Identification and analysis of Gene Regulatory Networks involved in the drought stress response in *Arabidopsis thaliana*

Sunitha Subramaniam

A thesis submitted for the degree of Doctor of Philosophy

School of Biological Sciences
University of Essex
April 2016

Abstract

There is a growing need to engineer an increased yield in food crops coupled with greater resistance to environmental stresses, such as drought, to meet the requirement of a growing population. As plant responses to drought (and other abiotic stress) are complex, many studies have attempted to understand the drought response in a holistic manner, for example, by analysing transcriptomics data obtained from plants subjected to drought. Preliminary work of obtaining time-series microarray data from a slow-drying experiment of *Arabidopsis* was used to construct Gene Regulatory Networks using Variational Bayesian State Space Modelling. This led to the identification of a number of transcription factors *RAP2.12*, *FD*, *BHLH038*, *ANL2*, and two unknown genes, *UKTF* and *POZ*, as possibly being important regulatory genes in the drought response. Loss- and gain-of-function mutants of these genes, as well as those of *AGL22* identified by Bechtold *et al.* (2016), were phenotyped under drought conditions. Only the flowering time *AGL22* showed a drought phenotype. Network connections in the gene network of *AGL22* were tested by qPCR. Drought responsive transcription factors, such as *DREB1A* and *WRKY20* were found to be induced by *AGL22*.

Acknowledgements

I am grateful to have worked with some wonderful people during my Ph.D. I would like to thank my supervisor Uli Bechtold for her support and guidance throughout my Ph.D. and I am also grateful to Phil Mullineaux for his advice over the years.

I would also like to thank all the lovely people (past and current) of the Plant Group in Essex and various other people in and around the School of Biological Sciences for an enjoyable environment to work in, and I am happy to call them friends.

Finally, I would like to thank my family and friends for being there for me throughout my Ph.D.

List of figures

List of tables

List of abbreviations

°C – degree Celsius
µg – microgram
µl – microliter
µmol – micromoles
A – photosynthetic assimilation rate
Abs - absorbance
att – attachment sites
BaR – Basta® resistance gene
bp – base pair
CaMV 35S – Cauliflower Mosaic Virus 35S promoter
cDNA – complementary DNA
cm – centimetre
Col-0 – Columbia-0
d^{-1} – per day
DNA – Deoxyribonucleic acid
F.W. – fresh weight
g – gram
g – relative centrifugal force
GP2S – Gaussian process two-sample test
GRN – Gene Regulatory Network
g_s – stomatal conductance
H2O2 – Hydrogen peroxide
KanR – Kanamycin resistance gene
LB – Left border of T-DNA
log – logarithm
ml – millilitre
mM – millimolar
mmol – millimoles
nmol – nanomoles
OCS – Octopine Synthase
P – probability value
RB – Right border of T-DNA
rLWC – relative leaf water content
RNA – Ribonucleic acid
RO – reverse osmosis

rSWC – relative soil water content

s – second

TCAP – Temporal Clustering by Affinity Propagation

T-DNA – Transfer DNA

U – units

VBSSM – Variational Bayesian State Space Modelling

WT – wild-type

Chapter 1

Introduction

1.1. Increased demand for water

With a projected population of 9.1 billion people by 2050 (UNDESA, 2009), there will be a need for an increase in food production by 70% to meet the growing demand (Bruinsma, 2009). But with an increasing population comes not only a higher demand for food, but also an increased demand for more water resources and more land space. At the same time predicted climate change will cause changes in rainfall patterns, thus varying the amount of available water (WWAP, 2015). The demand for energy will increase by 60% in 2050 (Steer, 2010), and hence the need for biofuels will result in more and more land being used in non-food crop production (Comprehensive Assessment of Water Management in Agriculture, 2007). While only 20% of agricultural land is irrigated, 50% of the world's food production comes from these irrigated lands. Thus, decreasing water and land availability, coupled with increasing atmospheric temperatures (and hence increasing water loss), will have a huge impact on food crop production in time to come. These predicted scenarios portray an increasingly urgent need to breed in or 'introduce' higher water-use efficiency in crops while also maintaining, or hopefully increasing, crop yields to feed the growing population (WWAP, 2015).

1.2. Plant adaptation responses to water deficit

Plants are regularly subjected to environmental stresses during their lifetime, and hence have evolved a number of adaptation mechanisms to deal with stresses (Chaves *et al.*, 2003; Lawlor, 2013). Plant responses to drought can involve escape, avoidance or tolerance of water-deficit situations. During drought escape, plants complete their lifecycle before the onset of drought, while water is still available, and hence ensure reproduction is successful.

Drought avoidance involves the plants maintaining higher cellular water potential in the presence of decreasing levels of soil moisture. To do this, plants undergo adjustments to minimise the loss of water from the leaves, while trying to maximise the gain of water from the soil through the roots. Water deficit causes a loss of turgor due to a reduction of cellular water content. Loss of turgor in leaves (due to a decreased supply of water to the shoot region) directly causes a decrease in the growth rate of cells, which results in a decrease in leaf expansion and leaf area. Reduced stomatal conductance occurs in response to increasing water deficit, and thus there is a reduction in water-loss due to transpiration. As drought progresses, older leaves are shed to further conserve water by

minimising transpiration, while also ensuring that plant resources are diverted to younger leaves for growth. The partitioning of water resources away from the shoot and leaves is to enable an increase in water supply to roots to support their growth deeper into the moister areas of soil (Chaves *et al.*, 2003).

In drought tolerance, the cell adapts to the lower water potential by inducing the expression of genes needed for the adaptation process. Osmolytes and sugars are produced to help balance the change in water potential (Zhang *et al.*, 1999; Serraj and Sinclair, 2002); the Late Embryogenesis Abundant (LEA) proteins help in protecting cellular structure from crystallisation during drought (Hara *et al.*, 2001; Bravo *et al.*, 2003) and ROS detoxifying enzymes are produced to modulate the production of ROS (Moran *et al.*, 1994; Sharma and Dubey, 2005). Stomatal closure, and consequently photosynthesis, is also regulated by the hormone Abscisic Acid (ABA) and a number of regulatory genes that control stomatal function and photosynthetic capacity under stress (Schroeder *et al.*, 2001).

The changes observed during drought tolerance are a consequence of perception of an occurring water-deficit, transduction of the signal, expression of drought tolerance-related genes and production of proteins and osmolytes needed to overcome the effects of the reduction in cellular water content. Thus, a better understanding of the molecular events occurring during tolerance to water deficit might provide the basic knowledge needed to engineer plants for greater tolerance to prolonged periods of water deficit, while also maintaining productivity.

1.3. Perception of water-deficit signal

The process involved in sensing the osmotic stress component of drought at the cellular level is not well understood in plants. However, in yeast, the perception of the high osmolarity stress signal through the High Osmolarity Glycerol (HOG) pathway has been studied (Tamás *et al.*, 2000), and a similar pathway may also exist in plants. Under normal osmotic conditions, the histidine kinase SLN1, phosphorylates an intermediate protein YPD1, another histidine kinase, which ‘relays’ the phosphorylation to regulator protein SSK1 (Reiser *et al.*, 2003). The phosphorylated SSK1 remains inactive under normal osmolarity conditions. When external osmolarity increases, SLN1 is inhibited and SSK1 becomes dephosphorylated and activated. SSK1 acts through the mitogen-activated protein kinase (MAPK) pathway to activate the HOG pathway through HOG1,

which mediates adaptation of the yeast cell to changes in osmolarity. Another protein, SHO1, is also activated in response to high external osmolarity and also activates the HOG pathway (Reiser *et al.*, 2003). In *Arabidopsis*, a gene homologous to that coding for SLN1 has been identified (Urao *et al.*, 1999). However, the whole pathway involved in sensing the onset of water deficit has yet to be characterised and many components of the pathway are still unknown.

1.4. ABA plays a key role in drought stress response

The transduction of the signal may be mediated through the production of ABA or may be ABA-independent (Shinozaki and Yamaguchi-Shinozaki, 2007). The action of ABA leads to the expression of genes that aid plants in acclimatising to environmental stresses. Levels of ABA have been found to increase in plants as drought progresses (Jiang and Zhang, 2002; Wang *et al.*, 2004). This corresponds with an increase in the expression of genes that help in the acclimation of the plant to increasing water deficit – genes that control stomatal closure, to help minimise water loss, genes that stimulate the production of osmolytes and LEA proteins, and other stress responsive genes (Shinozaki and Yamaguchi-Shinozaki, 2007).

In response to water deficit, a key enzyme in the biosynthesis of ABA, *9-cis-epoxycarotenoid dioxygenase3* (*NCED3*), is upregulated (Harb *et al.*, 2010). It has already been shown that the overexpression of this gene confers a higher degree of tolerance to plants, while the *nced* mutant is more sensitive to water deficit (Iuchi *et al.*, 2001). This shows the importance of ABA to drought stress response, but what exactly happens at the cellular level to elicit this reaction is still not known, and is an important area of study.

1.4.1. ABA mediates the expression of regulatory genes

ABA induces the expression of a number of drought-responsive genes such as response to *response to dessication22* (*RD22*; Yamaguchi-Shinozaki *et al.*, 1992), *RD29A* and *RD29B* (Yamaguchi-Shinozaki *et al.*, 1992; Yamaguchi-Shinozaki and Shinozaki, 1993), *LEA* genes (Galau *et al.*, 1986; Skriver and Mundy, 1990) and dehydrins (Close *et al.*, 1989). The induction of these genes was mediated by the binding of transcription factors to *cis*-acting elements in the promoter region of these genes. The DNA-binding elements

of ABA-responsive genes are the ABA-responsive elements (AREs; Mundy *et al.*, 1990), to which ARE-binding proteins (AREBs) or ARE-binding factors (ABFs) bind. Alternatively, ABA-responsive genes may also contain MYB or MYC recognition sites (Abe *et al.*, 1997) to which MYB or MYC proteins bind.

1.4.2. ABFs are important regulators of drought response

A number of bZIPs (basic leucine zippers) are involved in a plant's response to drought (Uno *et al.*, 2000), and one family of bZIPs are the ABFs (Choi *et al.*, 2000). *ABF1*, *ABF2*/*ABRE1*, *ABF3* and *ABF4*/*ABRE2* were found to be induced by abiotic stresses like drought, old and high salinity (Choi *et al.*, 2000), and all four of these ABFs were induced by drought (Yoshida *et al.*, 2015). Overexpression of *ABF3* and *ABF4* in Arabidopsis conferred drought tolerance in those plants (Kang *et al.*, 2002), while overexpression of a constitutively active form of *ABF2* also led to enhanced drought tolerance in the mutant plants compared to wild-type plants (Fujita *et al.*, 2005). On the other hand, the *abre1abre2abf3* triple knockout mutant exhibited increased drought sensitivity (Yoshida *et al.*, 2010), while the quadruple knockout *abre1abre2abf3abf1* displayed an even greater sensitivity to drought (Yoshida *et al.*, 2015).

1.5. Other regulatory sequences that modulate gene expression during the stress response

In addition to ABA-responsive genes, there are a number of genes that are drought-responsive but ABA-independent, such as *RD19*, *RD21* and *RD28* (Yamaguchi-Shinozaki *et al.*, 1992). These genes are activated by the binding of transcription factors to dehydration-responsive elements (DREs; Yamaguchi-Shinozaki and Shinozaki, 1994), known as DRE-binding proteins (DREBs). Thus, the induction of drought-responsive genes is mediated either in an ABA-dependent or an ABA-independent pathway. Many studies have focussed on loss-of-function or gain-of-function mutants of transcription factors that bind to *cis*-acting elements in the promoter region of drought-responsive genes.

1.5.1. Evidence of the importance of DREBs and other ERFs during drought

DREBs are a class of ethylene response factors (ERFs) that bind the DRE element present in the promoters of mostly stress-responsive genes, activating transcription of these genes (Liu *et al.*, 1998). In Arabidopsis, *DREB2A* was found to be induced by dehydration, while *DREB1A* was induced only by low temperature. However, only overexpression of *DREB1A* conferred drought and cold tolerance to the plants, while overexpression of *DREB2A* in Arabidopsis did not (Liu *et al.*, 1998). This was found to be due to the presence of a negative regulatory domain in the protein. Removal of this domain rendered the protein constitutively active, and overexpression of this altered protein conferred drought tolerance to the plants (Sakuma *et al.*, 2006). *DREB1D* was also found to be induced by drought and overexpressors of this gene in Arabidopsis were also drought tolerant (Haake *et al.*, 2002).

Overexpressors of other ERFs, which are induced by drought, have also been shown to affect the drought phenotype of these plants compared to the wild-type. In Arabidopsis, *ERF053* was shown to be induced by drought (Cheng *et al.*, 2012). However, overexpression of this gene did not confer drought tolerance to the plants. This was due to the presence of a RING-finger E3 ligase called *RGLG2*, and its homologue, *RGLG1*, that target *ERF053* for protein degradation. Overexpression of *ERF053* in an *rglg1/rglg2* double knockout allowed for accumulation of *ERF053*, and subsequently, a drought-tolerant phenotype was observed. Moreover, overexpression of *ERF053* in the double knockout background conferred those plants with greater drought tolerance compared to just the double knockout with wild-type levels of *ERF053* (Cheng *et al.*, 2012).

