study identified a genomic region in which two SNP markers significantly associated with measured NCFRB variation. Following an analysis of the delineated peak region for candidate genes likely to contribute to this phenotypic variation, using IWGSC annotation information, two putative aquaporins were selected (TraesCS6A01G405600.1_1 and TraesCS6A01G405700.1_1). In the light of this evidence supporting the validity of those genetic markers identified by this study's AT GEM analysis, it is again reiterated that an adaptation of the phenotyping experiment, similar to that used in the HSP40 analysis, but utilising the homozygous mutant SALK lines for aquaporin PIP1;5 T-DNA insertion, represents an attractive avenue for further potential work.

This study has successfully demonstrated the potential application of Associative Transcriptomics in analysing biomass trait data, generated from the treatment of hydroponically cultured *Triticum aestivum* landraces, to identify genetic markers associated with hyperosmotic stress responses. Although further work is yet required to confidently validate the role of TraesCS4B01G319100.1 in hyperosmotic stress responses, the morphological analysis completed in this study, using wheat TILLING mutant Kronos4610, serves as an example of how AT analysis can lead onto more classical genetic approaches. It is here stated that Associative Transcriptomics represents a previously unutilised method for the study of hyperosmotic stress responses in plants and perhaps even other abiotic stresses, such as heat stress.

Creating a unifying model to explain the different features that characterise plant responses to water deficient conditions has so far proved extremely difficult. This is owed to the vast quantity of molecular events involved, occurring in parallel, and doing so in a manner dependent on the tissue-type, the intensity of the stress, the length of application

and other environmental conditions. Many methodological challenges also lie in simulating water deficient conditions, determining plant water content and in relating particular molecular events to physiological condition, complicating the integration and comparison of results from different studies. The greatest limitation of the evidence presented in this study lies in how the hydroponic method utilised for the phenotyping experiment differs so greatly from the natural, soil-based conditions of canonical *T. aestivum* growth. From the 135 landraces utilised in this study, a comparison of total dry biomass identified the ten most 'drought susceptible' and 'drought tolerant' landraces. A comparison of the total dry biomasses of these landraces, following growth using a soil-based, water restricted regimen, would provide invaluable evidence as to the validity of this studies phenotyping results and the suitability of the identified genetic markers for consideration for use in a breeding programme.

A substantial amount of work is yet required to understand the underlying genetic basis of plant abiotic stress responses, such as hyperosmotic stress responses. Given the increasing concerns around climate change and global food security, studies such as this one may turn out to represent a powerful approach in identifying some of the as yet unrecognised genetic factors contributing to these responses, and play an important role in the ongoing work required to bolster global agriculture yields, and aid efforts to better feed the world.

Word count (excluding references, tables and figures):

Abstract: 283

Thesis: 21524

References:

Adams, KL., Cronn, R., Percifield, R., Wendel, JF. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proceedings of the National Academy of Sciences. 100(8): 4649-4654.

Ahmad, P., Rasool, S., Gul, A., Sheikh, S.A., Akram, N.A., Ashraf, M., Kazi, A.M., Gucel, S. (2016). Jasmonates: Multifunctional Roles in Stress Tolerance. Frontiers in Plant Science. 7: 813.

Alexandersson, E., Fraysse, L., Sjövall-Larsen, S., Gustavsson, S., Fellert, M., Karlsson, M., … Kjellbom, P. (2005). Whole gene family expression and drought stress regulation of aquaporins. Plant Molecular Biology. 59(3): 469-84.

Anjum, S., Xie, X., Wang, L., … M. S.-A. J. of, & 2011, U. (2011). Morphological, physiological and biochemical responses of plants to drought stress. African Journal of Agricultural Research. 6(9): 2026-2032.

Annicchiarico, P., & Pecetti, L. (1998). Yield vs. morphophysiological trait-based criteria for selection of durum wheat in a semi-arid Mediterranean region (northern Syria). Field Crops Research. 59(3): 163-167

Antonioa, C., Pinheiro, C., Manuela, M., Candido, C., Ricardo, P., Fernanda, M., Thomas-Oates. J., (2008). Analysis of carbohydrates in Lupinus albus stems on imposition of water deficit, using porous graphitic carbon liquid chromatography-electrospray ionization mass spectrometry. American Journal of Chromatography. 1187(1-2): 111-118.

Apel, K., & Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annual Review of Plant Biology. 55: 373-99

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., … Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 361(6403): 7191

Araus, J. L., Slafer, G. A., Reynolds, M. P., & Royo, C. (2002). Plant breeding and drought in C3 cereals: What should we breed for? Annals of Botany. 89: 925-40

Arsenault, J.-L., S. Pouleur, C. Messier, and R. Guay. (1995). WinRHIZO™, a root-measuring system with a unique overlap correction method. HortScience 30: 906.

Atkin, O. K., & Macherel, D. (2009). The crucial role of plant mitochondria in orchestrating drought tolerance. Annals of Botany. 103(4): 581–597.

Barghetti, A., Sjögren, L., Floris, M., Paredes, E. B., Wenkel, S., & Brodersen, P. (2017). Heat-shock protein 40 is the key farnesylation target in meristem size control, abscisic acid signaling, and drought resistance. Genes and Development. 31(22): 2282-2295.

Beerling, D.J., McElwain, J.C., Osborne, C.P., (1998). Stomatal responses of the “living fossil” Ginkgo biloba L. to changes in atmospheric CO2 concentrations. Journal of Experimental Botany. 49(326): 1603-1607.

Berryman, C.A., Eamus, D., Duff, G.A., (1994). Stomatal responses to a range of variables in two tropical tree species grown with CO2 enrichment. Journal of Experimental Botany. 45(5): 539-546.

Blasing, O. E. (2005). Sugars and Circadian Regulation Make Major Contributions to the Global Regulation of Diurnal Gene Expression in Arabidopsis. The Plant Cell. 17: 3257–3281.

Bodner, G., Nakhforoosh, A., & Kaul, H. P. (2015). Management of crop water under drought: a review. Agronomy for Sustainable Development. 35(2): 401–442.

Borghetti, M., Grace, J. and Raschi, A. (eds) (1993) Water Transport in Plants Under Climate Stress. Cambridge University Press, Cambridge.

Brandt, B., Munemasa, S., Wang, C., Nguyen, D., Yong, T., Yang, P.G., Poretsky, E., Belknap, T.F., Waadt, R., Alemań, F., Schroeder, J.I., (2015). Calcium specificity signaling mechanisms in abscisic acid signal transduction in arabidopsis guard cells. Elife. 4: e03599.

Chandler, J.M., (2001). Current molecular understanding of the genetically programmed process of leaf senescence. Physiologia Plantarum. 113(1): 1-8.

Chaves, M. M. (1991). Effects of water deficits on carbon assimilation. Journal of Experimental Botany. 42(1): 1-16.

Choi, W.-G., Toyota, M., Kim, S.-H., Hilleary, R., Gilroy, S., (2014). Salt stress-induced Ca2+ waves are associated with rapid, long-distance root-to-shoot signaling in plants. Proceedings of the National. Academy of Sciences. 111(17): 6497-6502.

Christmann, A., Weiler, E.W., Steudle, E., Grill, E., (2007). A hydraulic signal in root-to-shoot signalling of water shortage. Plant J. 52(1): 167-174.

Cutler, S., Ghassemian, M., Bonetta, D., Cooney, S., & McCourt, P. (1996). A protein farnesyl transferase involved in abscisic acid signal transduction in Arabidopsis. Science. 273(5279): 1239-1241.

Davies, W. J., Kudoyarova, G. and Hartung, W. (2005). Long-distance ABA signaling and its relation to other signaling pathways in the detection of soil drying and the mediation of the plant's response to drought, Journal of Plant Growth Regulation. 24: 285–295.

De Vita, P., Nicosia, O. L. D., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N., & Cattivelli, L. (2007). Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. European Journal of Agronomy. 26(1): 39-53.

de Zelicourt, A., Colcombet, J., Hirt, H., (2016). The Role of MAPK Modules and ABA during Abiotic Stress Signaling. Trends in Plant Science. 21(8): 677-685.

Demmig-Adams B, Adams III WW, Mattoo A, eds. (2006). Photoprotection, photoinhibition, gene regulation and environment. Advances in photosynthesis and respiration. Springer.

Demmig-Adams, B., & Adams, W. W. (1996). The role of xanthophyll cycle carotenoids in the protection of photosynthesis. Trends in Plant Science. 94(25): 14162–14167.

Dietz, K. J., Jacob, S., Oelze, M. L., Laxa, M., Tognetti, V., De Miranda, S. M. N., … Finkemeier, I. (2006). The function of peroxiredoxins in plant organelle redox metabolism. In Journal of Experimental Botany. 57(8): 1697-1709.

Durbak, A. R., Phillips, K. A., Pike, S., O'Neill, M. A., Mares, J., Gallavotti, A., … McSteen, P. (2014). Transport of Boron by the tassel-less1 Aquaporin Is Critical for Vegetative and Reproductive Development in Maize. The Plant Cell. 26: 2978–2995.

Else, M.A., Jackson, M.B., (1998). Transport of 1-aminocyclopropane-1-carboxylic acid (ACC) in the transpiration stream of tomato (Lycopersicon esculentum) in relation to foliar ethylene production and petiole epinasty. Australian Journal of Plant Physiology. 25(4): 453-458.

Estavillo, G.M., Crisp, P.A., Pornsiriwong, W., Wirtz, M., Collinge, D., Carrie, C., Giraud, E., Whelan, J., David, P., Javot, H., Brearley, C., Hell, R., Marin, E., Pogson, B.J., (2011). Evidence for a SAL1-PAP Chloroplast Retrograde Pathway That Functions in Drought and High Light Signaling in Arabidopsis. Plant Cell. 23(11): 3992-4012.

Fang, Y., Du, Y., Wang, J., Wu, A., Qiao, S., Xu, B., Zhang, S., Siddique, K.H.M., Chen, Y., (2017). Moderate Drought Stress Affected Root Growth and Grain Yield in Old, Modern and Newly Released Cultivars of Winter Wheat. Frontiers in Plant Science. 8: 672.

FAO, IFAD, UNICEF, WFP and WHO. (2018). The State of Food Security and Nutrition in the World (2018). Building climate resilience for food security and nutrition. Rome, FAO. Licence: CC BY-NC-SA 3.0 IGO.

FAO. (2016). Climate change and food security: risks and responses. Rome.

Fedoroff, N. V., Battisti, D.S., Beachy, R.N., Cooper, P.J.M., Fischhoff, D.A., Hodges, C.N., Knauf, V.C., Lobell, D., Mazur, B.J., Molden, D., Reynolds, M.P., Ronald, P.C., Rosegrant, M.W., Sanchez, P.A., Vonshak, A., Zhu, J.K., (2010). Radically rethinking agriculture for the 21st century. Science. 327(5967): 833-834.

Finn, R. N., Chauvigné, F., Hlidberg, J. B., Cutler, C. P., & Cerdà, J. (2014). The lineage-specific evolution of aquaporin gene clusters facilitated tetrapod terrestrial adaptation. PLoS ONE. 9(11): e113686.

Fitter, A.H., Hay, R.K.M., (2002). Environmental physiology of plants, Trends in Ecology and Evolution. Chapter 4.

Food and Agriculture Organisation of the United Nations FAOSTAT statistics database, (2017); www.fao.org/faostat/en/#data/FBS.

Furihata, T., Maruyama, K., Fujita, Y., Umezawa, T., Yoshida, R., Shinozaki, K., Yamaguchi-Shinozaki, K., (2006). Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. Proceedings of the National. Academy of Sciences 103(6): 1988-1993.

Gallagher, J.N., Biscoe, P. V., Hunter, B., (1976). Effects of drought on grain growth. Nature. 264: 541–542.