Similarly, the ERF gene *RAP2.4* was also found to be induced by drought and overexpression of this gene lead to enhanced drought tolerance in these plants compared to the wild-type (Lin *et al.*, 2008). Overexpression of tomato *JERF1* in rice (Zhang *et al.*, 2010) and tomato *JERF3* in tobacco (Wu *et al.*, 2008) led to greater drought tolerance in these mutants. On the other hand, overexpression of another ERF, *ERF7*, caused increased sensitivity to drought due to decreased sensitivity to abscisic acid, and thus insufficient stomatal closure (Song *et al.*, 2005). This shows that differential expression of DREBs and ERFs is crucial during drought and these genes play important regulatory roles during the drought response.

1.5.2. Members of the NF-Y family are involved in the drought response

The Nuclear Factor-Y (NF-Y) complex is formed from a heterotrimeric complex consisting of the NF-YA, NF-YB and NF-YC subunits (Forsburg and Guarente, 1989), and in plants there are a large number of genes coding for these subunits (Siefers *et al.*, 2009). Members of the NF-YA family, namely *NF-YA2*, *NF-YA3*, *NF-YA5*, *NF-YA7* and *NF-YA10*, have been shown to be differentially expressed by a range of abiotic stress, such as low nitrogen, drought, cold, heat, high glucose and abscisic acid (Levya-González *et al.*, 2012). Overexpression of *NF-YA5* (Li *et al.*, 2008), *NF-YA2*, *NF-YA3*, *NF-YA7* and *NF-YA10* (Levya-González *et al.*, 2012) in Arabidopsis showed increased drought tolerance in these plants while the knockout, *nf-ya5*, showed increased sensitivity to drought (Li *et al.*, 2008). Increased levels of NF-YAs resulted in repression of cell growth and elongation, leading to growth arrest, an acclimation response (Levya-González *et al.*, 2012).

Nelson *et al.* (2007) found that Arabidopsis *NF-YB1* was also responsive to drought and plants overexpressing this gene were found to be more drought tolerant than wild-type plants after severe drought. It was also seen that these plants were able to maintain higher water potential and higher photosynthetic rate under drought compared to the wild-type. Maize plant overexpressing *NF-YB2* (a homologue of Arabidopsis *NF-YB1*) were also drought tolerant in field conditions and had higher chlorophyll content, higher stomatal conductance and photosynthetic rate and lower average leaf temperature (Nelson *et al.*, 2007). These plants were also found to be better yielding under drought conditions compared to wild-type plants, with the highest yielding lines producing 50% more yield than wild-type plants under drought conditions. Drought avoidance involves reduced stomatal conductance, and subsequently reduced carbon assimilation rate, to avoid water loss and maintain normal water potential during drought (Chaves *et al.*, 2003). This has the effect of reduced yield due to lower carbon assimilation rate (Chaves and Pinheiro, 2011). However, overexpressors of *NF-YB2* were able to maintain higher stomatal conductance and photosynthetic rate during drought than the wild-type plants, and thus have better yield than these plants when subjected to drought (Nelson *et al.*, 2007).

Thus, a number of members of the NF-Y family of genes are differentially expressed during drought stress and, thus, important transcriptional regulators of drought response in plants.

1.5.3. WRKYs are also important genes involved during drought

A number of WRKYs are also differentially expressed during drought and other abiotic stresses (Seki *et al.*, 2002; Matsui *et al.*, 2008). A gain-of-function T-DNA insertional mutant of the gene *WRKY57* was found to be tolerant to drought stress (Jiang *et al.*, 2012), and a knockout of *WRKY63* was drought-sensitive drought due to reduced stomatal closure during drought. Also, the overexpression of the drought-inducible *GsWRKY20* from wild soybean in Arabidopsis led to enhanced drought tolerance. The gene *GsWRKY20* is homologous to Arabidopsis *WRKY63*. The drought-tolerant phenotype was due to increased sensitivity to ABA resulting in adequate stomatal closure and a decreased stomatal density in these mutants.

1.6. Signal transduction during the drought response

1.6.1. Members of the MAPK pathway have been identified

A number of genes have been identified in plants that are homologous to members of the MAP kinase signalling pathway characterised in other eukaryotes (Bartels and Sunkar, 2005). Some of these genes have also been found to be differentially expressed in response to dehydration or osmotic stress (which is a consequence of water deficit). MAPK and its associated molecules, MAPKK and MAPKKK, are well-known signalling molecules involved in the activation of cellular responses to environmental signals in eukaryotes. Homologues of these proteins have been identified in plants, mostly through homology studies. The *Arabidopsis* MAP kinase protein, MPK3, has been found to be induced in response to drought (Mizoguchi *et al.*, 1996); the proteins MPK4 and MPK6 were also found to be activated post-translationally in response to osmotic stress (Ichimura *et al.*, 2000). An *Arabidopsis* MAPKKK, *AtMEKK1*, has been found to be upregulated in response to water deficit (Mizoguchi *et al.*, 1996). Another kinase that is overexpressed during drought is OXI1 (Oxidative Signal-Induced1), which is believed to act upstream of MPK3 and MPK6 (Rentel *et al.*, 2004).

1.6.2. Other signal transduction-related molecules are differentially expressed in drought

Another family of protein kinases has also been found to be upregulated in response to drought conditions. SNF1 (Sucrose-Nonfermenting1)-related protein kinases (SnRKs; also called Open Stomata1 [OST1]), have been implicated in carbon metabolism, as well as stress signalling and may involve ABA-dependent and ABA-independent pathways of activating downstream stress-related target genes (Coello *et al.*, 2011). The Receptor-like Protein Kinase1 (RPK1) also appears to play an important role in the early stages of ABA signalling, although its exact mechanism and substrates are not yet known (Osakabe *et al.*, 2005; 2010).

Protein phosphatases, along with protein kinases, are also involved in transducing external stimuli into a genetic response. There are two types of phosphatases known – protein (serine/threonine) phosphatases and protein tyrosine phosphatases (de Nadal *et al.*, 2002). One sub-group of protein phosphatases, PP2C, has been found to mediate an ABA-dependent signal transduction, through ROS and Ca^{2+}, and is a negative regulator of ABA signalling (Murata *et al.*, 2001; Mustilli *et al.*, 2002). Mutants having a disrupted form of PP2C gene, *abi1* (*ABA-insenstive1*), have been found to be more drought-tolerant (Gosti *et al.*, 1999).

Despite the characterisation of many proteins involved in signalling the stress response, there are still gaps in our knowledge of the signalling process which have yet to be filled.

1.7. Stomatal closure in response to stress is mediated by ABA

ABA regulates the closing of stomata in response to water deficit and this process is mediated by a number of enzymes and signalling molecules that coordinate in a signalling cascade (Mishra *et al.*, 2006; Zhang *et al.*, 2009). The α subunit of a heterotrimeric G-protein (GPA1) was found to be upregulated under mild drought conditions. Also, Phospholipase Dα1 (PLDα1) was found to be overexpressed under the same conditions (Harb *et al.*, 2010). These findings corroborate a proposed mechanism of ABA-induced stomatal closure and inhibition of stomatal opening in response to drought (Mishra *et al.*, 2006). PLDα1 catalyses the hydrolysis of membrane phospholipids to form phosphatidic acid (PA), which acts as a signalling molecule. PA binds PP2C, and

consequently inhibits the inhibiting effect of the phosphatase, allowing stomatal closure. PA has also been found to act upstream of GTP-bound GPA1, which signals the inhibition of stomatal opening (Mishra *et al.*, 2006). PA appears to also interact and activate NADPH oxidase, which subsequently produces ROS (Zhang *et al.*, 2009). ROS are known to induce Ca^{2+} transients by activation of Ca^{2+} channels, which has further implications in signal transduction (Mori and Schroeder, 2004).

All the components of the pathway involved in stomatal response have not yet been characterised and it is also believed that SnRK2 plays a role in the ABA-mediated stomatal response and gene regulation (Mustilli *et al.*, 2002; Coello *et al.*, 2011). The binding of ABA to PYR/PYL/RCAR proteins, which act as ABA receptors (Ma *et al.*, 2009; Park *et al.*, 2009), causes the interaction between the ABA receptors and PP2C. This interaction prevents the dephosphorylation and inactivation of SnRK2, and hence it is free to activate downstream target genes through ABFs (Coello *et al.*, 2011).

1.8. The role of osmolytes and protective proteins in drought response

It has been observed that the concentration of osmolytes increase in response to environmental stresses and it has also been found that these osmolytes are responsible for conferring tolerance in known stress-tolerant plant species (Zhang *et al.*, 1999; Serraj and Sinclair, 2002). Not only do these osmolytes have osmoprotectant properties, by regulating osmotic potential during drought, but they also appear to have other protective properties, such as ROS scavenging and detoxification, and stabilisation of cellular components. Proteins, such as LEA proteins, proteases, redoxins, aquaporins and heat shock proteins, which have a similar protective role, have also been found to be induced during water deficit (Bartels and Sunkar, 2005).

1.8.1. Sugars and sugar alcohols

A number of sugars, such as sucrose, raffinose family of oligosaccharides (RFOs), trehalose and fructan, show increased levels during the drought response (Taji *et al.*, 2002; Ogbaga *et al.*, 2016). They are believed to elicit a role in maintaining the osmotic potential of the cell during low water potential, by stabilising proteins and membranes through hydrogen bonding (Peters *et al.*, 2007), and RFOs have also been demonstrated to have ROS-scavenging properties in *Arabidopsis* (Nishizawa *et al.*, 2008). Elevated

levels of sucrose and RFOs have been found in the resurrection plant *Xerophyta viscosa* (Peters *et al.*, 2007). Levels of sucrose are also increased in another resurrection plant *Craterostigma plantagineum* (Bianchi *et al.*, 1991). *C. plantagineum* accumulates a high concentration of 2-octulose, which is converted to sucrose under conditions of dehydration. Trehalose also has protective properties against drought, as observed when the trehalose phosphate synthase (TPS) gene and other genes involved in the biosynthesis of trehalose were induced in plants (Garg *et al.*, 2002; Avonce *et al.*, 2004). Increases in levels of fructan, through overexpression of fructan biosynthetic genes, has also resulted in an increase in stress tolerance in tobacco and sugar beet plants (Pilon-Smits *et al.*, 1995; Pilon-Smits *et al.*, 1999).

Overexpression of the mannitol-1-phosphate dehydrogenase gene in wheat, required for the biosynthesis of mannitol, induces an increased tolerance to drought stress. Mannitol appears to act through the scavenging of harmful hydroxyl radicals (Abebe *et al.*, 2003) and also through the protection of ROS-scavenging and detoxifying enzymes from the action of hydroxyl radicals (Shen *et al.*, 1997), while D-ononitol is believed to act similarly (Sheveleva *et al.*, 1997).

1.8.2. Proline

Proline may have many roles in ameliorating the effects of stress and inducing drought tolerance in plants, and accumulates in response to drought (Girousse *et al.*, 1996; An *et al.*, 2013). It is believed to aid in osmotic adjustment, help in the stabilisation of the membranes (Rhodes *et al.*, 1986; Verbruggen and Hermans, 2008) and scavenge ROS (Cuin and Shabala, 2007). Overexpression of the proline biosynthetic gene, Pyrroline-5-carboxylate synthase, in tobacco and rice has been shown to increase in tolerance to osmotic stress (Kavi Kishor *et al.*, 1995; Zhu *et al.*, 1998).

1.8.3. LEA proteins

Exogenously added ABA has been found to induce the expression of LEA proteins in plants and is associated with a higher degree of desiccation tolerance in these plants (Xu *et al.*, 1996; Porcel *et al.*, 2005). LEA proteins have been found to act as protective molecules. Group 1 LEA proteins may be involved in binding water during water deficit,

groups 2 and 4 may be responsible for stabilising membranes and proteins, and groups 3 and 5 may sequester ions that accumulate during drought (Bartels and Sunkar, 2005).

1.8.4. Heat shock proteins

Heat shock proteins (HSPs) are induced by drought and correspond with an increase in tolerance (Wang *et al.*, 2004; Aneja *et al.*, 2015). Overexpression of HSP17.6A in *Arabidopsis* results in a gain of tolerance in these plants during water deficit, possibly through protein stabilisation (Sun *et al.*, 2001).

1.8.5. Proteases

Proteases have been found to be induced in response to drought stress(Simova-Stoilova *et al.*, 2006). This could be to ensure the turnover of damaged proteins to enable the synthesis of new proteins needed in the drought response. Cysteine proteases are induced – ERD1 (Nakashima *et al.*, 1997) and DegP2 (Hausühl *et al.*, 2001) are two chloroplast-related proteases that appear to be upregulated during drought.

1.8.6. Detoxifying proteins

Stress inducible-aldehyde dehydrogenase (Sunkar *et al.*, 2003) and aldehyde reductase (Oberschall *et al.*, 2000) have been identified in *Arabidopsis* that are responsive to drought. These act to remove toxic aldehydes, produced as a result of lipid peroxidation and mediated by free radicals, to form carboxylic acids and alcohols, respectively. Peroxiredoxins are other enzymes that are capable of detoxifying toxic peroxides to alcohols, thus preventing harmful ROS-mediated cellular damage (Seki *et al.*, 2002; Islam *et al.*, 2015). Thioredoxins have been studied in other organisms to be important detoxifying-enzymes, however, not much is known about them in plants (Rey *et al.*, 1998; Cha *et al.*, 2014).

Although some studies have been done about the effects of osmolytes and protective proteins during the drought response, and the genes that control them, there is still much more to be elucidated about the genes that initiate the osmotic responses and the pathways that lead to their osmotic effects.

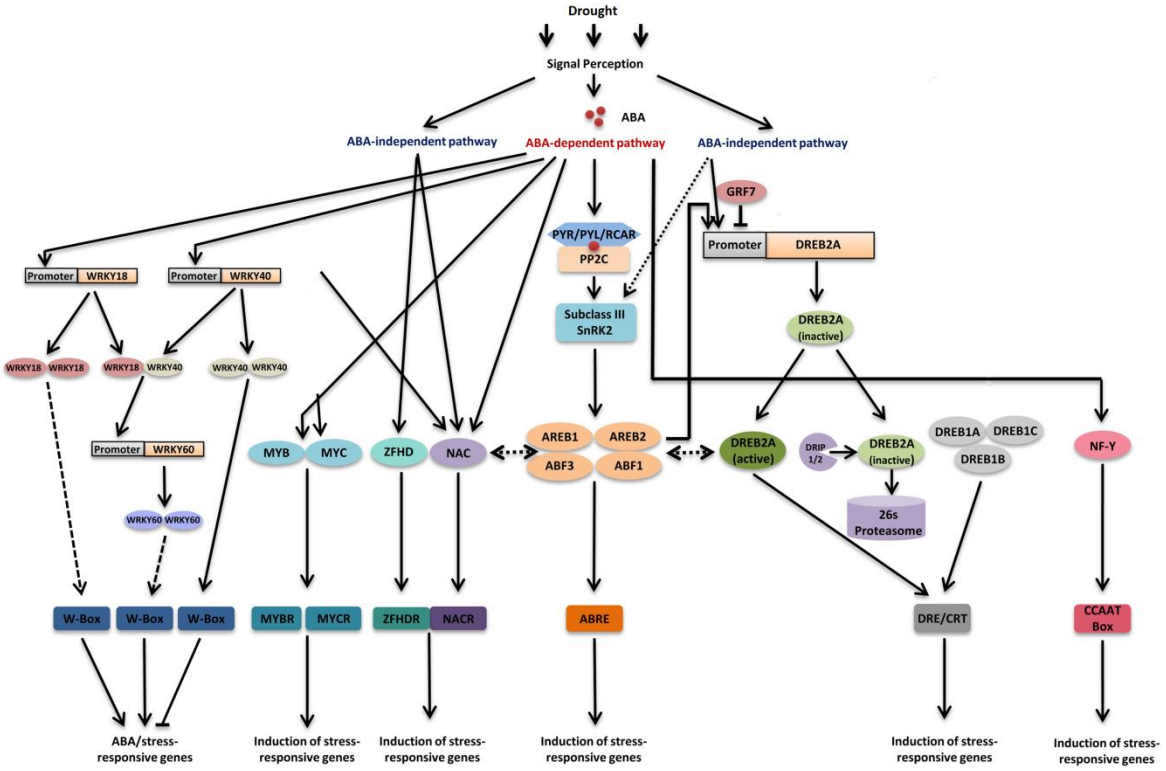


Figure 1.1. The components of the drought response in plants (adapted from Singh and Laxmi, 2015) from perception of the drought through to ABA-dependent and ABA-independent induction of transcription factors that regulate the expression of stress responsive genes.

1.9. Large-scale transcriptomics has been used to study drought transcriptomes

It is clear that a large number of regulatory and functional genes are differentially expressed during drought, and single gene studies do not allow for a holistic understanding of gene networks activated or repressed during drought. On the other hand, microarrays can be used to conduct genome-wide expression analyses (Eisen and Brown, 1999), and this can be used to analyse the drought transcriptome, as has been performed in a number of studies.

Among some of the earliest work on large-scale transcriptomics studies of drought-stressed plants was the work of Seki *et al.* (2001). In this study, drought stress was imposed on plant rosettes, by dehydration on Whatman 3MM paper, under growth conditions. RNA was extracted from these stressed rosettes and hybridised on a microarray spotted with the full-length cDNA for 1300 genes. A number of known drought stress-responsive genes were differentially expressed during drought, such as *RD29A*,

COR15A (*cold-regulated15A*), *KIN1* (*cold-induced1*), *KIN2*, *ERD10* (*early response to dehydration10*), *RD17*, *RD20*, *ERD7*, *ERD4*, *ERD14*, *ERD3*, *RD19A*, *RD22* and *P5CS* (*pyrroline-5-carboxylate synthase*), involved in proline biosynthesis (Yoshiba *et al.*, 1995). In addition to these 14 genes, another 30 genes, previously unknown to be drought-responsive, were also differentially expressed during drought. Thus, this study showed that the use of microarrays could be used to identify new genes involved during drought.

Seki *et al.* (2002) performed a time-series microarray analysis on rosettes subjected to dehydration, sampling at 1, 2, 5, 10 and 24 hours after the initiation of the stress. A full-length cDNA microarray was used to measure the expression levels of approximately 7000 genes. In this study, it was seen that 277 genes were induced during drought, while 79 genes were down-regulated. 40 of the 277 genes that were induced were transcription factors (TFs), showing that TFs play an important role in the drought response. In addition to TFs, genes coding for the biosynthesis of osmoprotectants such as proline and sugars, for LEA genes, an aquaporin and other membrane proteins were also seen. Genes involved in the biosynthesis of hormones such as ABA, ethylene, jasmonic acid and auxin, were also up-regulated during drought and genes involved in cellular signalling and transduction, such as those coding for Ca^{2+}-binding and calmodulin-binding proteins, protein kinases and phosphatases, were also induced. In total, genes belonging to 40 functional categories, and a large group of unknown genes, were induced during drought. This shows that the drought response is a complicated one involving many different groups of functional and regulatory genes. In addition to the induction of a large number of genes during drought, 79 genes were also down-regulated. The largest group of genes that were down-regulated were those involved in photosynthesis. 19 functional groups of genes were repressed during drought, and again a large number of down-regulated genes were unknown.

The type of drought experiment that was done in this study is unrepresentative of a normal drought that occurs in the field. It involved dehydration of plant rosettes for a few hours on the bench, while drought in the field lasts for days (Passioura, 2002). However, the use of a time-series analysis enabled the identification of genes that would not have been picked up from a single time-point study. Cold- and salinity-stressed plants were also analysed in a similar way and it was seen that more than half of the genes that were differentially expressed during drought were common to either cold or salinity or both. This shows that a large part of the stress response is common to different abiotic stresses, in addition to the many stress-specific genes that are induced or repressed (Seki *et al.*, 2002).

Another study by Kreps *et al.* (2002) was done in which plants were subjected to osmotic stress, using mannitol, when the photoperiod began, and leaf and root tissue were harvested at 3 hours and 27 hours after the stress. Cold and salinity genes were also analysed. A large number of genes common to all three stresses were differentially expressed in both the shoots and the roots 3 hours after the stresses were applied. This involved many transcription factors and signalling-related genes. 27 hours after the stresses, the number of genes differentially expressed in roots and shoots decreased drastically. Similar to Seki *et al.* (2002), this study also showed that it is important to study the effect of stress over a period of time rather than at a single point, as the transcriptome varies as the stress progresses.

In the drought analysis by Kilian *et al.* (2007), 24,000 genes were analysed in 18-day old plants. The plants were exposed to a stream of air for 15 minutes till 10% of the fresh weight was lost. Differentially expressed genes were also analysed in plants subjected to cold and UV-B light stress. The plants were returned to growth conditions after stressing them and the roots and shoots were harvested separately. Plant material was harvested 15, 30, 60 minutes, 3, 6, 12 and 24 hours after the stress. At the beginning of the analysis, many of the differentially expressed genes (such as transcription factors and signalling genes) were those that were shared between two or all of the stresses. Over time, the number of shared genes decreased, while stress-specific genes continued to constitute the majority of those that were differentially expressed in both roots and shoots. From this study and from Kreps *et al.* (2002), it is clear that there is an initial response that is common to abiotic stresses which then develops into a stress-specific response over time, and the analysis of the transcriptome at a single time-point after drought does not provide the full picture of the response.

The above studies used a variety of methods for imposing drought, such as quick dehydration of rosettes or seedlings, or using mannitol. These methods do not portray drought in the field accurately as they are not soil-based. Harb *et al.* (2010) performed a soil-based drought experiment to analyse the transcriptome under stress. They also compared a progressive drought experiment with a moderate drought experiment, where the relative soil water content is maintained at 30%. They believed that progressive drought is too severe and that moderate drought is more realistic to field conditions, and so performed a microarray analysis to compare the two types of drought stresses. Moderate drought was induced by withdrawing water till 2 g g^{-1} dry soil (30% field capacity) was reached, and then supplementing it to maintain this water content for the remainder of the experiment, while plants subjected to progressive drought were allowed

to wilt. 7648 genes were differentially expressed in wilting plants that were subjected to a progressive drought stress (pDr). In contrast, 2039 genes were differentially expressed on day 1 of the moderate drought (mDr day01), and 728 genes on day 10 (mDr day10).

Despite the difference in the number of genes differentially expressed between the two types of drought, they found that some of the same GO terms associated with drought were differentially expressed in both pDr and mDr day 01 (Harb *et al.*, 2010). These terms were response to water deprivation, response to ABA stimulus, osmotic, cold, and oxidative stresses, and DNA packaging, ribosome biogenesis and protein folding. Also, most of the differentially expressed genes in mDr day 01 were common to pDr, and were expressed in the same direction i.e. induced or repressed. The biggest difference in the expression levels of genes between the two types of stress was seen in those genes coding for cell wall modification enzymes – these were up-regulated in mDr day 01, but down-regulated in pDr – and in photosynthetic genes, which were down-regulated in pDr but showed no difference in expression in mDr day 01. These indicate an acclimation response in the plants subjected to moderate stress, as opposed to those subjected to progressive drought. However, as drought can be unpredictable, moderate drought represents an unlikely situation in the field.

As mentioned in this section, a number of large-scale transcriptome studies have been performed. These analyses have identified many more targets for gene manipulation to produce drought tolerant plants. However, these studies have used artificial methods for inducing drought or have not measured the changing dynamic in the drought transcriptome, using a time-series analysis. A drought experiment that encompasses both of these factors could provide valuable information about the initiation and progress of the drought response under conditions that resemble drought in the field.

1.10. Use of statistical algorithms to reverse-engineer gene regulatory networks

The huge volumes of data arising from genomic technologies has led to the need to study organisms at the ‘systems’ level, instead of at the single molecule level (Gutierrez *et al.*, 2005). Systems Biology aims to understand organisms from a whole-organism perspective, by integrating gene, protein, metabolite and other biological data to reverse engineer a network describing the interactions of each (Cho *et al.*, 2007). Gene regulatory networks (GRNs) indicate the interaction between different genes and provide a wider view of gene networks that can potentially reveal novel connections. It also

provides a dynamic view of the interactions, indicating activation and repression (Yuan *et al.*, 2008).

Statistical algorithms can be used to reverse-engineer GRNs based on transcriptomics data and the most common types of mathematical models used are ordinary differential equations, Boolean networks and Bayesian networks (Bansal *et al.*, 2007; Cho *et al.*, 2007). Ordinary differential equations represent the change in the level of a transcript as a function of the change in the levels of all other transcripts and also due to the effect of an external perturbation. Boolean networks represent the expression levels of genes as either '0' or '1' i.e. unexpressed state or expressed state. Both of these methods can only be applied to steady-state data and cannot be used on dynamic (time-series) data. Bayesian networks can either be static or dynamic; dynamic Bayesian networks are used to handle time-series data by applying Bayesian statistics to reverse-engineer the causal relationship between two 'nodes' or genes. Examples of networks that have reverse-engineered using each of these algorithms are given below (Gardner and Faith, 2005; Bansal *et al.*, 2007; Cho *et al.*, 2007).

The three-loop feedback network in the circadian clock in Arabidopsis was mathematically modelled using differential equations (Locke *et al.*, 2005). It predicted the presence of a gene, *Y*, in the network and also the dynamic of its expression during the photoperiod. Using this theoretical information, *Y* was soon identified experimentally to be *GIGANTEA*, an important gene involved in circadian rhythms in plants (Locke *et al.*, 2005; 2006).

Espinosa-Soto *et al.* (2004) used Boolean networks to model the dynamics of cell type determination during flower development, using the ABC model and available experimental results. Novel predictions were identified and two of these were independently verified experimentally by Schmid *et al.* (2003) and Gómez-Mena *et al.* (2005). A Boolean network was also used by Li *et al.* (2006) to model ABA induction of stomatal closure, by integrating data from various experimental studies. They were able to model the effect of perturbing parts of the network on ABA-mediated stomatal closure, and identify potentially novel conditions that regulate this.