Galmés, J., Medrano, H., & Flexas, J. (2007). Photosynthetic limitations in response to water stress and recovery in Mediterranean plants with different growth forms. New Phytologist. 175(1): 81-93.

Galmés, J., Ribas-Carbó, M., Medrano, H., Flexas, J., (2011). Rubisco activity in Mediterranean species is regulated by the chloroplastic CO2 concentration under water stress. Journal of Experimental Botany. 62(2): 653–665.

Geiger, D., Scherzer, S., Mumm, P., Stange, A., Marten, I., Bauer, H., Ache, P., Matschi, S., Liese, A., Al-Rasheid, K.A.S., Romeis, T., Hedrich, R., (2009). Activity of guard cell anion channel SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. Proceedings of the National Academies of Sciences. 106(50): 21425-21430.

Gibson, S. I. (2000) Plant sugar-response pathways. Part of a complex regulatory web. Plant Physiology. 124: 2000.

Gilbert, M. E., Zwieniecki, M. A. and Holbrook, N. M. (2011). Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. Journal of Experimental Botany. 62(8): 2875–2887.

Gimeno, T. E., Sommerville, K. E., Valladares, F., & Atkin, O. K. (2010). Homeostasis of respiration under drought and its important consequences for foliar carbon balance in a drier climate: Insights from two contrasting Acacia species. Functional Plant Biology. 37(4): 323-333.

Grant, O.M., Incoll, L.D., McNeilly, T., (2005). Variation in growth responses to availability of water in Cistus albidus populations from different habitats. Functional Plant Biology. 32(9): 817-829.

Gratani, L., Varone, L., & Bonito, A. (2007). Environmental induced variations in leaf dark respiration and net photosynthesis of Quercus ilex L. Photosynthetica. 45: 633.

Gregersen, P.L., Holm, P.B., (2007). Transcriptome analysis of senescence in the flag leaf of wheat (Triticum aestivum L.). Journal of Plant Biotechnology. 5(1): 192-206.

Gregorová, Z., Kováčik, J., Klejdus, B., Maglovski, M., Kuna, R., Hauptvogel, P., Matušíková, I., (2015). Drought-Induced Responses of Physiology, Metabolites, and PR Proteins in Triticum aestivum. Journal of Agriculture and Food Chemistry. 63(37): 8125-8133.

Groszmann, M., Osborn, H. L., & Evans, J. R. (2017). Carbon dioxide and water transport through plant aquaporins. Plant Cell and Environment. 40(6): 938-961.

Guo, R., Shi, L., Jiao, Y., Li, M., Zhong, X., Gu, F., Liu, Q., Xia, X., Li, H., (2018). Metabolic responses to drought stress in the tissues of drought-tolerant and drought-sensitive wheat genotype seedlings. AoB Plants. 10(2).

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.

Hamanishi, E.T., Thomas, B.R., Campbell, M.M., (2012). Drought induces alterations in the stomatal development program in Populus. Journal of Experimental Botany. 63(13): 4959-4971.

Hardie, D. G., Schaffer, B. E., & Brunet, A. (2016). AMPK: An Energy-Sensing Pathway with Multiple Inputs and Outputs. Trends in Cell Biology. 26(3): 190-201.

Harlan, J.R., (1975). Our Vanishing Genetic Resources. Science. 188(4188): 617-621.

Harper, A.L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., Hattori, C., Werner, P., Bancroft, I., (2012). Associative transcriptomics of traits in the polyploid crop species Brassica napus. Nature Biotechnology. 30: 798-802.

Heinen, R. B., Bienert, G. P., Cohen, D., Chevalier, A. S., Uehlein, N., Hachez, C., Chaumont, F. (2014). Expression and characterization of plasma membrane aquaporins in stomatal complexes of Zea mays. Plant Molecular Biology. 86(3): 335-350.

Herms, D.A., Mattson, W.J., (2004). The Dilemma of Plants: To Grow or Defend. The Quarterly Review of Biology. 67(3): 283-335.

Hernandez MP, Sulliva WP, Toft DO (2002) The assembly and intermolecular properties of the Hsp70-Hop-Hsp90 molecular chaperone complex. J Biol Chem. 277(41): 38294-38304.

Horie, T., Kaneko, T., Sugimoto, G., Sasano, S., Panda, S. K., Shibasaka, M., & Katsuhara, M. (2011). Mechanisms of water transport mediated by PIP aquaporins and their regulation via phosphorylation events under salinity stress in barley roots. Plant and Cell Physiology. 52(4): 663-675.

Hou, Q., Ufer, G., Bartels, D., (2016). Lipid signalling in plant responses to abiotic stress. Plant, Cell and Environment. 39(5): 1029-1048.

Hua, D., Wang, C., He, J., Liao, H., Duan, Y., Zhu, Z., Guo, Y., Chen, Z., Gong, Z., (2012). A Plasma Membrane Receptor Kinase, GHR1, Mediates Abscisic Acid- and Hydrogen Peroxide-Regulated Stomatal Movement in Arabidopsis. Plant Cell. 24(6): 2546-2561.

Huang, X., & Han, B. (2014). Natural Variations and Genome-Wide Association Studies in Crop Plants. Annual Review of Plant Biology. 65: 531-551.

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Han, B. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nature Genetics. 44: 32–39.

Hungate, R.E., (1934). THE COHESION THEORY OF TRANSPIRATION. PLANT Physiol. 7(2):285–295.

International Plant Genetic Resources Institute (IPGRI, 1985) Descriptors for Wheat (Revised), IPGRI, Rome, Italy.

IPCC, (2014): Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate

Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland.

Jalakas, P., Huang, Y. C., Yeh, Y. H., Zimmerli, L., Merilo, E., Kollist, H., & Brosché, M. (2017). The role of ENHANCED RESPONSES TO ABA1 (ERA1) in arabidopsis stomatal responses is beyond ABA signaling. Plant Physiology. 174: 665–671.

Jaleel, C. A., Gopi, R., Sankar, B., Gomathinayagam, M., & Panneerselvam, R. (2008). Differential responses in water use efficiency in two varieties of Catharanthus roseus under drought stress. Comptes Rendus - Biologies. 331(1): 42-47.

Ji, H., Li, X., (2014). ABA mediates PEG-mediated premature differentiation of root apical meristem in plants. Plant Signalling and Behaviour. 9(11): e977720.

Ji, H., Liu, L., Li, K., Xie, Q., Wang, Z., Zhao, X., Li, X., (2014). PEG-mediated osmotic stress induces premature differentiation of the root apical meristem and outgrowth of lateral roots in wheat. Journal of Experimental Botany. 65(17): 4863-4872.

Ji, X., Shiran, B., Wan, J., Lewis, D.C., Jenkins, C.L.D., Condon, A.G., Richards, R.A., Dolferus, R., (2010). Importance of pre-anthesis anther sink strength for maintenance of grain number during reproductive stage water stress in wheat. Plant, Cell and Environment. 33(6): 926-942.

Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., Han, B. (2013). A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (Setaria italica). Nature Genetics. 45: 957-961.

Johnson, B. D., Schumacher, R. J., Ross, E. D., & Toft, D. O. (1998). Hop modulates hsp70/hsp90 interactions in protein folding. Journal of Biological Chemistry. 273: 3679-3686.

Jonak, C., Kiegerl, S., Ligterink, W., Barker, P. J., Huskisson, N. S., & Hirt, H. (2002). Stress signaling in plants: a mitogen-activated protein kinase pathway is activated by cold and drought. Proceedings of the National Academy of Sciences. 93(20): 11274-11279.

Khan, M. S. (2011). The role of DREB transcription factors in abiotic stress tolerance of plants. Biotechnology and Biotechnological Equipment. 25(3): 2433-2442.

Kitsios, G., & Doonan, J. H. (2011). Cyclin dependent protein kinases and stress responses in plants. Plant Signaling and Behaviour. 6(2): 204–209.

Koch, K. E. (1996). CARBOHYDRATE-MODULATED GENE EXPRESSION IN PLANTS, Annual Review of Plant Physiology and Plant Molecular Biology. 47: 509-540.

Kooyers, N. J. (2015). The evolution of drought escape and avoidance in natural herbaceous populations. Plant Science. 234: 155-162.

Kreps, J. A. (2002). Transcriptome Changes for Arabidopsis in Response to Salt, Osmotic, and Cold Stress. Plant Physiology. 130(4): 2129-2141.

Kump, K. L., Bradbury, P. J., Wisser, R. J., Buckler, E. S., Belcher, A. R., Oropeza-Rosas, M. A., Holland, J. B. (2011). Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nature Genetics. 43: 163–168.

Kurzbaum, E., Kirzhner, F., Armon, R., (2010). A simple method for dehydrogenase activity visualization of intact plant roots grown in soilless culture using tetrazolium violet. Plant Root. 4: 12-16.

Lagerwerff, J. V., Ogata, G.E.N., Eagle, H.E., (1961). Control of osmotic pressure of culture solutions with polyethylene glycol. Science. 133(3463): 1486-1487.

Lake, J.A., Quick, W.P., Beerling, D.J., Woodward, F.I., (2001). Signals from mature to new leaves. Nature. 411(6834): 154.

Lawlor, D. W., & Tezara, W. (2009). Causes of decreased photosynthetic rate and metabolic capacity in water-deficient leaf cells: A critical evaluation of mechanisms and integration of processes. Annals of Botany. 103(4): 561–579.

Lee, S. B., Kim, H., Kim, R. J., & Suh, M. C. (2014). Overexpression of Arabidopsis MYB96 confers drought resistance in Camelina sativa via cuticular wax accumulation. Plant Cell Reports. 33(9): 1535-1546.

Lesk, C., Rowhani, P., Ramankutty, N., (2016). Influence of extreme weather disasters on global crop production. Nature. 529: 84–87.

Lev-Yadun, S., Gopher, A., & Abbo, S. (2000). The cradle of agriculture. Science. 288(5471): 1602-1603.

Li, J., Ban, L., Wen, H., Wang, Z., Dzyubenko, N., Chapurin, V., Wang, X. (2015). An aquaporin protein is associated with drought stress tolerance. Biochemical and Biophysical Research Communications. 459(2): 208-213.

Li, J., Ban, L., Wen, H., Wang, Z., Dzyubenko, N., Chapurin, V., Wang, X. (2015). An aquaporin protein is associated with drought stress tolerance. Biochemical and Biophysical Research Communications. 459(2): 208-213.

Li, P., Li, Y. J., Zhang, F. J., Zhang, G. Z., Jiang, X. Y., Yu, H. M., & Hou, B. K. (2017). The Arabidopsis UDP-glycosyltransferases UGT79B2 and UGT79B3, contribute to cold, salt and drought stress tolerance via modulating anthocyanin accumulation. Plant Journal. 89(1): 85-103.

Li, A., Wang, X., Leseberg, C. H., Jia, J., & Mao, L. (2008). Biotic and abiotic stress responses through calcium-dependent protein kinase (CDPK) signaling in wheat (Triticum aestivum L.). Plant Signaling and Behaviour. 3(9): 654-6.

Li, Y., Ye, W., Wang, M., Yan, X., (2009). Climate change and drought: a risk assessment of crop-yield impacts. Climate Research. 39: 31-46.

Liu, J.X., Howell, S.H., (2016). Managing the protein folding demands in the endoplasmic reticulum of plants. New Phytologist. 211(2): 418-28.

Liu, Y., Ji, X., Nie, X., Qu, M., Zheng, L., Tan, Z., … Wang, Y. (2015). Arabidopsis AtbHLH112 regulates the expression of genes involved in abiotic stress tolerance by binding to their E-box and GCG-box motifs. New Phytologist. 207(3): 692-709.

Lonbani, M., Arzani, A., (2011). Morpho-physiological traits associated with terminal drought stress tolerance in triticale and wheat. Agronomic Research. 9(1–2): 315–329.

Loss, S. P., & Siddique, K. H. M. (1994). Morphological and Physiological Traits Associated with Wheat Yield Increases in Mediterranean Environments. Advances in Agronomy. 52: 229-276.

Luu, D. T., & Maurel, C. (2005). Aquaporins in a challenging environment: Molecular gears for adjusting plant water status. Plant, Cell and Environment. 28(1): 85-96.

Ma, X., Xin, Z., Wang, Z., Yang, Q., Guo, S., Guo, X., Cao, L., Lin, T., (2015). Identification and comparative analysis of differentially expressed miRNAs in leaves of two wheat (Triticum aestivum L.) genotypes during dehydration stress. BMC Plant Biol. 15: 21.

Ma, Y., Dai, X., Xu, Y., Luo, W., Zheng, X., Zeng, D., Pan, Y., Lin, X., Liu, H., Zhang, D., Xiao, J., Guo, X., Xu, S., Niu, Y., Jin, J., Zhang, H., Xu, X., Li, L., Wang, W., Qian, Q., Ge, S., Chong, K., (2015). COLD1 confers chilling tolerance in rice. Cell. 160(6): 1209-1221.

Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A., Grill, E., (2009). Regulators of PP2C phosphatase activity function as abscisic acid sensors. Science. 324(5930): 1064-1068.

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K. S., … Praud, S. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. Science. 345(6194): 1250092.

McDowell, N. G. (2011). Mechanisms Linking Drought, Hydraulics, Carbon Metabolism, and Vegetation Mortality. Plant Physiology. 155: 1051–1059.

Miao, Y., Lv, D., Wang, P., Wang, X.-C., Chen, J., Miao, C., & Song, C.-P. (2006). An Arabidopsis Glutathione Peroxidase Functions as Both a Redox Transducer and a Scavenger in Abscisic Acid and Drought Stress Responses. THE PLANT CELL. 18(10): 2749-2766.

Michaletti, A., Naghavi, M. R., Toorchi, M., Zolla, L., & Rinalducci, S. (2018). Metabolomics and proteomics reveal drought-stress responses of leaf tissues from spring-wheat. Scientific Reports. 8: 5710.

Michel, B.E., Kaufmann, M.R., (1973). The Osmotic Potential of Polyethylene Glycol 6000. Plant Physiology. 51: 914-916.

Miller, G., Schlauch, K., Tam, R., Cortes, D., Torres, M.A., Shulaev, V., Dangl, J.L., Mittler, R., (2009). The plant NADPH oxidase RBOHD mediates rapid systemic signaling in response to diverse stimuli. Scientific Signaling. 2(84): 45.

Miller, G., Suzuki, N., Ciftci-Yilmaz, S., & Mittler, R. (2010). Reactive oxygen species homeostasis and signalling during drought and salinity stresses. Plant, Cell and Environment. 33(4): 453-467.

Misselwitz, B., Staeck, O., & Rapoport, T. A. (1998). J proteins catalytically activate Hsp70 molecules to trap a wide range of peptide sequences. Molecular Cell. 2(5): 593-603.

Mittler, R., Vanderauwera, S., Gollery, M., & Van Breusegem, F. (2004). Reactive oxygen gene network of plants. Trends in Plant Science. 9(10): 490-498.

Moragues, M., Moral, L. F. G. Del, Moralejo, M., & Royo, C. (2006a). Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin I: Yield components. Field Crops Research. 95(2–3), 194-205.

Moragues, M., García Del Moral, L. F., Moralejo, M., & Royo, C. (2006b). Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin: II. Biomass production and allocation. Field Crops Research. 95(2-3): 182-193.

Morgan, J.M., (1977). Differences in osmoregulation between wheat genotypes. Nature. 270, 234-235.

Morgan, J.M., Condon, A.G., (1986). Water use, grain yield, and osmoregulation in wheat. Australian Journal of Plant Physiology. 13(4): 523-532.

Motzo, R., & Giunta, F. (2007). The effect of breeding on the phenology of Italian durum wheats: From landraces to modern cultivars. European Journal of Agronomy. 26(4): 462-470

Mwadzingeni, L., Shimelis, H., Tesfay, S., & Tsilo, T. J. (2016). Screening of Bread Wheat Genotypes for Drought Tolerance Using Phenotypic and Proline Analyses. Frontiers in Plant Science. 7: 1276.

National Academies of Sciences, Engineering, and Medicine. (2016). Attribution of Extreme Weather Events in the Context of Climate Change. Washington, DC: The National Academies Press.

Newton, A. C., Akar, T., Baresel, J. P., Bebeli, P. J., Bettencourt, E., Bladenopoulos, K. V., … Patto, M. C. V. (2009). Cereal landraces for sustainable agriculture. In Sustainable Agriculture. 30(2): 237-269.

Ng, S., De Clercq, I., Van Aken, O., Law, S.R., Ivanova, A., Willems, P., Giraud, E., Van Breusegem, F., Whelan, J., (2014). Anterograde and retrograde regulation of nuclear

genes encoding mitochondrial proteins during growth, development, and stress. Molecular Plant. 7(7): 1075-93.

Nicotra, A.B., Hermes, J.P., Jones, C.S., Schlichting, C.D., (2007). Geographic variation and plasticity to water and nutrients in Pelargonium australe. New Phytologist. 176(1): 136-49.

Nonami, H. (1998). Plant water relations and control of cell elongation at low water potentials. Journal of Plant Research. 111(3): 373-382.

Noreen, S., Fatima, K., Athar, H.U.R., Ahmad, S., Hussain, K., (2017). Enhancement of physio-biochemical parameters of wheat through exogenous application of salicylic acid under drought stress. Journal of Animal and Plant Science. 27(1): 153-163.

Normand, S., Treier, U. A., Randin, C., Vittoz, P., Guisan, A., & Svenning, J. C. (2009). Importance of abiotic stress as a range-limit determinant for European plants: Insights from species responses to climatic gradients. Global Ecology and Biogeography. 18(4): 437-449.

Norton, R.E., Hamm, J., Langenheim, J.H., (2008). Plant Resins: Chemistry, Evolution, Ecology, and Ethnobotany. Journal of American Institute for Conservation.

Otto, F.E.L., (2016). Extreme events: The art of attribution. Nature Climate Change. 6: 342-343.

Park, S.Y., Fung, P., Nishimura, N., Jensen, D.R., Fujii, H., Zhao, Y., Lumba, S., Santiago, J., Rodrigues, A., Chow, T.F.F., Alfred, S.E., Bonetta, D., Finkelstein, R., Provart, N.J., Desveaux, D., Rodriguez, P.L., McCourt, P., Zhu, J.K., Schroeder, J.I., Volkman, B.F., Cutler, S.R., (2009). Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. Science. 324(5930): 1068-71.

Passioura, J.B., (2002). Soil conditions and plant growth. Plant, Cell and Environment. 25(2): 311-318.

Patonnier, M.P., Peltier, J.P., Marigo, G., (1999). Drought-induced increase in xylem malate and mannitol concentrations and closure of Fraxinus excelsior L. stomata. Journal of Experimental Botany. 50(336): 1223-1229.

Pego, J. V., Kortstee, A. J., Huijser, C., & Smeekens, S. C. M. (2000). Photosynthesis, sugars and the regulation of gene expression. Journal of Experimental Botany. 51: 407-416.

Petersen, G., Seberg, O., Yde, M., & Berthelsen, K. (2006). Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). Molecular Phylogenetics and Evolution. 39(1): 70-82.

Piffanelli, P., Ramsay, L., Waugh, R., Benabdelmouna, A., D'Hont, A., Hollricher, K., Jørgensen, J.H., Schulze-Lefert, P., Panstruga, R., (2004). A barley cultivation-associated polymorphism conveys resistance to powdery mildew. Nature. 430(7002): 887-91.

Pinheiro, C. and Chaves, M. M. (2011). Photosynthesis and drought: Can we make metabolic connections from available data? Journal of Experimental Botany. 62(3): 869-882.

Pinheiro, C., Chaves, M. M. and Ricardo, C. P. (2001). Alterations in carbon and nitrogen metabolism induced by water deficit in the stems and leaves of Lupinus albus L. Journal of Experimental Botany. 52(358): 1063-1070.

Pires, J.C., Zhao, J., Schranz, M.E., Leon, E.J., Quijada, P.A., Lukens, L.N., Osborn, T.C., (2004). Flowering time divergence and genomic rearrangements in resynthesized Brassica polyploids (Brassicaceae). Biological Journal of the Linnean Society. 82(4): 675–688.

Poorter, H., Villar, R., (1997). The Fate of Acquired Carbon in Plants: Chemical Composition and Construction Costs. Plant Resource Allocation. 39-72.

Popescu, A. A., Harper, A. L., Trick, M., Bancroft, I., & Huber, K. T. (2014). A novel and fast approach for population structure inference using Kernel-PCA and optimization. Genetics. 198(4): 1421-31.

Pratt, W. B., Morishima, Y., & Osawa, Y. (2008). The Hsp90 chaperone machinery regulates signaling by modulating ligand binding clefts. Journal of Biological Chemistry. 283(34): 22885-22889.

Qaseem, M. F., Qureshi, R., & Shaheen, H. (2019). Effects of Pre-Anthesis Drought, Heat and Their Combination on the Growth, Yield and Physiology of diverse Wheat (Triticum aestivum L.) Genotypes Varying in Sensitivity to Heat and drought stress. Scientific Reports. 9: 6955.

Ramel, F., Sulmon, C., Gouesbet, G., & Couée, I. (2009). Natural variation reveals relationships between pre-stress carbohydrate nutritional status and subsequent responses to xenobiotic and oxidative stress in Arabidopsis thaliana. Annals of Botany. 104(7): 1323-1337.

Riboni, M., Galbiati, M., Tonelli, C., & Conti, L. (2013). GIGANTEA enables drought escape response via abscisic acid-dependent activation of the florigens and SUPPRESSOR of OVEREXPRESSION of CONSTANS11. Plant Physiology. 162: 1706-1719.

Riehl, S., Zeidi, M., & Conard, N. J. (2013). Emergence of Agriculture in the Foothills of the Zagros Mountains of Iran. Science. 341(6141): 65-67.

Rojas, O., Vrieling, A., Rembold, F., (2011). Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. Remote Sensing of Environment. 115(2): 343-352.

Rolland, F., Baena-Gonzalez, E. and Sheen, J. (2006). SUGAR SENSING AND SIGNALING IN PLANTS: Conserved and Novel Mechanisms. Annual Review of Plant Biology. 57: 675-709.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Running, M. P., Lavy, M., Sternberg, H., Galichet, A., Gruissem, W., Hake, S., … Yalovsky, S. (2004). Enlarged meristems and delayed growth in plp mutants result from lack of CaaX prenyltransferases. Proceedings of the National Academy of Sciences of the United States of America. 101(20): 7815-7820.

Sakakibara, H., (2006). CYTOKININS: Activity, Biosynthesis, and Translocation. Annual. Review of Plant Biology. 57: 431-449.

Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R., & Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. Nature Reviews Genetics. 3: 429-441.

Sangwan, V., Örvar, B.L., Beyerly, J., Hirt, H., Dhindsa Rajinder, S., (2002). Opposite changes in membrane fluidity mimic cold and heat stress activation of distinct plant MAP kinase pathways. Plant Journal. 31(5): 629-638.

Schmalenbach, I., Zhang, L., Reymond, M., (2014). The relationship between flowering time and growth responses to drought in the Arabidopsis Landsberg erecta x Antwerp-1 population. Frontiers in Plant Science. 5: 609.