A Dynamic Bayesian method called Variational Bayesian State Space Modelling (VBSSM; Beal *et al.*, 2005) was used by Breeze *et al.* (2011) to model whole transcriptome data from a time-series senescence experiment. The model was able to correctly predict the positive effect that the gene *ANAC092* would have on the expression

of known target genes. It also identified *STZ* as a highly connected gene in the network that regulates the expression of several genes during senescence, providing a candidate gene for experimental analysis.

Windram *et al.* (2012) used a Nonlinear Dynamic System called Causal Structure Identification (CSI; Penfold *et al.*, 2012) to model time-series transcriptional data from Arabidopsis plants subjected to *Botrytis cinerea* infection. This study identified *TGA3* as an important gene potentially involved in the response to Botrytis infection, which was experimentally verified, as knockouts of the gene were more susceptible to infection that the wild-type plants. Hickman *et al.* (2013) modelled the time-series data from Breeze *et al.* (2011) and Windram *et al.* (2012) using CSI, and combining this with Yeast-1-Hybrid data, identified genes such as *MYB* genes that act upstream of the transcription factors, *ANAC019*, *ANAC055* and *ANAC072*. These connections were experimentally verified using microarray analyses of mutants of the *MYB* genes.

1.11. Objectives of the project

This thesis is based on the work of Bechtold *et al.* (2016) in which a progressive drought experiment was performed with wild-type Arabidopsis Col-0, and time-series transcriptomics analyses were conducted using microarrays. In that work, differentially expressed genes were clustered in groups using hierarchical clustering, and these groups of genes were modelled into networks using the Dynamic Bayesian model VBSSM, just as in Breeze *et al.* (2011), and a network was generated around the flowering time gene, *AGAMOUS-LIKE22* (*AGL22*) / *SHORT VEGETATIVE PHASE* (*SVP*).

Using the same list of differentially expressed obtained in Bechtold *et al.* (2016), the genes were clustered using a temporal clustering algorithm and clusters were selected to model gene networks during drought using VBSSM. ‘Hub’ genes in the networks were selected for further analysis by phenotyping loss- and gain-of-function mutants of these genes. Finally, some of the networks were selected to verify gene connections between hub genes and potential downstream genes.

The main objectives of this project were:

1. Generation of GRNs using the time-series transcriptomics data from Bechtold *et al.* (2016) and identification of hub genes that represent potentially important genes involved during the drought response.
2. Phenotyping loss- and gain-of-function mutants of these hub genes under drought conditions to verify a potential role in drought stress signalling.
3. Evaluation of the GRNs generated using VBSSM by testing the potential regulation of network genes by the hub genes.

Chapter 2

Materials and Methods

2.1. Bioinformatics techniques

2.1.1. Temporal Clustering of differentially expressed genes

A list of differentially expressed genes with a Gaussian process two-sample test (GP2S; Stegle *et al.*, 2010) score ≥ 5 (as described in Chapter 3) was used for further analysis. The genes were grouped into clusters using the temporal clustering algorithm Temporal Clustering by Affinity Propagation (TCAP; http://www.wsbc.warwick.ac.uk/stevenkiddle/tcap .html; Kiddle *et al.*, 2009) that was run in MATLAB® (MathWorks®, USA). TCAP clusters genes with similar expression profiles including early-, middle- and late-responding genes as well as those with inverse profiles.

2.1.2. Enrichment of Biological Process terms

The clusters were enriched for terms related to Biological Processes using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (http://david.abcc.ncifcrf.gov/; Huang *et al.*, 2009) to identify groups of related functional genes clustered by TCAP.

2.1.3. Modelling drought-responsive gene networks by Variational Bayesian State-Space Modelling

Gene clusters from TCAP were selected for further analysis, based on the results from DAVID. Individual clusters were used to model a number of gene regulatory networks (GRNs) using Variational Bayesian State-Space Modelling (VBSSM; Beal *et al.*, 2005). A graphical user interface version of VBSSM (available from http://www2.warwick.ac.uk/fac/sci/systemsbiology/research/software/) was run in MATLAB® using the following settings: Number of seeds = 15; Maximum dimension of hidden states = 20; Maximum number of iterations = 2000. The gene networks predicted by the models were viewed using Cytoscape 2.8.2.

2.2. Molecular Biology techniques

2.2.1. Plant material and growth conditions

Plants were grown in compost (Levington F2+S, The Scotts Company, Ipswich, UK) and maintained in a controlled environment – 8/16 hour light/dark cycle at 23 °C and 60% relative humidity, under a photosynthetic photon flux density of ~200 µmol m^{-2} s^{-1}.

2.2.2. Primer design

All primers used in this thesis were designed using the website Primer Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/; Ye *et al.*, 2012) which is available from the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/). Information about all the primers used in this thesis is given in Appendix C.

2.2.3. Identification of knockout mutants from T-DNA insertional lines

2.2.3.1. T-DNA insertional lines were obtained from NASC

Multiple lines of putative T-DNA insertional mutants for each gene of interest were obtained from the Nottingham Arabidopsis Stock Centre (NASC, http://arabidopsis.info/). SALK (Alonso *et al.*, 2003), SAIL (Sessions *et al.*, 2002) and GABI-Kat (Rosso *et al.*, 2003) lines in a Col-0 background were used in this thesis, and a list of all T-DNA insertional mutants screened are given in Appendix B.

2.2.3.2. Genomic DNA extraction

Genomic DNA was extracted from plants using the protocol described by Edwards *et al.* (1991), with some modifications. A single leaf was taken from each plant and was ground using micropestles, along with 500 µl of DNA extraction buffer (200 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 mm EDTA, 0.5% SDS). The samples were then centrifuged at 12,000 g for 10 minutes, and the resulting supernatant was mixed 1:1 with isopropanol. This was then centrifuged at 12,000 g for 10 minutes, the pellet was washed with 70% ethanol and the dried pellet was resuspended in 50 µl of Reverse Osmosis (RO) water.

2.2.3.3. PCR to identify T-DNA insertional mutants

A Polymerase Chain Reaction (PCR) was performed to verify the presence of the T-DNA insertion within the gene of interest. *Taq* DNA polymerase (recombinant) (#EP0402, ThermoFisher Scientific, Loughborough, UK) was used for the PCR to a final concentration of 2.5 U/µl; the accompanying 10X *Taq* Buffer (100 mM Tris-HCl [pH 8.8 at 25 °C], 500 mM KCl, 0.8% [v/v] Nonidet P40) was used at a final concentration of 1X and 2.5 mM of $MgCl_2$ was also added. A mix of deoxyribonucleotide triphosphates (dNTPs; #R0191, ThermoFisher Scientific) were added to a final concentration of 200 µM and 100-200 ng of genomic DNA was used per reaction. The forward and reverse primers (sequence information in Appendix C) were added to a final concentration of 500 nM each and the volume was made up to 25 µl with RO water. Figure 2.1 shows the primers used to verify the presence of the T-DNA insertion and the zygosity of the plants.

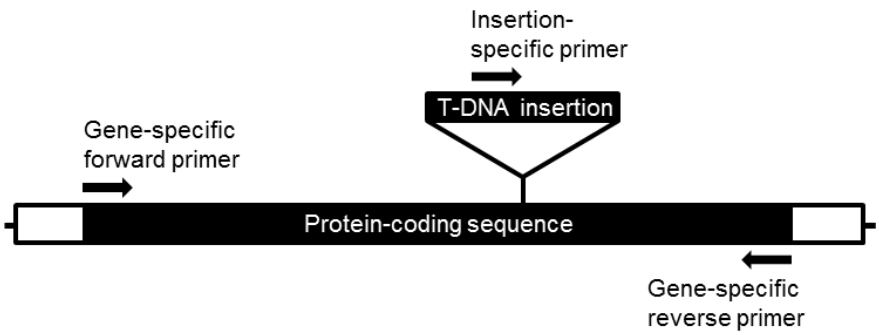


Figure 2.1. Strategy (indicated by the block arrows) used to screen T-DNA insertions

The conditions used for the PCR reactions were: an initial denaturation at 98 °C for 3 minutes; 35 cycles of denaturation at 98 °C for 30 seconds, primer annealing at 60 °C for 30 seconds and extension at 72 °C for 3 minutes; final extension at 72 °C for 10 minutes.

2.2.3.4. Agarose Gel Electrophoresis and Visualisation of PCR products

The PCR products were mixed with 6X DNA loading dye (10 mM Tris-HCl [pH 7.6], 0.03% Bromophenol Blue, 0.03% Xylene Cyanol FF, 60% Glycerol, 60 mM EDTA) in the ratio of 6:1, respectively. These were run on a 1% agarose gel in TBE buffer (10.8 g Tris, 5.5 g Boric acid, 4 ml of 500 mM EDTA [pH 8.0]) at 110V for 35 minutes, alongside the GeneRuler™ DNA Ladder Mix (#SM0331; ThermoFisher Scientific). The agarose gel was

stained with SafeView (#NBS-SV1, NBS Biologicals Ltd., Huntingdon, UK) in a 1:10,000 dilution and the DNA was visualised using the GeneGenius Bioimaging system (SYNGENE, SYNOPTICS Ltd., Cambridge, UK), as per the supplier's instructions.

2.2.3.5. Gel extraction of PCR product from agarose gel

PCR products were purified by extracting them from the agarose gel using the GeneJET Gel Extraction Kit (#K0691, ThermoFisher Scientific), as per the supplier's instructions, with the exception that 20 µl of RO water was used to elute the DNA from the column instead of Elution Buffer.

2.2.3.6. TA Cloning® of PCR product for sequencing

To identify the position of the T-DNA insertion within the gene of interest, the PCR product obtained with the insertion-specific primer was first cloned into the pCR®2.1 vector to allow for sequencing. TA Cloning® was performed using the TA Cloning® Kit (#K2020-20, ThermoFisher Scientific) according to the supplier's instructions, before being transformed into chemically competent One Shot® OmniMAX™ 2-T1^{R} *Escherichia coli* cells (#C8540-03, ThermoFisher Scientific).

2.2.3.7. Transformation of chemically competent *E. coli* cells

E. coli OmniMAX™ cells were transformed with the pCR®2.1 plasmid using the heat-shock method. 2 µl of the plasmid was gently mixed with a 50 µl aliquot of *E. coli* competent cells and was incubated on ice for 30 minutes. Heat-shock was performed at 42 °C for 30 seconds, and immediately incubated on ice for 5 minutes. 250 µl of Luria-Bertani (LB) medium (#L2408, Melford Laboratories Ltd., Ipswich, UK) was then added to the cells and this was shaken at 37 °C for 1 hour. 40 µl of 100 mM IPTG (#I6758, Sigma-Aldrich®, Dorset, UK) and 40 µl of 40 mg/ml X-Gal (#B4252, Sigma-Aldrich®) were spread on to LB Agar plates containing 50 µg/ml of Kanamycin (#K0126, Melford Laboratories Ltd.), to screen positive transformants by blue-white screening. 200 µl of the transformed cells were then plated till the plates were dry, and these were incubated overnight at 37 °C.

2.2.3.8. Colony PCR to verify positive transformants

Colony PCR was performed on white colonies to verify positive transformants. The PCR reaction was set-up as described in section 2.2.3.3 and the M13 forward and reverse primers were used for the reaction. Instead of using genomic DNA as the template, white colonies were picked with clean tips, streaked onto a fresh LB Agar plate and dipped into the PCR mixture. The annealing temperature used for the M13 forward and reverse primers was 60 °C. The PCR products were run on a 1% agarose TBE gel, as mentioned in section 2.2.3.4.

2.2.3.9. Extraction of plasmid DNA

10 ml of LB medium containing 50 µg/ml of Kanamycin was inoculated with a positive colony from the freshly streaked plate, and incubated overnight at 37 °C. The grown culture was spun down at 3000 g for 20 minutes to pellet the cells and plasmid DNA was extracted using the QIAprep® Spin Miniprep Kit (#27104, Qiagen, Manchester, UK), as per the recommended protocol.

2.2.3.10. Quantification of nucleic acids

The plasmid was quantified and the purity ascertained by spectrophotometry, using the NanoDrop™ ND-1000 Spectrophotometer (ThermoFisher Scientific) as per the manufacturer's instructions.

2.2.3.11. Sequencing to determine the position of the T-DNA insertion

The pCR®2.1 plasmid transformed with the PCR product obtained using the insertion-specific primer (as described in section 2.2.3.6) was sequenced between the M13 forward and reverse primers to identify the position of the T-DNA insertion in the gene of interest. 100 ng/µl of plasmid was sent to the sequencing company GATC Biotech (Konstanz, Germany) and the sequence was analysed using ContigExpress (Vector NTI Advance 10, ThermoFisher Scientific).

2.2.3.12. RNA extraction from plant leaf tissue

Total RNA from identified T-DNA insertional mutants was extracted using TRI Reagent® solution (#AM9738, ThermoFisher Scientific), as per the supplier's instructions with some changes. 100 mg of leaf tissue was harvested, immediately frozen in liquid nitrogen and subsequently finely ground using a mortar and pestle. 1 ml of TRI Reagent® was added to the tissue and incubated for 5 minutes at room temperature (RT). 200 µl of Chloroform was added and the solution was mixed well and again incubated for 5 minutes at RT. The samples were centrifuged at 12,000 g for 20 minutes at 4 °C. The aqueous phase was transferred to a new tube and mixed 1:1 with Isopropanol, before centrifuging at 12,000 g for 20 minutes at 4 °C to pellet the RNA. The RNA pellet was then washed with 1 ml of ice-cold 70% Ethanol at 12,000 g for 5 minutes at 4 °C and air-dried to remove excess Ethanol. The RNA was resuspended in 26 µl of RO water and DNase-treated.