Schmidhuber, J., Tubiello, F.N., (2007). Global food security under climate change. Proceedings of the National Academy of Sciences. 104(50): 19703-19708.

Sečenji, M., Lendvai, Á., Miskolczi, P., Kocsy, G., Gallé, Á., Szucs, A., Hoffmann, B., Sárvári, É., Schweizer, P., Stein, N., Dudits, D., Györgyey, J., (2010). Differences in root

functions during long-term drought adaptation: Comparison of active gene sets of two wheat genotypes. Plant Biology. 12(6): 871-882.

Shakirova, F., Allagulova, C., Maslennikova, D., Fedorova, K., Yuldashev, R., Lubyanova, A., Bezrukova, M., Avalbaev, A., (2016). Involvement of dehydrins in 24-epibrassinolide-induced protection of wheat plants against drought stress. Plant Physiology and Biochemistry. 108: 539-548.

Sharp, R. E. (2002). Interaction with ethylene: Changing views on the role of abscisic acid in root and shoot growth responses to water stress. Plant, Cell and Environment. 25(2): 211-222.

Shavrukov, Y., Kurishbayev, A., Jatayev, S., Shvidchenko, V., Zotova, L., Koekemoer, F., Langridge, P. (2017). Early Flowering as a Drought Escape Mechanism in Plants: How Can It Aid Wheat Production? Frontiers in Plant Science. 17(8): 1950.

Shepherd, T.G., (2016). A Common Framework for Approaches to Extreme Event Attribution. Current Climate Change Reports. 2: 28–38.

Shiferaw, B., Smale, M., Braun, H.J., Duveiller, E., Reynolds, M., Muricho, G., (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Security.

Skirycz, A., De Bodt, S., Obata, T., De Clercq, I., Claeys, H., De Rycke, R., Andriankaja, M., Van Aken, O., Van Breusegem, F., Fernie, A.R., Inze, D., (2010). Developmental Stage Specificity and the Role of Mitochondrial Metabolism in the Response of Arabidopsis Leaves to Prolonged Mild Osmotic Stress. Plant Physiology. 152: 226-244.

Slot, M., Zaragoza-Castells, J., & Atkin, O. K. (2008). Transient shade and drought have divergent impacts on the temperature sensitivity of dark respiration in leaves of Geum urbanum. Functional Plant Biology. 35(11): 1135-1146.

Sobeih, WY., Dodd, IC., Bacon, MA., Grierson, D., Davies, WJ. (2004). Long-distance signals regulating stomatal conductance and leaf growth in tomato (Lycopersicon esculentum) plants subjected to partial root-zone drying. Journal of Experimental Botany. 55(407): 2353-2363.

Sonder, Kai, (2016). Global map of wheat mega-environments. CIMMYT Research Data & Software Repository Network, V4

Stapper, M., & Fischer, R. A. (1990). Genotype, sowing date and plant spacing influence on high-yielding irrigated wheat in southern New South Wales. III.* Potential yields and optimum flowering dates. Australian Journal of Agricultural Research. 41(6): 1043-1056.

Stitt, M., Lunn, J., & Usadel, B. (2010). Arabidopsis and primary photosynthetic metabolism - More than the icing on the cake. Plant Journal. 61(6): 1067-91.

Stothard, P (2000) The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques. 28(6): 1102-1104.

Sun, L., Zhang, Q., Wu, J., Zhang, L., Jiao, X., Zhang, S., Zhang, Z., Sun, D., Lu, T., Sun, Y., (2014). Two Rice Authentic Histidine Phosphotransfer Proteins, OsAHP1 and OsAHP2, Mediate Cytokinin Signalling and Stress Responses in Rice. Plant Physiology. 165(1): 335-345.

Sun, X., Li, Y., Cai, H., Bai, X., Ji, W., Ding, X., & Zhu, Y. (2012). The Arabidopsis AtbZIP1 transcription factor is a positive regulator of plant tolerance to salt, osmotic and drought stresses. Journal of Plant Research. 125(3): 429-438.

Sun, Y., & Yu, D. (2015). Activated expression of AtWRKY53 negatively regulates drought tolerance by mediating stomatal movement. Plant Cell Reports. 34(8): 1295-1306.

Takahashi, M., & Asada, K. (1988). Superoxide production in aprotic interior of chloroplast thylakoids. Archives of Biochemistry and Biophysics. 267(2): 714-22.

Tenhaken, R., (2015). Cell wall remodeling under abiotic stress. Frontiers in Plant Science. 5: 771.

Tezara, W., Mitchell, V.J., Driscoll, S.D., Lawlor, D.W., (1999). Water stress inhibits plant photosynthesis by decreasing coupling factor and ATP. Nature. 401: 914–917.

Tichá, I., (1982). Photosynthetic characteristics during ontogenesis of leaves: 7. Stomata density and sizes. Photosynthetica. 16(3): 375- 471.

Tilman, D. (1998). The greening of the green revolution. Nature. 396: 211-212.

Trillo, N., & Fernández, R. J. (2005). Wheat plant hydraulic properties under prolonged experimental drought: Stronger decline in root-system conductance than in leaf area. Plant and Soil. 277: 277–284.

United Nations. (2017). UN, 2017. World Popul. Prospect.

Vassileva, V., Signarbieux, C., Anders, I., & Feller, U. (2011). Genotypic variation in drought stress response and subsequent recovery of wheat (Triticum aestivum L.). Journal of Plant Research. 124(1): 147-154.

Verkman AS: (1992). Water channels in cell membranes. Annual Review of Physiology. 54:97-108.

Villa, T. C. C., Maxted, N., Scholten, M., & Ford-Lloyd, B. (2005). Defining and identifying crop landraces. Plant Genetic Resources. 3(3): 373-384.

Wang, M., & Casey, P. J. (2016). Protein prenylation: Unique fats make their mark on biology. Nature Reviews Molecular Cell Biology. 17(2): 110-122.

Wang, Y., Ying, J., Kuzma, M., Chalifoux, M., Sample, A., McArthur, C., Huang, Y. (2005). Molecular tailoring of farnesylation for plant drought tolerance and yield protection. Plant Journal. 43(3): 413-424.

Wang, R., Zhang, Y., Kieffer, M., Yu, H., Kepinski, S., & Estelle, M. (2016). HSP90 regulates temperature-dependent seedling growth in Arabidopsis by stabilizing the auxin co-receptor F-box protein TIR1. Nature Communications. 7: 10269.

Wheeler, T., von Braun, J., (2013). Climate Impacts on Global Food Security. Science. 341(6145): 508-513.

Wilkinson, S. and Davies, W. J. (2008). Manipulation of the apoplastic pH of intact plants mimics stomatal and growth responses to water availability and microclimatic variation. Journal of Experimental Botany. 59(3): 619-631.

Williams, M., Rastetter, E.B., Fernandes, D.N., Goulden, M.L., Wofsy, S.C., Shaver, G.R., Melillo, J.M., Munger, J.W., Fan, S.M., Nadelhoffer, K.J., (1996). Modelling the soil-plant-atmosphere continuum in a Quercus-acer stand at Harvard forest: The regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. Plant, Cell Environment. 19(18): 911-927.

Woodward, F.I., (1987). Stomatal numbers are sensitive to increases in CO2 from pre-industrial levels. Nature. 327: 617–618.

Wu, J. R., Wang, L. C., Lin, Y. R., Weng, C. P., Yeh, C. H., & Wu, S. J. (2017). The Arabidopsis heat-intolerant 5 (hit5)/enhanced response to aba 1 (era1) mutant reveals the crucial role of protein farnesylation in plant responses to heat stress. New Phytologist. 213(3): 1181-1193.

Yang, W., Li, Y., Yin, Y., Qin, Z., Zheng, M., Chen, J., Luo, Y., Pang, D., Jiang, W., Li, Y., Wang, Z., (2017). Involvement of ethylene and polyamines biosynthesis and abdominal phloem tissues characters of wheat caryopsis during grain filling under stress conditions. Scientific Reports. 7: 46020.

Yuan, F., Yang, H., Xue, Y., Kong, D., Ye, R., Li, C., Pei, Z. M. (2014). OSCA1 mediates osmotic-stress-evoked Ca 2+ increases vital for osmosensing in Arabidopsis. Nature. 514(7522): 367-371.

Zeuthen, T., & Klaerke, D. A. (1999). Transport of water and glycerol in aquaporin 3 is gated by H+. Journal of Biological Chemistry. 274(31): 21631-21636.

Zhang, J. and Davies, W.J., 1987. Increased synthesis of ABA in partially dehydrated root tips and ABA transport from roots to leaves. Journal of Experimental Botany. 38(12): 2015-2023.

Zhang, J., Nguyen, H.T., Blum, A., 1999. Genetic analysis of osmotic adjustment in crop plants. Journal of Experimental Botany. 50(332): 291–302.

Zhang, X.-C., Millet, Y. A., Cheng, Z., Bush, J., & Ausubel, F. M. (2015). Jasmonate signalling in Arabidopsis involves SGT1b–HSP70–HSP90 chaperone complexes. Nature Plants. 1: 15049.

Zhang, Y., Pan, J., Huang, X., Guo, D., Lou, H., Hou, Z., Su, M., Liang, R., Xie, C., You, M., Li, B., 2017. Differential effects of a post-anthesis heat stress on wheat (Triticum aestivum L.) grain proteome determined by iTRAQ. Sci. Rep. 7(1): 3468.

Zheng, B., Biddulph, B., Li, D., Kuchel, H., & Chapman, S. (2013). Quantification of the effects of VRN1 and Ppd-D1 to predict spring wheat (Triticum aestivum) heading time across diverse environments. Journal of Experimental Botany. 64(12): 3747–3761.

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., … Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. Nature Genetics. 42: 355–360.

Zhu, J.-K., 2002. Salt and Drought Stress Signal Transduction in Plants. Annual Review of Plant Biology. 53: 247-273.

Zhu, J.K., 2016. Abiotic Stress Signaling and Responses in Plants. Cell. 167(2): 313-324.

Appendix:

			A	B	C	D			
1			4	3	2	1			
2			3	2	1	4			
3			2	1	4	3			
4			1	4	3	2			

Supplementary Figure 1: Schematic depicting the landrace position allocations of hydroponic samples in the Latin square design used in the PEG6000 and PEG8000 pilot experiments.

		A			B			C	
1		1			2			3	
2		3			1			2	
3		2			3			1	
4		1			2			3	

Supplementary Figure 2: Schematic depicting the position of Cadenza and Kronos lines in hydroponics in the morphological analysis experiments.

Supplementary Figure 3: Image depicting the microbial growth typical of the polyethylene glycol 8000 (PEG8000) treatment boxes.

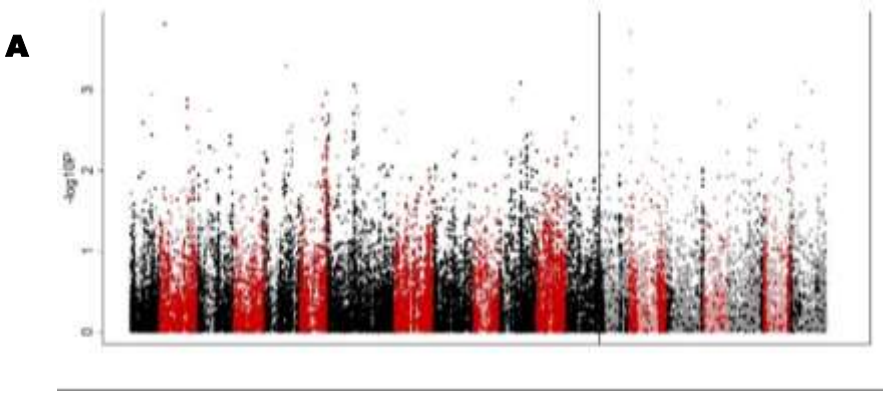
Supplementary Table 1: Information for the 135 YoGI panel landraces that were phenotyped in this study.