2.2.3.13. DNase treatment of RNA

The extracted RNA was subjected to DNase treatment to ensure that the sample was DNA-free. Recombinant DNase I (rDNase I; #AM2235, ThermoFisher Scientific) was used to treat the RNA. To each sample, the 10X DNase I buffer was added to a final concentration of 1X along with 2 U of rDNase I, and the samples were incubated at 37 °C for 30 minutes. 2.4 mM of EDTA was then added and the rDNase I was heat inactivated at 75 °C for 5 minutes.

2.2.3.14. cDNA preparation from RNA

RNA was quantified, as mentioned in section 2.2.3.10, and 1000 ng of RNA was taken into a separate tube to convert to cDNA. 0.2 µg of Random Hexamer Primer (#S0142, ThermoFisher Scientific) was added to the RNA and made up to 12 µl. This was incubated at 65 °C for 10 minutes, after which it was immediately placed on ice. The RNA was reverse-transcribed using the RevertAid Reverse Transcriptase (#EP0442, ThermoFisher Scientific). The 5X buffer was added to a final concentration of 1X, dNTPs were added to give a final concentration of 1 mM and 200 U of the enzyme were added. The volume was made up to 20 µl using RO water and the tubes were incubated under the following conditions: 25 °C for 10 minutes, 42 °C for 60 minutes and 70 °C for 10 minutes.

2.2.3.15. Expression analysis using RT-PCR

The semi-quantitative Reverse Transcription-PCR (RT-PCR) was performed on the T-DNA insertional mutants to verify that the gene of interest was knocked out. The forward and reverse primers used were specific to the start and end of the protein-coding sequence, and thus the full-length sequence was analysed. The sequence information for each primer is given in Appendix C. The PCR was setup as mentioned in section 2.2.3.3, with the exception that 0.5 µl of undiluted cDNA and 0.5 µl of DreamTaq DNA Polymerase (#EP0702, ThermoFisher Scientific) were used. $MgCl_2$ was not added to the mix as it was already incorporated in the buffer. The PCR conditions used were the same as in section 2.2.3.3.

2.2.3.16. Expression analysis using qPCR

Quantitative PCR (qPCR) was used on T-DNA insertional mutants that appeared to have increased levels of the gene of interest when analysed by RT-PCR. The SensiFAST™ SYBR No-ROX Kit (#BIO-98002, Bioline Reagents Ltd., London, UK) was used to perform the qPCR. The 2X SensiFAST™ SYBR No-ROX Mix was diluted to 1X and the forward and reverse primers were added to a final concentration of 400 nM each. 1 µl of 1:5 diluted cDNA was added to each reaction and made up to 20 µl with RO water. The sequences of the primers used for qPCR are given in Appendix C. The following PCR conditions were used for the reactions: Initial denaturation and DNA polymerase activation at 95 °C for 2 minutes; 50 cycles of denaturation at 95 °C for 5 seconds, annealing at 60 °C for 10 seconds and extension at 72 °C for 5 seconds. A melting curve was done from 65 °C to 90 °C in 0.5 °C increments for 5 seconds, to verify the amplification of just a single product. All replicates are biological replicates and stated in the legends.

2.2.4. Production of lines overexpressing the gene of interest

To construct overexpression vectors for the genes of interest, the Gateway® Technology (Invitrogen™ from ThermoFisher Scientific) was used. This provides a quick and efficient method of cloning a gene of interest into a number of vectors for functional characterisation of the gene (Hartley *et al.*, 2000). It involves cloning the gene into the

TOPO® vector and subsequently the use of site-specific recombination to clone the gene into a compatible vector of choice using LR Clonase™ II.

2.2.4.1. Amplification of the gene of interest for Gateway® Cloning

The full-length protein-coding sequence of the genes of interest was amplified from wild-type Col-0 cDNA by PCR. The Phusion® Hot Start II High-Fidelity DNA polymerase (#F-5495S, ThermoFisher Scientific) was used as per the supplier's instructions for a 20 µl reaction, and using the recommended PCR conditions. The forward primer contained the CACC sequence at the 5' end to ensure directional cloning into the TOPO® vector. The sequences of the forward and reverse primers used in the PCR are given in Appendix C. 100-200 ng of cDNA was added to each reaction and an annealing temperature of 60 °C was used. The PCR products were run on a 1% agarose TBE gel and purified by gel extraction as mentioned in section 2.2.3.4.

2.2.4.2. TOPO® cloning of the amplified gene

The PCR product was ligated into the pENTR™/D-TOPO® vector (#K2400-20, ThermoFisher Scientific) as per the supplier's instructions. The vector was transformed by heat-shock into *E. coli* OmniMAX™ cells, plated on Kanamycin (50 µg/ml) LB Agar plates and incubated overnight at 37 °C. Positive transformants were identified by colony PCR with M13 forward and reverse primers. The positive colonies were inoculated in 10 ml cultures and grown overnight to extract the plasmid. A restriction digestion was performed to check the positive clones. The sequence and directionality of the genes in the TOPO® vector were confirmed by sequencing with the M13 forward and reverse primers.

2.2.4.3. Restriction Digestion of TOPO® clones

The TOPO® clones were further verified by restriction digestion of the plasmid. The restriction enzymes to be used were determined from the sequence of the plasmid with the gene insert and using the application PlasmaDNA (University of Helsinki) available from http://research.med.helsinki.fi/plasmadna/. The enzymes and corresponding buffers used for each of the constructs were obtained from ThermoFisher Scientific and are given in Table 2.1. To 1000 ng of the plasmid, 1 µl containing 10 U of each enzyme was added,

and the 10X buffer was added to a final concentration of 1X. The volume of the reaction was made up to 20 µl with RO water and incubated at 37 °C for 1 hour. For the double digests, the correct buffer and the amount of enzyme to be added were determined using the DoubleDigest Calculator (https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/thermo-scientific-restriction-modifying-enzymes/restriction-enzymes-thermo-scientific/double-digest-calculator-thermo-scientific.html, ThermoFisher Scientific). The products of the restriction digestion were run on a 1% agarose TBE gel.

Table 2.1 Table showing the restriction enzymes used to confirm the cloning of the gene of interest into the pENTR™ vector.

Gene construct	**Restriction enzymes used (with Catalogue number)**	**Enzyme Buffer used**
AGL22	*MluI* (#ER0561)	Buffer R
ANL2	*MluI*	Buffer R
BHLH038	*MluI* and *HindIII* (#ER0501)	Buffer R
POZ	*MluI* and *HindIII*	Buffer R
RAP2.12	*NotI* (#ER0591) and *PstI* (#ER0611)	Buffer O
UKTF	*NotI* and *PstI*	Buffer O
FD	*EcoRI* (#ER0271) and *PvuII* (#ER0631)	Buffer R

2.2.4.4. LR Cloning of the gene into a Destination vector

The gene of interest was cloned into the destination vector pEarleyGate (Earley *et al.*, 2006) by LR Cloning, using the Gateway® LR Clonase™ II Enzyme Mix (#11791020, ThermoFisher Scientific) as per the supplier's instructions. The binary vector pEarleyGate contains the Cauliflower Mosaic Virus (CaMV) 35S promoter and the termination sequences from the *Octopine Synthase* (*OCS*) Gene, for overexpression of the gene of interest. The vector contains the *Basta Resistance* (*BaR*) gene for *in planta* selection, and the *Kanamycin Resistance* (*KanR*) gene for selection in bacteria. The constructs were transformed by heat-shock into *E. coli* OmniMAX™ cells, plated on LB Agar plates containing 50 µg/ml of Kanamycin and incubated overnight at 37 °C. Positive transformants were identified by colony PCR with the 35S promoter-specific forward primer (Appendix C) and the gene-specific reverse primer. A restriction digestion of the

plasmid extracted from the positive transformants was also done to further verify the results of the colony PCR, and the products of the digestion were analysed on a 1% TBE gel.

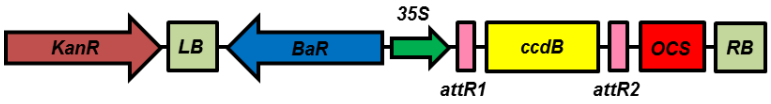


Figure 2.2. Vector map of the destination vector pEarleyGate. The pEarleyGate vector (Earley *et al.*, 2006) contains the CaMV 35S promoter and the OCS terminator, between which the gene of interest is inserted by recombination between the attR sites on pEarleyGate and the attL sites of the pENTR™ vector. This is contained between the right and left T-DNA border sequences, along with the *BaR* gene for *in planta* selection using Basta®. The vector also contains the *KanR* gene for selection in *E. coli* using Kanamycin.

2.2.4.5. Transformation of *Agrobacterium tumefaciens*

The pEarleyGate expression constructs were transformed into electro-competent *Agrobacterium tumefaciens*, strain GV3101/pMP90 (Koncz and Schell, 1986) by electroporation. 1 µl of the plasmid was gently mixed with a 50 µl aliquot of the cells and transferred to ice-cold 2 mm electroporation cuvettes (#732-1136, VWR International, Lutterworth, UK). Electroporation was carried out at 1800 volts using the EasyJecT Prima Electroporator (Equibio Ltd., Ashford, UK). 1 ml of ice-cold LB medium was added to the cells and the samples were shaken at 28 °C for 3 hours. 200 µl of the cells were then plated on LB Agar plates containing 50 µg/ml of Rifampicin (#13292-46-1, Melford Laboratories Ltd.), 25 µg/ml of Gentamycin (#1405-41-0, Melford Laboratories Ltd.) and 50 µg/ml of Kanamycin. The plates were incubated at 28 °C for 2 days and positive transformants were verified by colony PCR using the same primers used to screen *E. coli* colonies after LR cloning.

2.2.4.6. Transformation of Arabidopsis by Agrobacterium-mediated transformation

Positive colonies were grown overnight at 28 °C in 10 ml LB medium containing 50 µg/ml of Rifampicin, 25 µg/ml of Gentamycin and 50 µg/ml of Kanamycin. These were used to inoculate 500 ml cultures and wild-type Col-0 Arabidopsis plants were transformed with the overexpression vectors using the floral-dip method described by Clough and Bent (1998). The seeds (T_1) of the transformed plants were screened on soil by watering the seedlings regularly with 0.62 mM of glufosinate-ammonium (Kaspar; Bayer CropScience,

Cambridge, UK). Positive transformants were analysed by qPCR, as mentioned in section 2.2.3.16, to verify overexpression of the gene of interest. Homozygous lines were obtained by screening T_2 or T_3 seeds to confirm there is no segregation in the lines.

2.3. Plant Phenotyping techniques

2.3.1. Drought stress conditions

The drought experiment was performed as described by Bechtold *et al.* (2016). To perform the drought experiments, 7x7x8 cm pots (Desch Plantpak, Maldon, UK) were filled with equal weight of a 6:1 mixture of compost and Vermiculite (Verve; B&Q, Colchester, UK) and 2-3 seeds were sown per pot. Control pots were filled to the same weight, soaked in water for a few hours till saturation and weighed to estimate 100% relative soil water content (rSWC). The pots were allowed to dry out completely and the pots were weighed to estimate 0% rSWC. 6 weeks after sowing, half of the plants were maintained as normally watered control plants, while water was completely withdrawn from the remaining plants. The pots subjected to drought were weighed every day and the rSWC was measured using the formula:

$$\frac{\text{Pot weight} - \text{Empty pot weight} - \text{Dry soil weight}}{\text{Saturated pot weight} - \text{Dry soil weight}} \times 100$$

The drought was continued until the rSWC reached 20%. Plants were either sampled for the experiments described below, or to estimate biomass (rosette, stalks and siliques, and seed). Plants that were used to harvest leaf material were discarded immediately. If the plants were used to measure plant biomass, then the plants were re-watered and a normal watering regime was maintained until the first silique matured, at which point the plants were bagged and allowed to dry out. The dried rosette, reproductive structures (stalks and siliques) and seeds were separated and weighed to estimate vegetative, reproductive and total biomass.

2.3.2. Calculation of rosette area

Digital images of the plants were taken using a Panasonic DMC-GF6K camera and the images were analysed using the image processing software ImageJ

(https://imagej.nih.gov/ij/) to calculate the rosette area of the plants, as described by Bechtold *et al.* (2010).

2.3.3. Thermal imaging of drought-stressed and watered plants

Thermal images of both droughted and watered plants were taken at multiple points during the drought using the TH7100 Thermal Tracer (NEC Avio Infra-red Technologies Co. Ltd., Tokyo, Japan), as described by Bechtold *et al.* (2010). The images were analysed using the TH91-719 software (NEC San-ei Instruments, Ltd., Tokyo, Japan) to extract the average temperature of the rosettes.