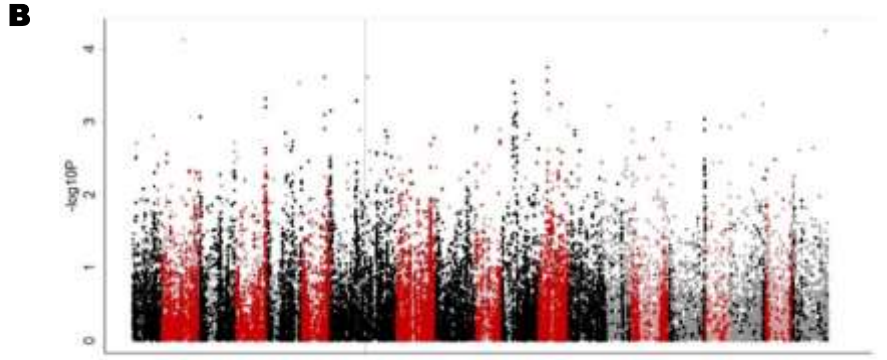
Landrace Number	Country of Origin	Landrace Name
1	China	YoGI_002
2	Kenya	YoGI_004
3	Bolivia	YoGI_005
4	Pakistan	YoGI_006
5	Nepal	YoGI_007
6	China	YoGI_008
7	Kenya	YoGI_009
8	Kenya	YoGI_010
9	Kenya	YoGI_011
10	Kenya	YoGI_012
11	Kenya	YoGI_013
12	United States	YoGI_014
13	South Africa	YoGI_015
14	South Africa	YoGI_016
15	South Africa	YoGI_017
16	South Africa	YoGI_018
17	South Africa	YoGI_019
18	South Africa	YoGI_020
19	India	YoGI_021
20	India	YoGI_022
21	India	YoGI_023
22	India	YoGI_024
23	India	YoGI_025
24	India	YoGI_026
25	Turkey	YoGI_027
26	India	YoGI_028
27	Egypt	YoGI_029
28	Egypt	YoGI_030
29	Bolivia	YoGI_031
30	Brazil	YoGI_032
31	Pakistan	YoGI_033
32	South Africa	YoGI_034
33	Lebanon	YoGI_035
34	Argentina	YoGI_036
35	Australia	YoGI_037
36	Argentina	YoGI_038
37	Australia	YoGI_040
38	United States	YoGI_042

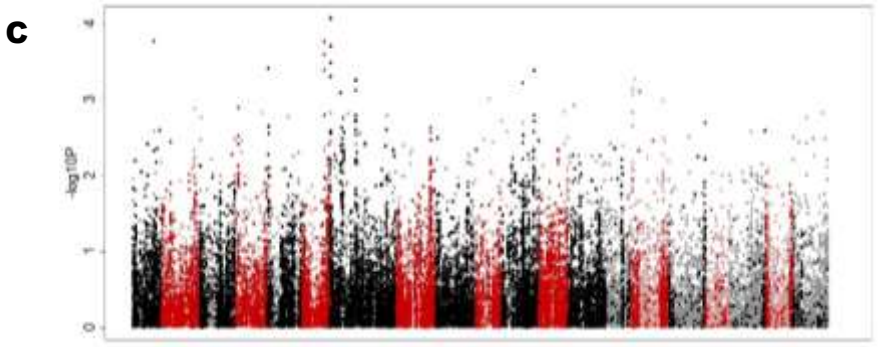
39	United States, Minnesota	YoGI_043
40	Australia	YoGI_046
41	United States	YoGI_048
42	Australia	YoGI_049
43	China	YoGI_050
44	Uruguay	YoGI_051
45	United States	YoGI_052
46	Turkey	YoGI_053
47	United States	YoGI_054
48	Egypt	YoGI_055
49	United States	YoGI_056
50	Argentina	YoGI_058
51	Brazil	YoGI_059
52	Yemen	YoGI_060
53	Turkey	YoGI_061
54	Chile	YoGI_062
55	Ecuador	YoGI_063
56	Australia	YoGI_067
57	United States	YoGI_069
58	Mexico	YoGI_072
59	Mexico	YoGI_073
60	Mexico	YoGI_075
61	United States	YoGI_078
62	Iraq	YoGI_083
63	Australia	YoGI_084
64	Australia	YoGI_087
65	India	YoGI_088
66	India	YoGI_089
67	India	YoGI_101
68	India	YoGI_102
69	Italy	YoGI_109
70	India	YoGI_114
71	France	YoGI_117
72	China	YoGI_118
73	Morocco	YoGI_119
74	Spain	YoGI_120
75	Spain	YoGI_125
76	India	YoGI_127
77	India	YoGI_137
78	Egypt	YoGI_140
79	Tunisia	YoGI_142
80	Spain	YoGI_143
81	Burma	YoGI_144
82	China	YoGI_145

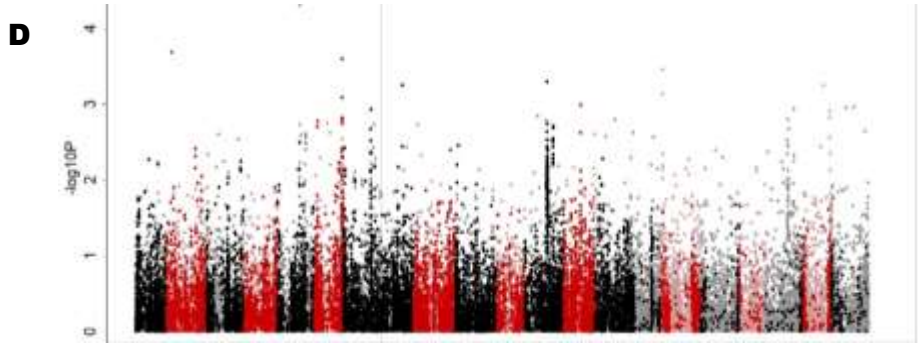
83	Iran	YoGI_150
84	Spain	YoGI_151
85	India	YoGI_152
86	Morocco	YoGI_153
87	Canary Islands	YoGI_155
88	Spain	YoGI_156
89	Greece	YoGI_158
90	Cyprus	YoGI_160
91	Cyprus	YoGI_161
92	Turkey	YoGI_162
93	Turkey	YoGI_163
94	Egypt	YoGI_164
95	Burma	YoGI_167
96	Palestine	YoGI_191
97	India	YoGI_197
98	China	YoGI_199
99	China	YoGI_201
100	Afghanistan	YoGI_205
101	Afghanistan	YoGI_207
102	Afghanistan	YoGI_208
103	Afghanistan	YoGI_211
104	Morocco	YoGI_218
105	Australia	YoGI_220
106	Spain	YoGI_222
107	Greece	YoGI_224
108	Greece	YoGI_225
109	Greece	YoGI_226
110	Iran	YoGI_228
111	Portugal	YoGI_230
112	Iran	YoGI_235
113	Greece	YoGI_256
114	India	YoGI_257
115	China	YoGI_258
116	Iran	YoGI_261
117	China	YoGI_264
118	India	YoGI_267
119	India	YoGI_268
120	Algeria	YoGI_271
121	Italy	YoGI_291
122	Tunisia	YoGI_299
123	Tunisia	YoGI_300
124	Lesotho	YoGI_308
125	Mongolia	YoGI_309
126	Paraguay	YoGI_310

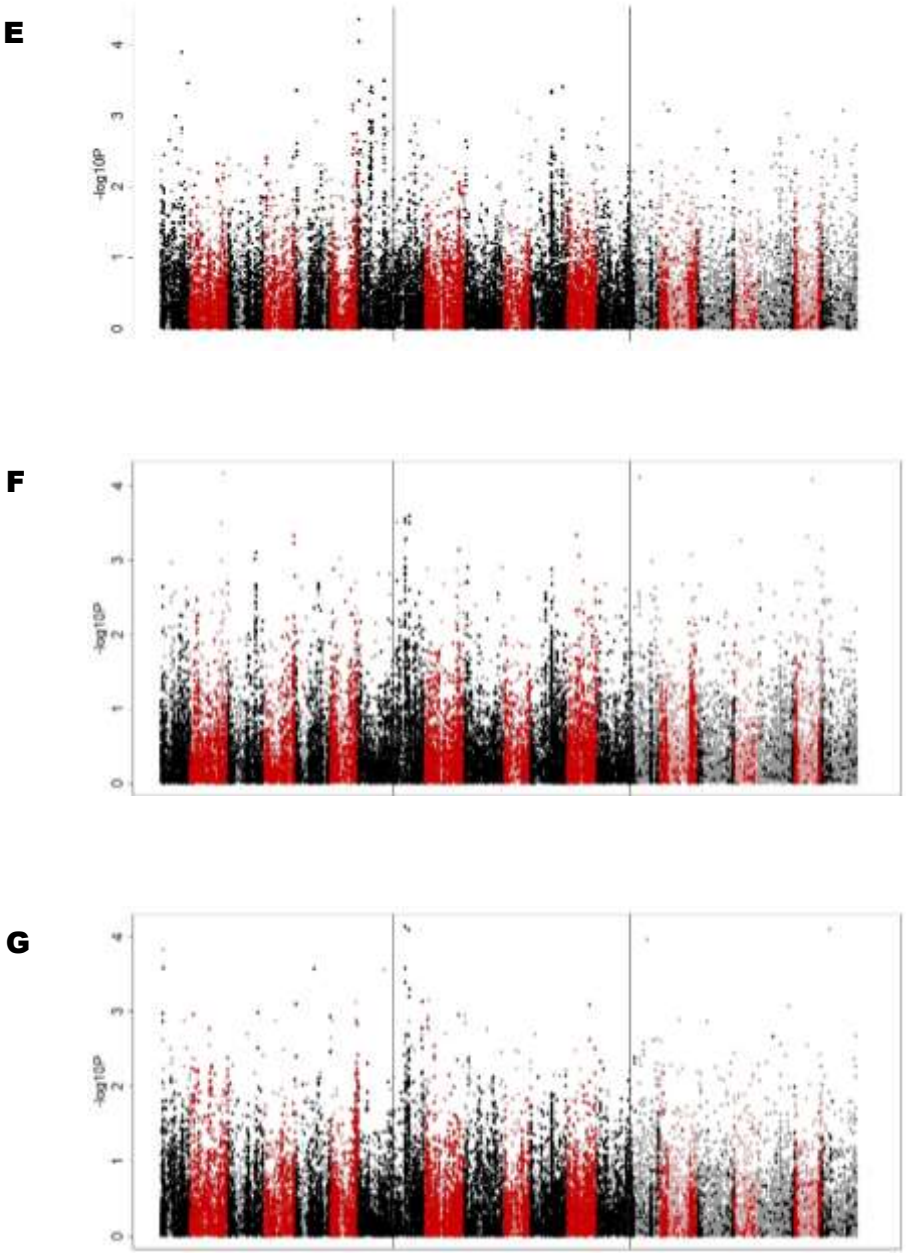
127	New Zealand	YoGI_311
128	New Zealand	YoGI_314
129	Angola	YoGI_318
130	Saudi Arabia	YoGI_320
131	Mongolia	YoGI_323
132	Peru	YoGI_324
133	Sudan	YoGI_325
134	Mongolia	YoGI_330
135	Colombia	YoGI_333