2.3.4. Instantaneous measurements of assimilation rate and stomatal conductance

Instantaneous measurements of assimilation rate (A) and stomatal conductance (g_s) were performed under growth conditions using a CIRAS-1 (PP Systems, Amesbury, USA) as described by Lawson and Weyers (1999). The measurements were taken at multiple points during the drought for both droughted and watered plants.

2.3.5. Measurement of hydrogen peroxide

100 mg of leaf material was harvested in liquid nitrogen and ground on ice with 500 µl of 100 mM HCl. The samples were centrifuged at 12,000 g for 10 minutes at 4 °C. The supernatant was purified through an activated charcoal column to remove interfering contaminants in the sample. Hydrogen Peroxide (H_2O_2) in the samples was measured using Homovanillic acid (HVA; #306-08-1, Sigma-Aldrich®; Guilbault *et al.*, 1967). A master mix of 100 µl of 50mM HVA and 50 µl of 20 U/ml horseradish peroxidase (#P8375, Sigma-Aldrich®) was made and made up to 5 ml with 50mM HEPES. To 45 µl of the master mix, 5 µl of the extract was added and measured at an excitation of 315 nm and an emission of 425 nm using a FLUOstar Omega microplate reader (BMG LABTECH, Offenburg, Germany). The values obtained were compared against a standard curve of known concentrations of H_2O_2 to determine the concentration of H_2O_2 in the samples.

2.3.6. Determination of electrolyte leakage

Electrolyte leakage was determined as described by Shi *et al.* (2012). 100 mg of fresh leaf material was used to measure electrolyte leakage in droughted and watered plants. The leaf material was cut into three pieces with a clean, sharp blade and washed three times with RO water to remove surface and cut-generated electrolytes. The cut leaf pieces were incubated in 10 ml of RO water for 24 hours in the dark. The initial conductivity of the solution (C_i) was measured using the FG3 FiveGo™ Portable Conductivity Metre (Mettler Toledo™, ThermoFisher Scientific), as per the manufacturer's instructions. The samples were boiled for 20 minutes in a boiling water bath, allowed to cool to RT and then the conductivity of the lysed tissue (C_{max}) was measured. The electrolyte leakage of the plants was determined using the formula: (C_i / C_{max}) x 100.

2.3.7. Calculation of relative leaf water content

Relative leaf water content was measured as described by Bechtold *et al.* (2016). Rosettes of plants subjected to drought and their corresponding controls were detached from the rest of the plant and the fresh weight was measured. They were then soaked in wet tissue paper for 1 week at 4 °C to saturate them with water. The weight of the saturated rosettes was taken and they were subsequently dried out completely and the dry weight taken. The relative leaf water content was calculated using the formula:

$$\frac{\text{Fresh Weight} - \text{Dry Weight}}{\text{Saturated Weight} - \text{Dry Weight}} \times 100$$

2.3.8. Measurement of Chlorophyll and Carotenoid content

50 mg of leaf material was frozen on dry ice and used to measure the level of Chlorophyll a, Chlorophyll b, Carotenoid content and Total Chlorophyll in droughted and watered plants using the equations of Hill *et al.* (1985). The samples were ground in 1 ml of 80% Acetone using a mortar and pestle kept on ice. The ground sample was centrifuged at 10,000 g for 5 minutes at 4 °C. 200 µl of the supernatant was directly used to measure absorbance at 470 nm, 646 nm, 652 nm and 663 nm using a SPECTROstar Omega microplate reader (BMG LABTECH). Chlorophyll and Carotenoid content per fresh weight (µg/g) were measured using the following formulae:

$$\text{Chlorophyll a} = \frac{(12.15 \times A_{663}) - (2.55 \times A_{646})}{\text{Fresh Weight}}$$

$$\text{Chlorophyll b} = \frac{(18.29 \times A_{646}) - (4.58 \times A_{663})}{\text{Fresh Weight}}$$

$$\text{Carotenoids} = \frac{((1000 \times A_{470}) - (3.27 \times \text{Chlorophyll a}) - (104 \times \text{Chlorophyll b})) / 229}{\text{Fresh Weight}}$$

$$\text{Total Chlorophyll} = \frac{(27.8 \times A_{652})}{\text{Fresh Weight}}$$

2.3.9. Measurement of Anthocyanin content

Anthocyanin content was measured as described by Nakata *et al.* (2013). 50 mg of leaf material was frozen on dry ice and ground in 1ml (20 times the volume of fresh weight) extraction buffer using a mortar and pestle. The extraction buffer consisted of 45% Methanol and 5% Acetic Acid. This was then centrifuged at 12,000 g for 5 minutes at 4 °C. The supernatant was used directly to measure absorbance at 530 nm and 657 nm. The anthocyanin content (per g of fresh weight) in the samples was measured using the formula:

$$A_{530}/\text{g F.W.} = (A_{530}) - (0.25 \times A_{657}) \times 20$$

2.3.10. Measurement of flowering time

The flowering time of the droughted and watered plants was estimated by two methods: the number of leaves at emergence of the primary inflorescence (Bechtold *et al.*, 2010) and the number of days from sowing till the emergence of the primary inflorescence.

Chapter 3

Modelling drought time-series transcriptome data to identify ‘hub’ genes in gene regulatory networks

3.1. Introduction

A number of studies have been done that have measured whole transcriptome changes during drought (Seki *et al.*, 2002; Matsui *et al.*, 2008; Harb *et al.*, 2010, among others). However, these studies have only analysed a single time-point after the plant has experienced severe drought stress, and have not measured the changing dynamic as the drought progresses. The work described in this thesis is based on the work of Bechtold *et al.* (2016) in which a slow-drying time-series drought experiment was performed and whole transcriptome analysis was done using microarrays.

Differentially expressed genes from this analysis were used to model gene regulatory networks (GRNs) using Variational Bayesian State Space Modelling (VBSSM; Beal *et al.*, 2005). GRNs are generated that consist of 'hub' genes that are shown to have an interaction, either directly or indirectly, with other genes in the model, and thus the model is based around the 'hub' gene. The output of the modelling produces a directed diagram consisting of genes or 'nodes' connected to other nodes by 'edges' that indicate either activation or repression of the downstream gene. Nodes with a large number of connections originating from it are called 'hubs' and are the focus of further studies.

VBSSM has been used previously to generate GRNs for senescence (Breeze *et al.*, 2011). A model for the interactions of *ANAC092* was produced which correctly predicted known interactions, as well as generating novel interactions for further experimentation. Thus, VBSSM has good potential in identifying novel interactions between genes that would not normally be deduced based on transcriptome data only. One disadvantage of VBSSM is that only a limited number of genes (not more than 100) can be modelled at a time, and hence it is useful to cluster the list of differentially expressed genes into groups of smaller numbers of genes.

Clustering partitions a set of genes into groups based on the similarity of the genes. The similarity is usually determined based on simple vector distances, like Euclidean distance or Pearson's correlation coefficient (Thalamuthu *et al.*, 2006; Yona *et al.*, 2006). These algorithms find genes whose expression profiles look similar, i.e. they have highly correlated expressions, but are not suitable to picking out genes that regulate others. Genes that are regulated by other genes may have similar profiles to the regulating gene, but with a time lag, and these genes will not be clustered together using the above clustering algorithms (Kiddle *et al.*, 2010). Temporal Clustering by Affinity Propagation (TCAP) allows for the clustering of genes that have similar expression profiles even if the

profiles or delayed or inverted, and thus can account for genes that are regulated by other genes (Kiddle *et al.*, 2010).

This chapter describes the selection of genes for further analysis using TCAP and then the selection of clusters for modelling GRNs using VBSSM. Knockout lines and overexpressors for the hub genes were screened for subsequent phenotyping of these mutants under drought conditions, to elucidate the role, if any, of the identified hub genes in the drought response and to verify the GRN.

3.2. Selection and hierarchical clustering of list of genes and cluster annotation

In the work of Bechtold *et al.* (2016), a time-series transcriptomics analysis was performed on drought-stressed Arabidopsis Col-4 plants. These plants were subjected to a progressive drought in which water was withdrawn until a relative soil water content of 20% was reached. This took 14 days, and each day drought-stressed and watered control plants were sampled during the middle of the photoperiod. RNA was extracted from these plant samples and used to measure the expression of >32,000 genes during drought and watered. All the genes were subjected to a Gaussian process two-sample test (GP2S; Stegle *et al.*, 2010) to identify differentially expressed genes between the two conditions, and each gene was scored accordingly and ranged from 50.49 to -22.12 (Bechtold *et al.*, 2016).

The work described in this thesis made use of the above list of genes. Genes with a GP2S score ≥ 5 were selected and the expression profiles of these genes were visually inspected to confirm differential expression under drought. Genes were considered differentially expressed when there was no overlap of the respective gene expression profiles between the drought-stressed and normally watered plants for at least two consecutive time-points. A total of 2190 unique genes were extracted from this list (this list of genes is given in Appendix A) and used for further analysis.

The list of 2190 genes was clustered using the clustering algorithm Temporal Clustering by Affinity Propagation (TCAP; Kiddle *et al.*, 2010). TCAP clusters genes based on similarity in expression profiles, including those with a time lag or an inverted expression profile, and hence is able to cluster genes that are not only co-regulated but potentially regulate each other. Using TCAP, all 2190 genes were clustered into 153 clusters, and each cluster consists of up- or down-regulated genes with similar gene expression profiles. The expression profiles of all the genes in Cluster 1, which was used subsequently to model gene networks in drought, are shown as an example in Figure 3.1. It shows how genes of similar expression profiles, though opposite in their expression pattern, are clustered together in Cluster 1. TCAP also aligns all the genes within the cluster, and from Figure 3.1B it can be seen that, overall, Cluster 1 is induced during drought.

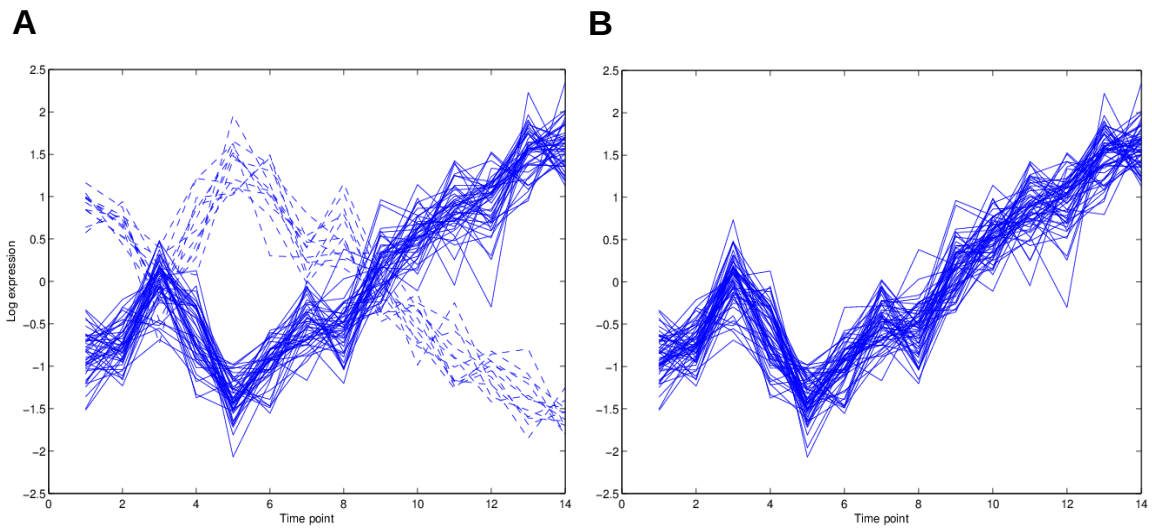


Figure 3.1. An example of a gene cluster obtained from TCAP. The expression profiles of the genes in the cluster are shown; the expression profiles of the majority of the genes in the cluster are shown in continuous lines, while genes with an inverted expression profile are shown in dashed lines.

The 153 clusters obtained from the output of TCAP were analysed for over-represented terms related to Biological Processes using the Database for Annotation, Visualization and Integrated Discovery (DAVID v6.7; Huang *et al.*, 2009). Enriched terms were identified for only 88 of these clusters, however, only 26 clusters were considered to have significantly enriched terms, with a P-value < 0.01 and a False Discovery Rate (FDR) ≤ 3. Table 3.1 shows significantly enriched terms for all 26 clusters and the list of genes in each cluster is given in Appendix A.

Table 3.1. The significantly over-represented terms for each of the clusters obtained from TCAP.