A
-log10P
0
1
2
3


B
-log10P
0
1
2
3
4

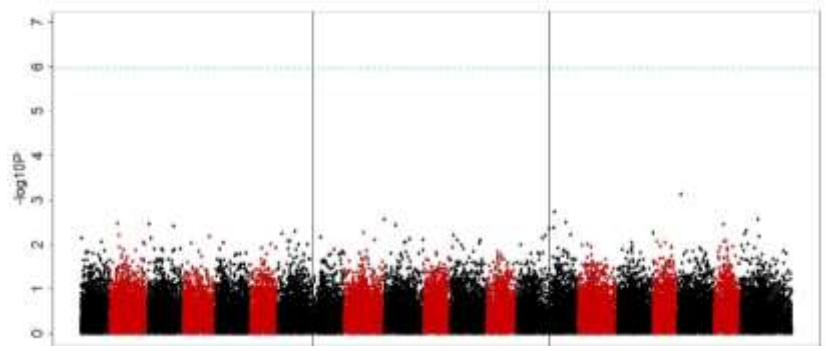

C
-log10P
0
1
2
3
4


D
-log10P
0
1
2
3
4

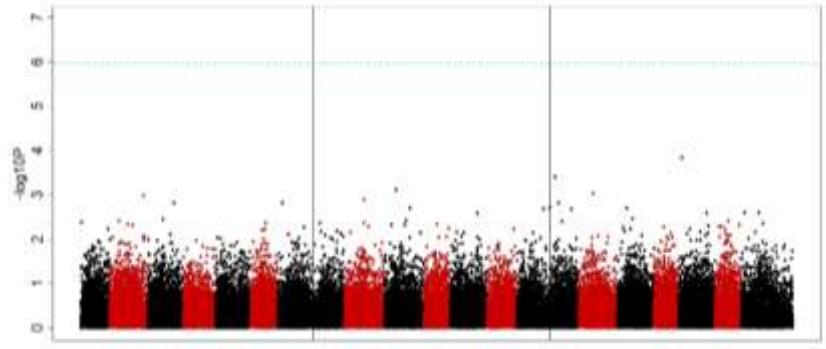


Supplementary Figure 4: Manhattan plots depicting SNP marker trait association test -log10 p-values for normalised change in fresh shoot biomass (A), fresh total biomass (B), dry root biomass (C), dry shoot biomass (D), total dry biomass (E), fresh root fraction (F) and dry root fraction (G) trait data. Analysis was performed in R 3.5.1 using GAPIT.

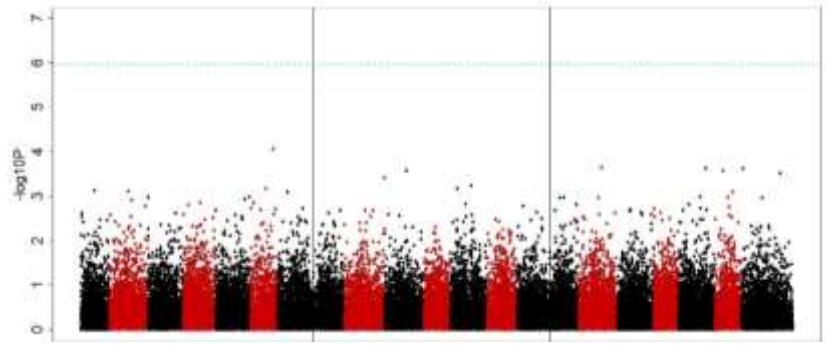
A


-log10P
0
1
2
3
4
5
6
7

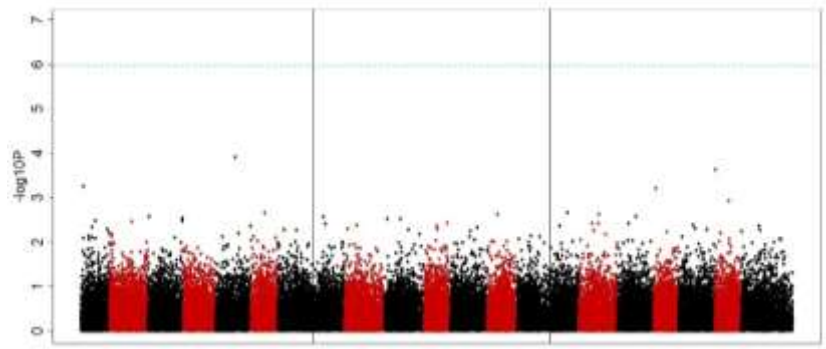
B

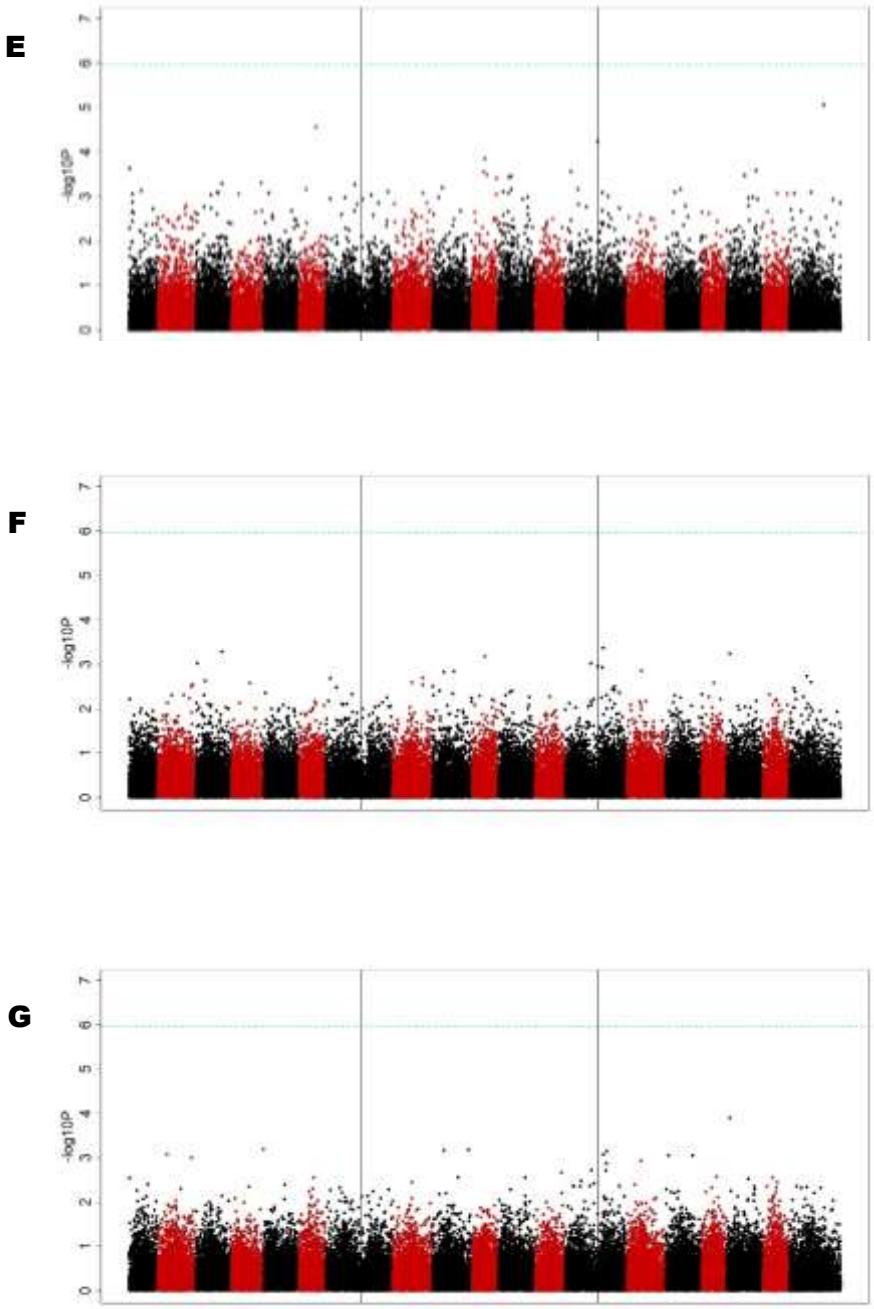

-log10P
0
1
2
3
4
5
6
7

C


-log10P
0
1
2
3
4
5
6
7

D


-log10P
0
1
2
3
4
5
6
7



Supplementary Figure 5: Manhattan plots depicting GEM marker trait association test -log10 p-values for normalised change in fresh shoot biomass (A), fresh total biomass (B), dry root biomass (C), dry shoot biomass (D), total dry biomass (E), fresh root fraction (F) and dry root fraction (G) trait data. Analysis was performed in R 3.5.1 using Regress Multi.

Supplementary Table 2: Top NCBI protein sequence identity matches for TraesCS6A01G405600.1 without taxa specification (*Aegilops tauschii/Tritcum urartu*) and with Arabidopsis taxa specification (*Arabidopsis thaliana*).

Sequence Annotation	Species	Query Cover (%)	Percentage Identity (%)	Sequence Reference
Aquaporin PIP1-5-like	*Aegilops tauschii subsp. tauschii*	100	100	XP_020158265.1
Aquaporin PIP1-5	*Triticum urartu*	100	98.97	EMS50328.1
Plasma membrane intrinsic protein 1;5	*Arabidopsis thaliana*	99	85.91	XP_020873658.1

Supplementary Table 3: Twenty-one gene expression markers from Regress_multi analysis of NCFRB data with FDR p-values lower than 0.05.

Sequence Name	Functional Annotation Description	FDR p-value
TraesCS1A01G163700.2	MATH domain containing protein	7.63E-07
TraesCS4D01G027600.1	Calcium-dependent protein kinase	2.15E-05
TraesCS2B01G205800.1	IgA FC receptor	4.19E-05
TraesCS7A01G002500.1	Serine carboxypeptidase, putative	8.96E-05
TraesCS1D01G259700.1	Serine/threonine-protein kinase	0.00046
TraesCS4B01G192200.1	Protein kinase family protein, putative, expressed	0.001185
TraesCS7D01G111600.1	Flowering locus T	0.001439
TraesCS4B01G319100.1	Chaperone protein dnaJ	0.001687
TraesCS3A01G488400.1	Calcium binding protein	0.002621
TraesCS6D01G209500.1	Glycolipid transfer protein domain-containing protein	0.002988
TraesCS3D01G312700.1	Methylthioribulose-1-phosphate dehydratase	0.003538
TraesCS5D01G128200.1	MYB transcription factor	0.004327