Cluster Number	Direction of expression	Over-represented terms
1	Up	GO:0009056~catabolic process
2	Down	GO:0034728~nucleosome organization GO:0006334~nucleosome assembly GO:0031497~chromatin assembly GO:0065004~protein-DNA complex assembly GO:0006323~DNA packaging GO:0006333~chromatin assembly or disassembly GO:0034622~cellular macromolecular complex assembly GO:0044085~cellular component biogenesis GO:0006325~chromatin organization GO:0034621~cellular macromolecular complex subunit organization GO:0051276~chromosome organization GO:0065003~macromolecular complex assembly GO:0043933~macromolecular complex subunit organization GO:0015979~photosynthesis GO:0022607~cellular component assembly
8	Up	GO:0044237~cellular metabolic process GO:0009987~cellular process
11	Up	GO:0031537~regulation of anthocyanin metabolic process GO:0043455~regulation of secondary metabolic process GO:0050896~response to stimulus GO:0042221~response to chemical stimulus GO:0006950~response to stress
14	Down	GO:0006412~translation GO:0034645~cellular macromolecule biosynthetic process GO:0009059~macromolecule biosynthetic process GO:0010467~gene expression GO:0044267~cellular protein metabolic process GO:0019538~protein metabolic process GO:0044249~cellular biosynthetic process GO:0009058~biosynthetic process GO:0044260~cellular macromolecule metabolic process GO:0043170~macromolecule metabolic process

		GO:0042254~ribosome biogenesis GO:0022613~ribonucleoprotein complex biogenesis GO:0044237~cellular metabolic process GO:0044238~primary metabolic process GO:0044085~cellular component biogenesis GO:0008152~metabolic process GO:0009987~cellular process
16	Down	GO:0009733~response to auxin stimulus GO:0009725~response to hormone stimulus
18	Down	GO:0006412~translation GO:0034645~cellular macromolecule biosynthetic process GO:0009059~macromolecule biosynthetic process GO:0044249~cellular biosynthetic process GO:0009058~biosynthetic process
19	Up	GO:0015995~chlorophyll biosynthetic process GO:0015994~chlorophyll metabolic process GO:0006779~porphyrin biosynthetic process GO:0033014~tetrapyrrole biosynthetic process GO:0006778~porphyrin metabolic process GO:0033013~tetrapyrrole metabolic process GO:0046148~pigment biosynthetic process GO:0042440~pigment metabolic process
20	Up	GO:0009737~response to abscisic acid stimulus
21	Down	GO:0006325~chromatin organization GO:0051276~chromosome organization GO:0006334~nucleosome assembly GO:0034728~nucleosome organization GO:0031497~chromatin assembly GO:0065004~protein-DNA complex assembly GO:0006323~DNA packaging GO:0006333~chromatin assembly or disassembly
28	Down	GO:0045087~innate immune response GO:0006955~immune response GO:0002376~immune system process
40	Down	GO:0044249~cellular biosynthetic process GO:0009058~biosynthetic process

		GO:0006412~translation
42	Up	GO:0080028~nitrile biosynthetic process GO:0050898~nitrile metabolic process
48	Down	GO:0006412~translation GO:0044249~cellular biosynthetic process GO:0034645~cellular macromolecule biosynthetic process GO:0009059~macromolecule biosynthetic process GO:0009058~biosynthetic process GO:0010467~gene expression GO:0044267~cellular protein metabolic process GO:0019538~protein metabolic process GO:0044237~cellular metabolic process GO:0044238~primary metabolic process
49	Down	GO:0005975~carbohydrate metabolic process GO:0044262~cellular carbohydrate metabolic process GO:0044238~primary metabolic process GO:0006796~phosphate metabolic process GO:0006793~phosphorus metabolic process
55	Down	GO:0006412~translation
69	Up	GO:0042398~cellular amino acid derivative biosynthetic process GO:0009813~flavonoid biosynthetic process GO:0009812~flavonoid metabolic process GO:0006575~cellular amino acid derivative metabolic process GO:0009718~anthocyanin biosynthetic process GO:0046283~anthocyanin metabolic process GO:0019748~secondary metabolic process GO:0009699~phenylpropanoid biosynthetic process GO:0009698~phenylpropanoid metabolic process GO:0006519~cellular amino acid and derivative metabolic process GO:0019438~aromatic compound biosynthetic process GO:0006725~cellular aromatic compound metabolic process GO:0046148~pigment biosynthetic process GO:0042440~pigment metabolic process
87	Up	GO:0009698~phenylpropanoid metabolic process GO:0006575~cellular amino acid derivative metabolic process GO:0009813~flavonoid biosynthetic process

		GO:0009812~flavonoid metabolic process GO:0006725~cellular aromatic compound metabolic process GO:0019748~secondary metabolic process GO:0009699~phenylpropanoid biosynthetic process GO:0006519~cellular amino acid and derivative metabolic process GO:0042398~cellular amino acid derivative biosynthetic process GO:0019438~aromatic compound biosynthetic process
94	Down	GO:0006412~translation
103	Up	GO:0044272~sulfur compound biosynthetic process GO:0006790~sulfur metabolic process GO:0016144~S-glycoside biosynthetic process GO:0019761~glucosinolate biosynthetic process GO:0019758~glycosinolate biosynthetic process GO:0019760~glucosinolate metabolic process GO:0016143~S-glycoside metabolic process GO:0019757~glycosinolate metabolic process GO:0016138~glycoside biosynthetic process GO:0016137~glycoside metabolic process GO:0034637~cellular carbohydrate biosynthetic process GO:0016051~carbohydrate biosynthetic process GO:0019748~secondary metabolic process GO:0044262~cellular carbohydrate metabolic process
131	Up	GO:0050794~regulation of cellular process GO:0050789~regulation of biological process
142	Down	GO:0009165~nucleotide biosynthetic process GO:0034404~nucleobase, nucleoside and nucleotide biosynthetic process GO:0034654~nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process GO:0009117~nucleotide metabolic process GO:0006753~nucleoside phosphate metabolic process GO:0055086~nucleobase, nucleoside and nucleotide metabolic process GO:0044271~nitrogen compound biosynthetic process
152	Up	GO:0042180~cellular ketone metabolic process

The gene clusters that are upregulated during drought include those that are involved in various cellular metabolic processes (clusters 1, 8, 131, 152), regulation of anthocyanin metabolism and stress response (clusters 11, 69), chlorophyll biosynthetic genes (cluster 19), response to abscisic acid stimulus (cluster 20), nitrile biosynthetic processes (cluster 42), flavonoid biosynthetic processes (clusters 69, 87) and glucosinolate biosynthetic processes (cluster 103). On the other hand, genes that are downregulated during drought are those involved in nucleosome and chromatin organisation and assembly (clusters 2, 21), translation (clusters 14, 18, 40, 48, 55, 94), response to auxin (cluster 16), innate immune response (cluster 28), carbohydrate metabolic processes (cluster 49) and nucleotide biosynthetic processes (cluster 142).

In this work, four clusters of genes were selected to model gene regulatory networks using VBSSM. These are Clusters 1, 2, 8 and 16. The salient aspects of these gene clusters that explain the rationale for choosing these clusters are described below.

Cluster 1 contains genes that are enriched for catabolic processes. A closer look at the genes that make up this category reveals that they are involved in proteolysis and protein catabolic processes. These genes encode the following proteins: nuclear pore localisation protein (NPL4) involved in proteasome-mediated ubiquitin-dependent protein catabolic process (UniProt); F-box protein, FBW2, which forms a part of the SCF (ASK-cullin-F-box) E3 ubiquitin ligase complex (Earley *et al.*, 2010); a ubiquitin-conjugating enzyme 32 (UBC32); Autophagy-related protein 3 (ATG3), an E2 protein (Yamaguchi *et al.*, 2012); Regulatory Particle AAA-ATPASE 2B (RPT2B), that forms a part of the proteasome complex (Lee *et al.*, 2011). In addition to proteolysis-related genes, genes involved in other catabolic processes such as homogentisate 1, 2-dioxygenase (HGO), involved in tyrosine catabolism (Dixon and Edwards, 2006); Pheophytinase (PPH), involved in chlorophyll breakdown (Ren *et al.*, 2010); Peroxin 10 (Pex10), involved in importing proteins to the peroxisome (Schumann *et al.*, 2003) are also in this cluster. Thus, it appears that genes involved in different catabolic processes are present in cluster 1, and most of the genes in this cluster are up-regulated during drought, indicating there is increased catabolic activity in cells experiencing water deficit.

Cluster 8 was enriched for cellular metabolic processes and the genes in cluster 8 are involved in a range of different cellular metabolic processes. The gene products in this cluster include a number of transcription factors, a calmodulin-binding protein, a CBL-interacting protein kinase, Asparagine Synthetase, (ASN2) and Trehalose-6-Phosphate Phosphatase H (TPPH). Thus, it appears that there are a number of functionally different

genes in this cluster that are possibly co-expressed and co-regulated, and so the genes in this cluster were also chosen to model gene regulatory networks involved in the drought response.

In addition to these two gene clusters, clusters 2 and 16 were also chosen for the modelling. The genes in cluster 2 are involved in chromatin and nucleosome organisation and assembly and DNA packaging. These genes are down-regulated during drought, indicating that there is increased gene expression activity in drought-stressed plants. Cluster 16 consists of genes that are involved in response to auxin stimulus, in particular Small Auxin Up-regulated RNA (SAUR) genes. This group of genes is induced by auxin and promotes cell expansion (Spartz *et al.*, 2012). Thus, during drought, cell expansion is inhibited through the down-regulation of these genes. Modelling these clusters could help identify drought-mediated regulatory genes involved in gene expression and plant growth.

3.3. Gene networks were modelled and 'hub' genes identified from these models

Gene regulatory networks (GRNs) were modelled using Variational Bayesian State Space Modelling (VBSSM; Beal *et al.*, 2005). VBSSM is a dynamic Bayesian algorithm that uses time-series transcriptomics data to generate models of gene regulation. These models could help identify potentially important drought-responsive regulatory genes. One restriction of VBSSM is the number of genes that can be input (less than 100 genes), and so clustering all 2190 genes using TCAP not only helped to group together genes of similar expression profiles, but also created smaller genes lists that are suitable to input into VBSSM. The algorithm was run in MATLAB® with the settings: Number of seeds = 15; Maximum dimension of hidden states = 20; Maximum number of iterations = 2000. The networks with a confidence of 95% were used and viewed using Cytoscape 2.8.2.

Six GRNs were generated from the four clusters and the list of genes in each network is given in Appendix A. Two GRNs were generated using the genes in Cluster 1 which led to the identification of two hub genes, one from each network: an unknown hub gene whose protein contains the BTB/POZ domain (henceforth known as *POZ*), and *Flowering Locus D* (*FD*), a bZIP (Basic Leucine Zipper) protein involved in the positive regulation of flowering (Abe *et al.*, 2005). *POZ* was connected to all the genes in Cluster 1, while *FD* was connected to 55 out of a total of 56 genes. Both models generated from Cluster 1 are shown in Figures 3.3A and 3.4A. The modelling of genes from Cluster 2 generated a

gene network model that identified another unknown gene, designated as *UKTF* (Figure 3.5A), which was connected to 28 of the 57 genes in Cluster 2. From the genes in Cluster 8, two GRNs were generated and from these models two hub genes were identified: *Related to AP2.12* (*RAP2.12*; Figure 3.6A), a member of the ERF (Ethylene Response Factor) subfamily of the ERF/AP2 transcription factor family which has been found to be stress-responsive (Licausi *et al.*, 2011; Papdi *et al.*, 2015); *Basic Helix-Loop-Helix038* (*BHLH038*; Figure 3.7A), a protein that regulates iron homeostasis (Wang *et al.*, 2007). *RAP2.12* was connected to 23 of the 40 genes in the cluster, while *BHLH038* was connected to 36 genes. In the final GRN from genes in Cluster 16, *Anthocyaninless2* (*ANL2*; Figure 3.8A), a protein involved in the accumulation of anthocyanin and in root development (Kubo *et al.*, 1999), was identified as a hub gene in the gene network and was connected to all 24 genes in the cluster. In addition to these six genes, the flowering time gene *Agamous-like22* (*AGL22*; Hartmann *et al.*, 2000) was identified as a hub gene by Bechtold *et al.* (2016), and together with the above six, was also included in the analysis described in this thesis.

The expression profiles of all seven genes, as determined using the time-series transcriptomics analysis (Bechtold *et al.*, 2016), are shown in Figures 3.2-3.8. Genes were classified as early- or late-responding depending on whether differential expression occurred between days 1 to 7 or between days 8 to 13, respectively, as similarly described in Bechtold *et al.* (2016). Among the early-responsive genes were *AGL22* (Figure 3.2) and *UKTF* (Figure 3.5B), which were induced at days 7 and 6, respectively. The other genes were late-responding genes: *POZ* was induced on day 8 (Figure 3.3B), *BHLH038* was down-regulated on day 9 (Figure 3.7B), *FD* (Figure 3.4B) and *RAP2.12* (Figure 3.6B) were differentially expressed on day 10, and *ANL2* was induced on day 12 (Figure 3.8B).