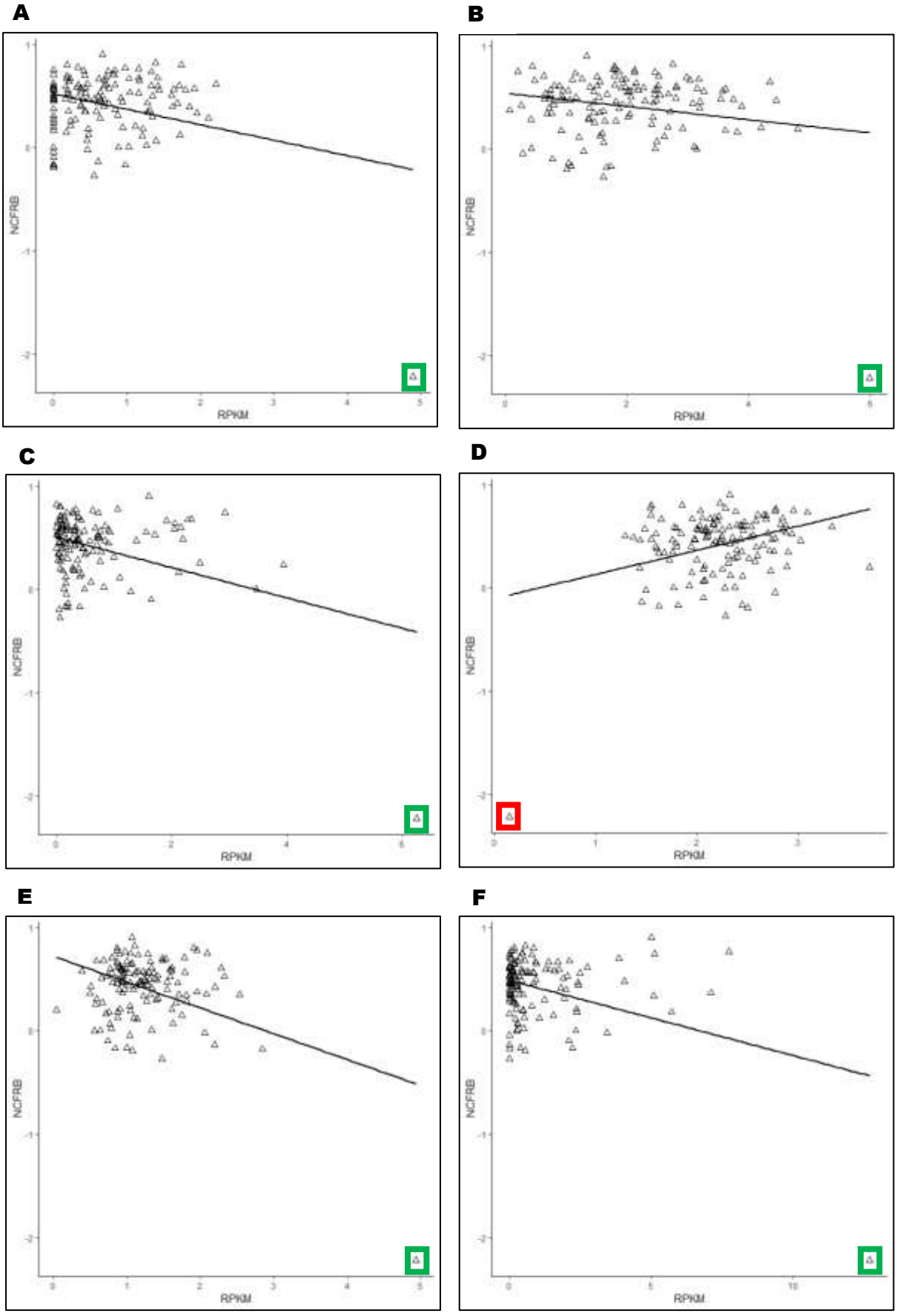
TraesCS6A01G262000.1	Senescence-associated family protein (DUF581)	0.00426
TraesCS1D01G037500.1	Disease resistance protein RPM1	0.014068
TraesCS2B01G468300.1	Phosphoenolpyruvate carboxykinase (ATP)	0.01825
TraesCS4A01G305300.1	Ribosomal protein	0.021707
TraesCS1A01G259900.1	Serine/threonine-protein kinase	0.022396
TraesCS3D01G176100.1	Germin-like protein 1	0.025768
TraesCS1B01G244000.1	Complex 1 protein (LYR family)	0.029748
TraesCS2D01G211000.1	F-box domain	0.031834
TraesCS3B01G013100.1	UDP-glucuronosyl/UDP-glucosyltransferase	0.040246

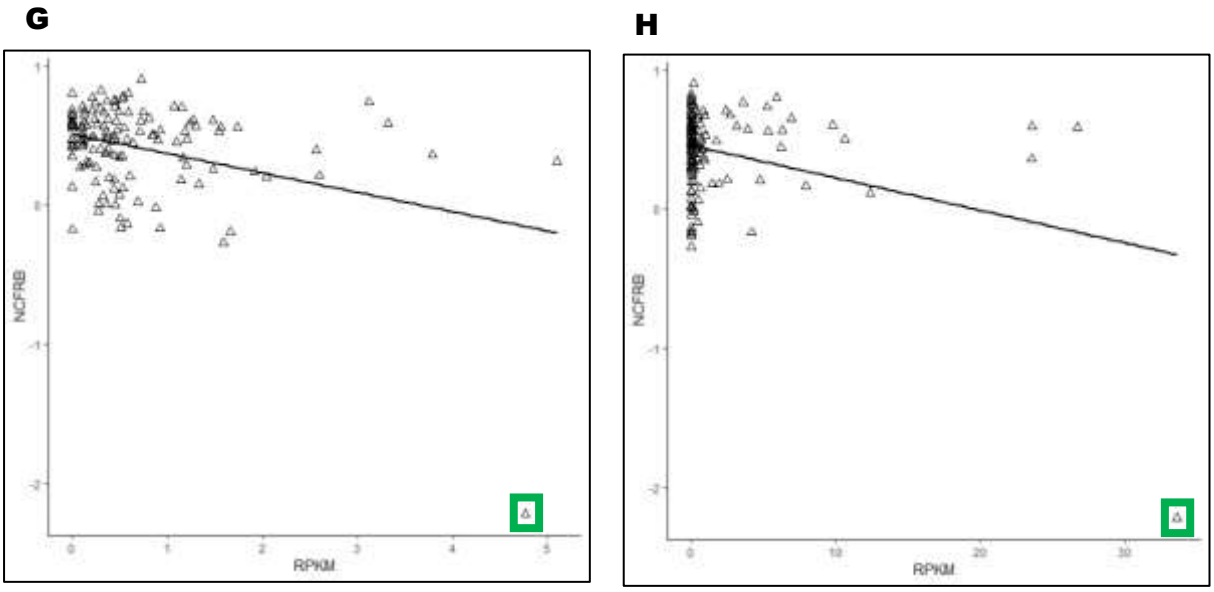
Supplementary Table 4: List of designed primer sequences used to amplify PIP1;5 DNA extracted from SALK lines for genotyping.

SALK Line	Primer Name	Primer Sequence
SALK056898	LP	CTTGCTGCTCTGTACCATCAG
SALK056898	RP	GGGTTTTGTTTTGTATTGCAAG
SALK056041	LP	GTTCACAAGAAGATCGCTTGC
SALK056041	RP	GCCTCATTTTGTCACATCTCC

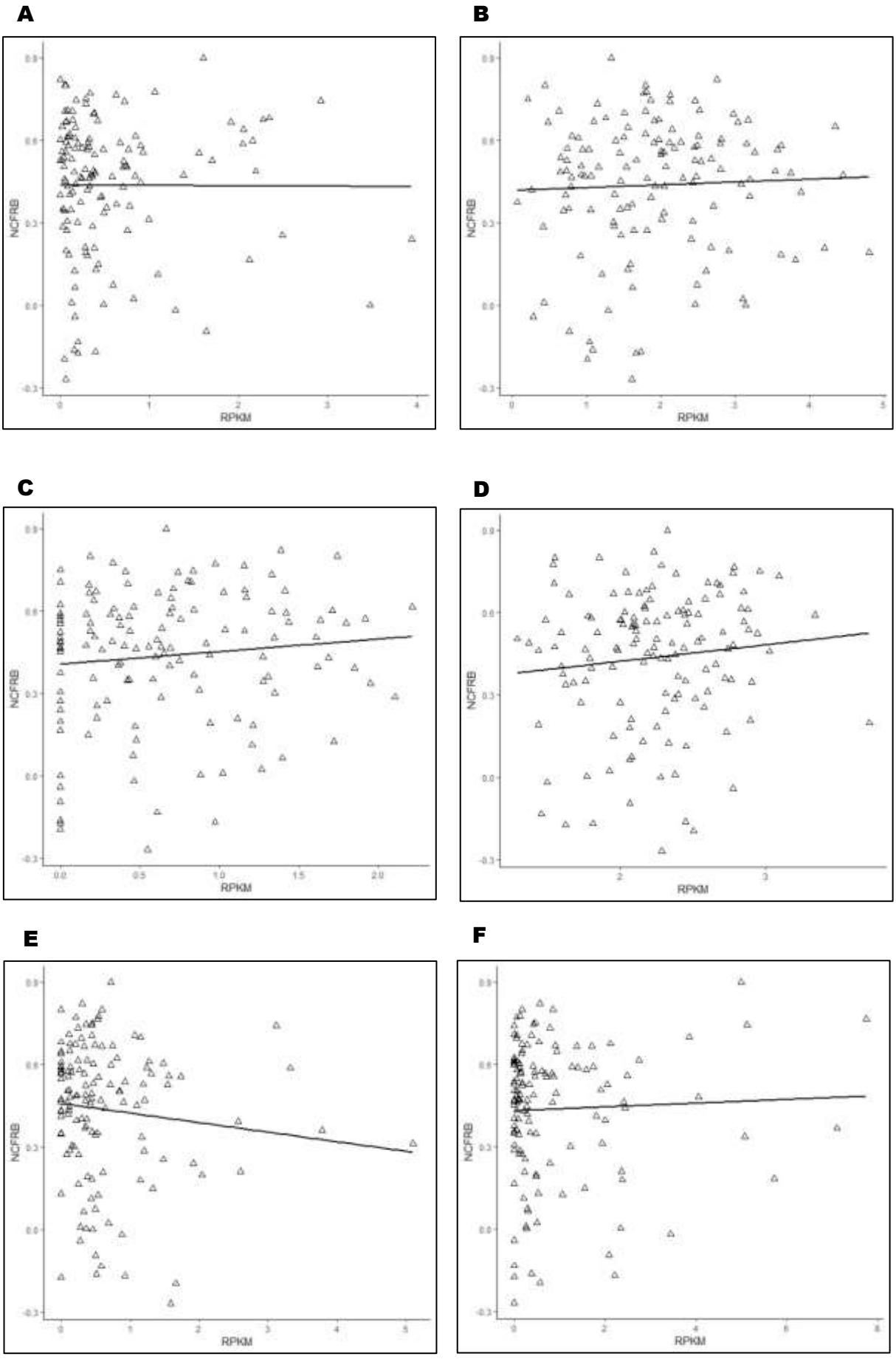
Supplementary Table 5: List of designed primer sequences used to amplify a section of the Kronos4610 TraesCS4B01G319100.1 homolog. Left primer 1 and right primer 1 were the most successful and were used to generate PCR products that were sent for sequencing.

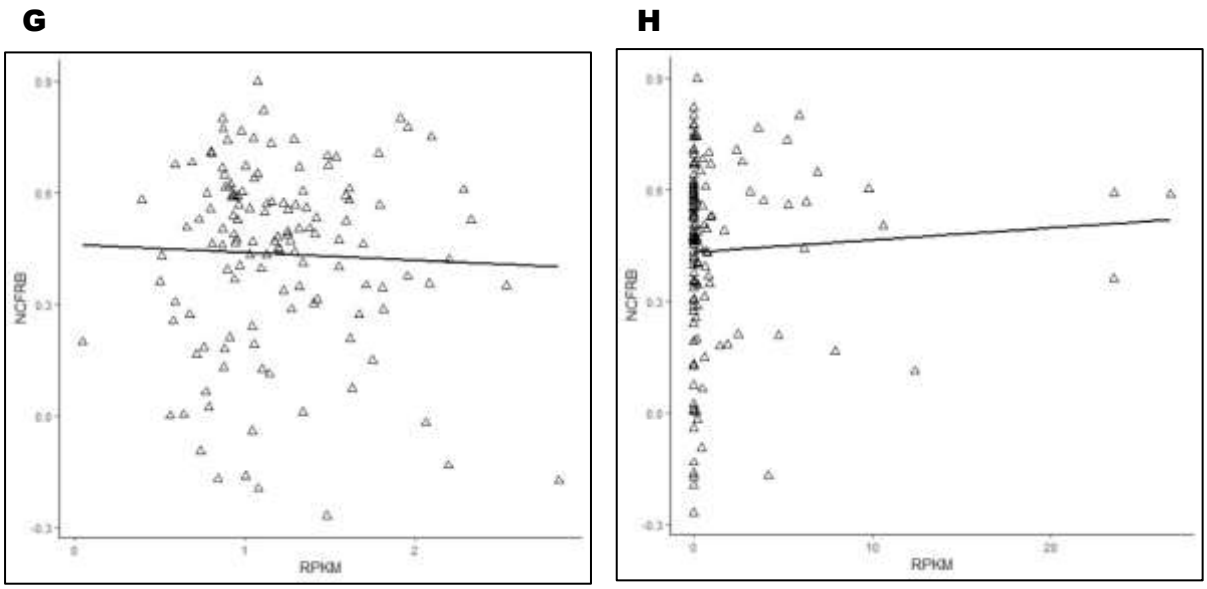
Primer Name	Primer Sequence
Left Primer 1	ATGTCTGTGCTGATCCAGAG
Right Primer 1	ACCTTTCTCCCATTTAATCTGG
Left Primer 2	CGTACTGAGGAATCTTTTCAGAGT
Right Primer 2	TACCAATCCTGAAATCACCGC


A
B
C
D
E
F
NCFRB
RPKM

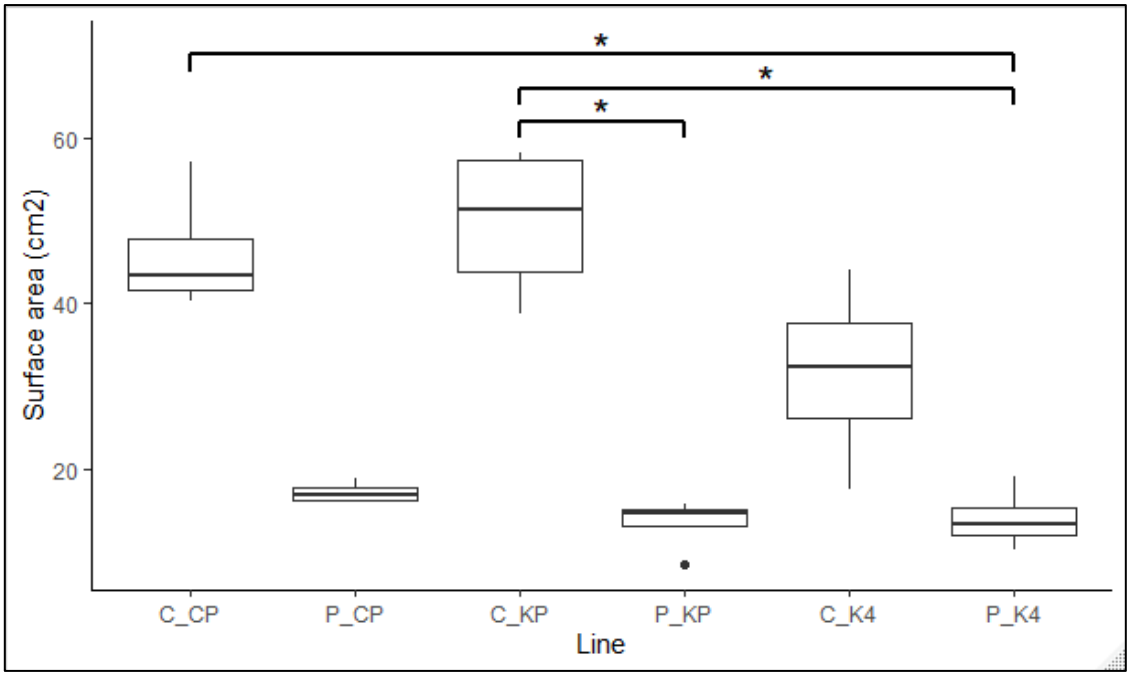


Supplementary Figure 6: Scatter plots showing the relationship between RPKM and NCFRB for the eight candidate genes investigated for the 134 landraces measured in the panel screen experiment. A, TraesCS1D01G259700_1; B, TraesCS3B01G013100_1; C, TraesCS3D01G312700_1; D, TraesCS4B01G319100_1; E, TraesCS4D01G027600.1; F, TraesCS5D01G128200_1; G, TraesCS6A01G262000_1; H, TraesCS7D01G111600_1. YoGI 28 data points indicated in green, putative HSP40 of interest indicated in red.

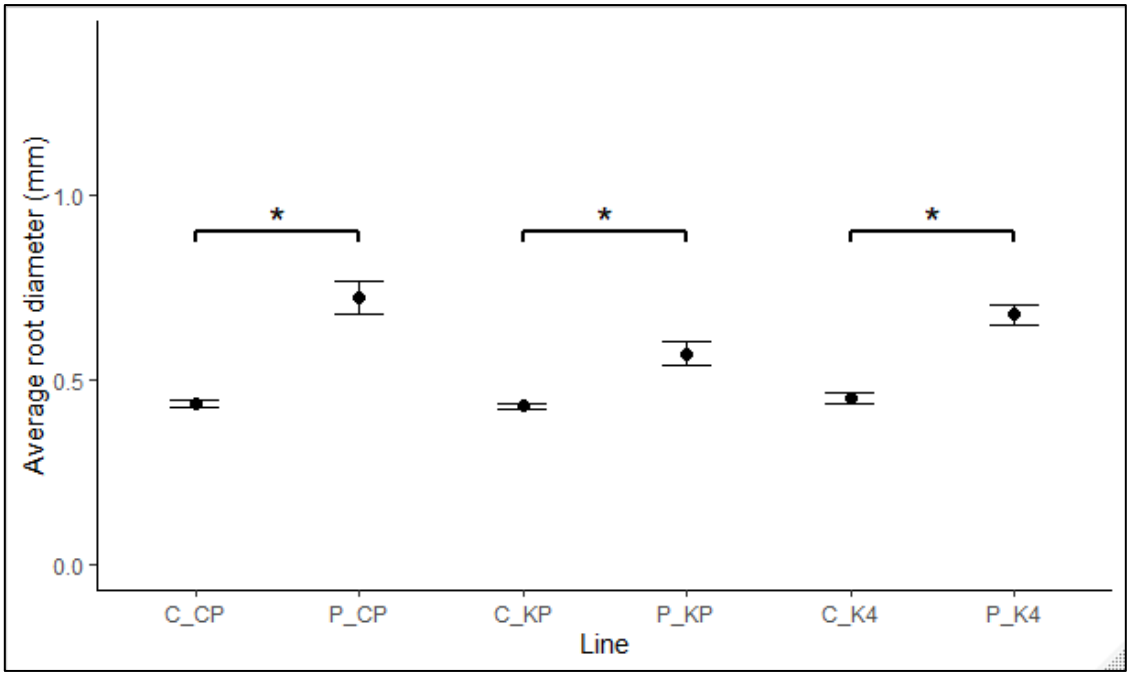

A
B
C
D
E
F
NCFRB
RPKM



Supplementary Figure 7: Scatter plots showing the relationship between RPKM and NCFRB for the eight candidate genes investigated for the 134 landraces measured in the panel screen experiment (YoGI 28 removed). A, TraesCS1D01G259700_1; B, TraesCS3B01G013100_1; C, TraesCS3D01G312700_1; D, TraesCS4B01G319100_1; E, TraesCS4D01G027600.1; F, TraesCS5D01G128200_1; G, TraesCS6A01G262000_1; H, TraesCS7D01G111600_1.



Supplementary Figure 8: Plot comparing root surface area (cm2) of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (C) and supplemented with 7.5% (w/v) polyethylene glycol 8000 (P). Cadenza parent (_CP), Kronos parent (_KP) and Kronos4610 (_K4). All lines, N=4 (Kruskal-Wallis, $p < 0.05$ (*)). Root surface area data was generated using WinRHIZO.



Supplementary Figure 9: Plot comparing average root diameter (mm) of *T. aestivum* individuals grown in static hydroponic systems containing 55% Hoagland's solution, 45% water (C) and supplemented with 7.5% (w/v) polyethylene glycol 8000 (P). Cadenza parent (_CP), Kronos parent (_KP) and Kronos4610 (_K4). All lines, N=4 (ANOVA, $p < 0.05$ (*), $p < 0.1$ (*)). Average root diameter data was generated using WinRHIZO. Other significant relationships are excluded from plot for clarity but can be found in Supplementary table 6.

Supplementary Table 6: Showing results of Tukey multiple comparisons of means 95% family-wise confidence level test for ANOVA test completed, for which results are presented in Supplementary Figure 9.

Lines	FDR-adjusted p-value
P_CP-C_CP	0.0000053
P_KP-C_CP	0.0201418
P_KP-P_CP	0.0075364
C_K4-P_CP	0.0000038
P_KP-C_KP	0.0000113
P_K4-C_KP	0.0137711
C_K4-P_KP	0.0000375
P_K4-C_K4	0.0001212

R Scripts:

```
title: Script to take the mean of any duplicates (according to landrace) in FreshDry

author: "EHSmith"

date: "8 March 2019"

#Subset the data into Control and PEG.

library(readxl)

FreshDry <- read_excel("R:/rsrch/ah1309/mscr/ehs510/Masters/Associative Transcriptomics Experiment/Data/Primary Data/FreshDry.xlsx")

FreshDry

Control1 = subset(FreshDry, FreshDry$Condition == "Control")

Control <- na.omit(Control1)

Control

PEG1 = subset(FreshDry, FreshDry$Condition != "Control")

PEG <- na.omit(PEG1)

PEG

#Checks for duplicates in PEG. To do the same for Control just substitute the Control data

library(plyr)

PEG_duplicates <- duplicated(PEG$Landrace)

count(PEG_duplicates == TRUE)

#Using the dpylr package, replace any duplicates with the mean of their value-would recommend double checking the results against what you think they should be.

library(dplyr)

############ PEG Duplicates ##################

PEG_fixed = ddply(PEG, .(Landrace), function(x) c(
mean(x[,4]),mean(x[,5]),mean(x[,6]),mean(x[,7]),mean(x[,8]),mean(x[,9]),mean(x[,10]),mean(x[,11]), na.rm=TRUE))

colnames(PEG_fixed)
```

```
colnames(PEG_fixed) <- c("Landrace", "FreshMeanRoot", "FreshMeanShoot",
"FreshMeanTotal","DryRootMean","DryShootMean","DryMeanTotal","FreshRF","DryRF")

############ Control Duplicates #################

Control_fixed = ddply(Control, .(Landrace), function(x) c(
mean(x[,4]),mean(x[,5]),mean(x[,6]),mean(x[,7]),mean(x[,8]),mean(x[,9]),mean(x[,10]),mean(x[,11]),
na.rm=TRUE))

colnames(Control_fixed)

colnames(Control_fixed) <- c("Landrace", "FreshMeanRoot", "FreshMeanShoot",
"FreshMeanTotal","DryRootMean","DryShootMean","DryMeanTotal", "FreshRF","DryRF")

#if any 'TRUE' in PEG_duplicates, something's gone wrong

PEG_duplicates = duplicated(PEG_fixed$Landrace)

PEG_duplicates

PEG_fixed

#Dataframe containing mean landrace values for PEG treated samples

Control_duplicates = duplicated(Control_fixed$Landrace)

Control_duplicates

Control_fixed

#Dataframe containing mean landrace values for no treatment control samples

write.table(PEG_fixed,"R:PEG_fixed.txt",sep="\t", row.names=FALSE)

write.table(Control_fixed,"Control_fixed.txt",sep="\t", row.names=FALSE)
```

R Script 1: Script used to calculate landrace mean trait values for PEG and control treatment

Acknowledgements:

I would like to offer my warm thanks to those who so cheerily populated L2 during my brief spell at the University, with particular mention addressed to the time and efforts so disinterestedly imparted by James Ronald, Kayla McCarthy and even Amanda Davis.

Most especially, I wish to thank Andrea, over and again, for her tireless adaptations to often challenging circumstances, and for the ever many unspoken lessons, that if to be fortunate, I should not forget with haste.