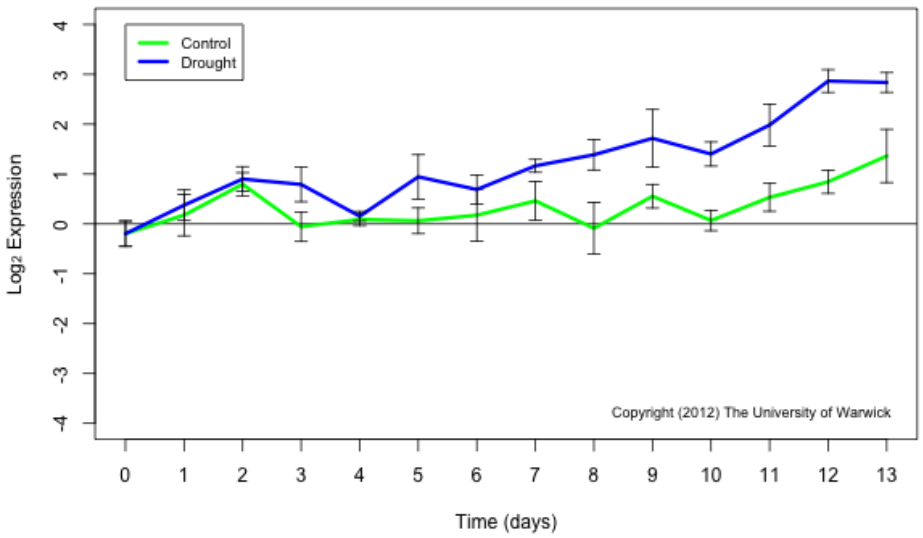


Figure 3.2. The expression profile of *AGL22* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

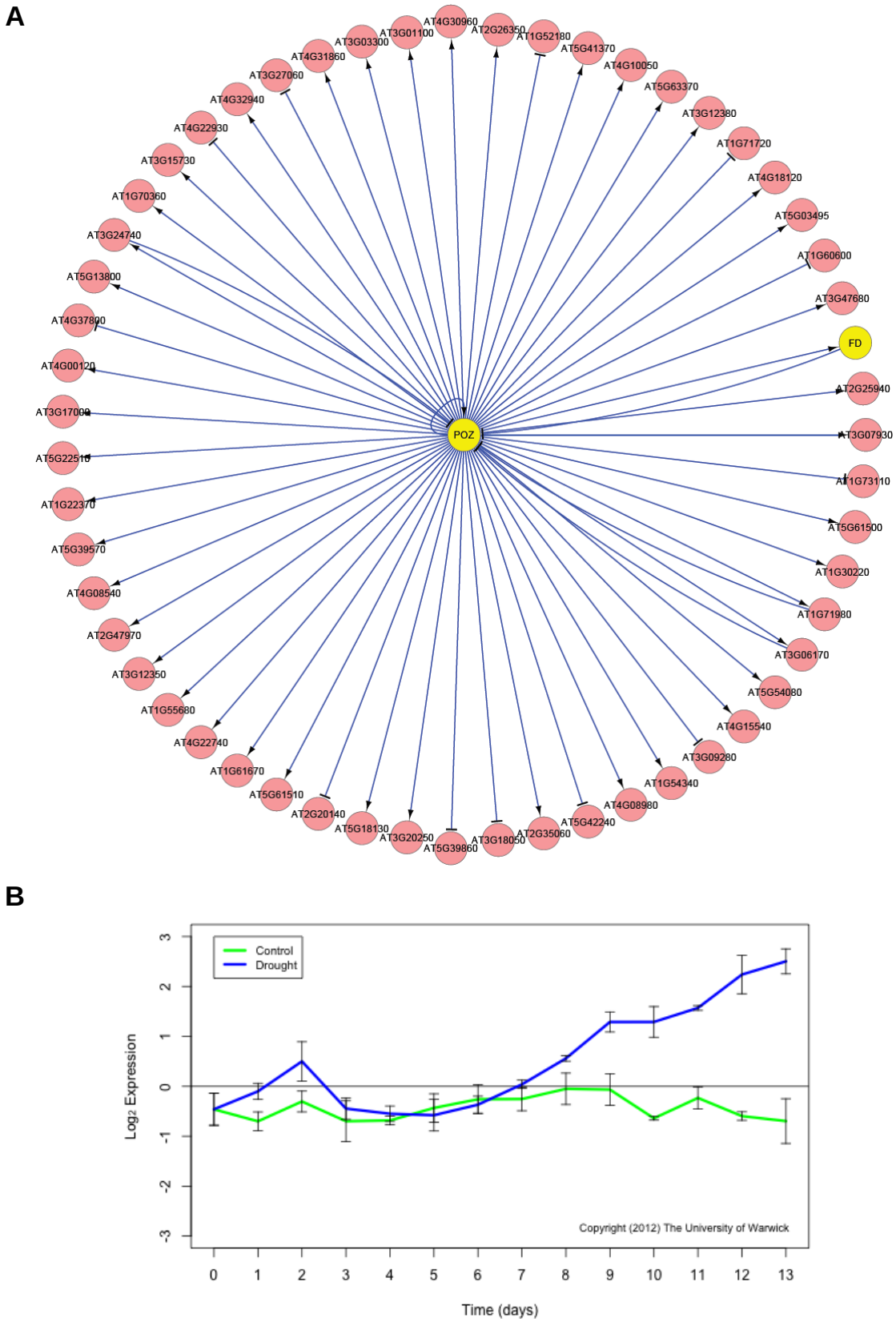


Figure 3.3. Identification of the hub gene *POZ*.
(A) Gene regulatory network modelled from cluster 1 using VBSSM which identified *POZ* as a hub gene. The network with a confidence level of 95% is shown here. All 56 genes in the cluster were connected to *POZ*, including *FD*. (B) The expression profile of *POZ* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

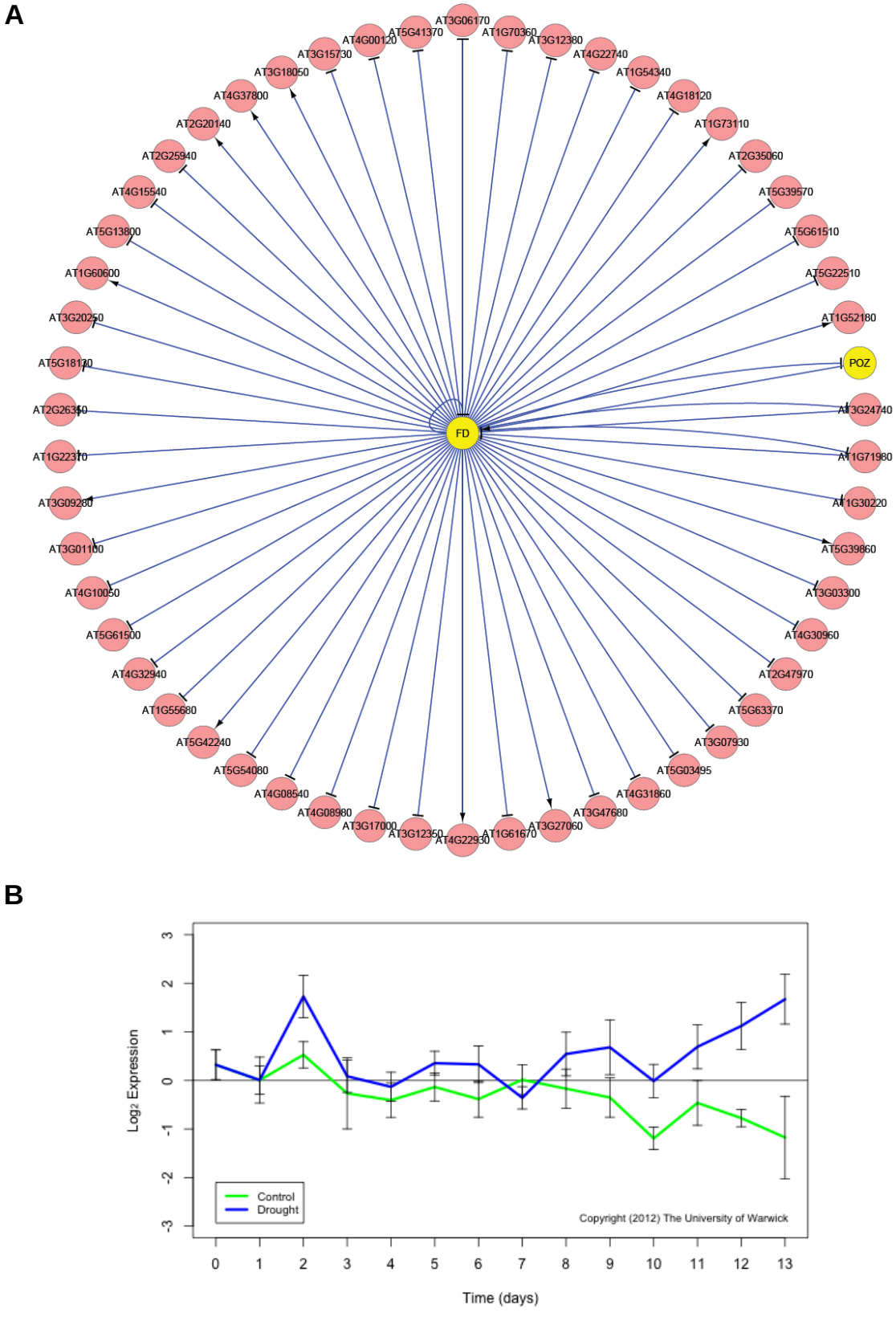


Figure 3.4. Identification of the hub gene *FD*.
(A) Gene regulatory network modelled from cluster 1 using VBSSM which identified *FD* as a hub gene. The network with a confidence level of 95% is shown here. 55 of the 56 genes in the cluster were connected to *FD*, including *POZ*. (B) The expression profile of *FD* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

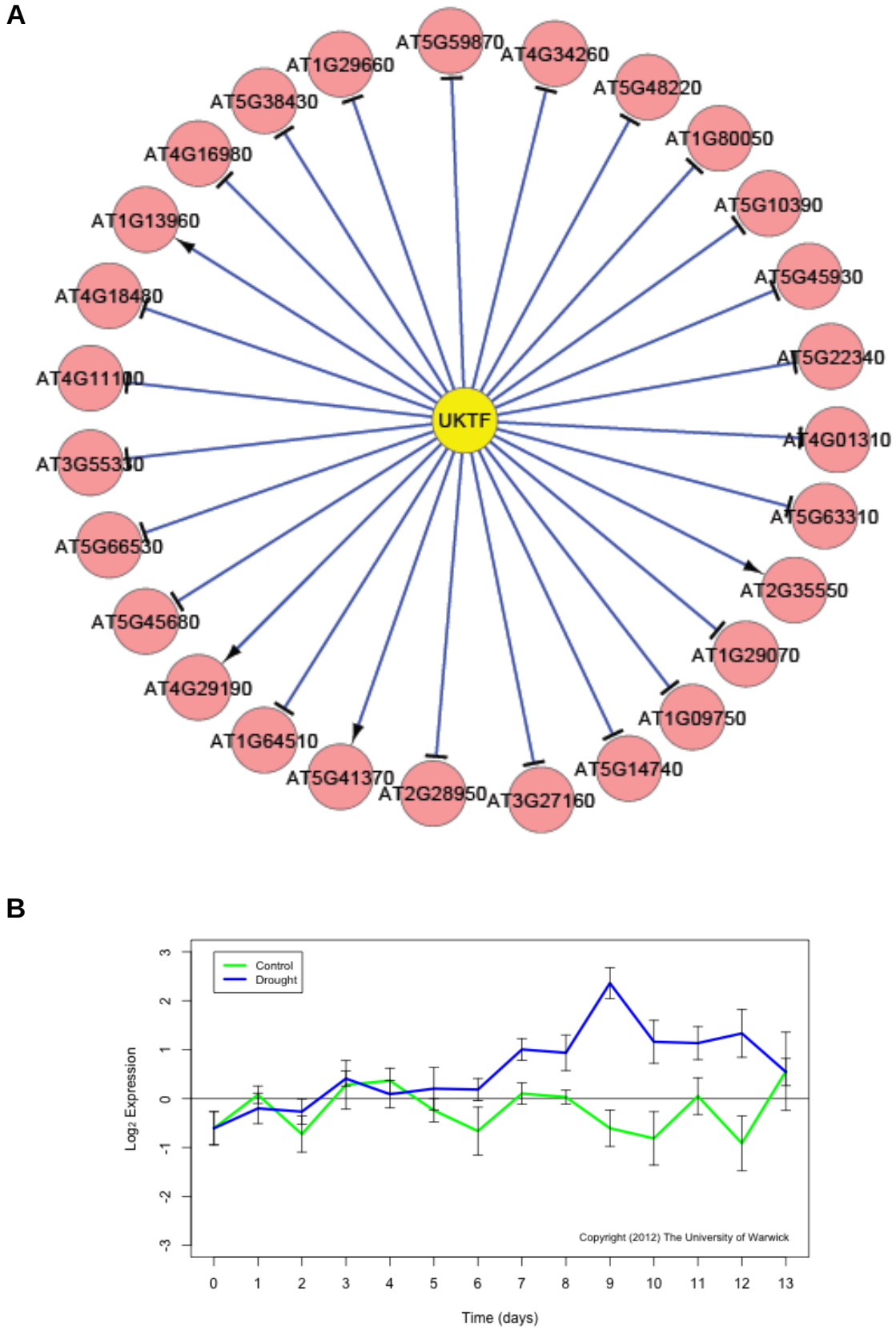


Figure 3.5. Identification of the hub gene *UKTF*.
(A) Gene regulatory network modelled from cluster 2 using VBSSM which identified *UKTF* as a hub gene. The network with a confidence level of 95% is shown here. 28 of the 57 genes in the cluster were connected to *UKTF*. (B) The expression profile of *UKTF* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

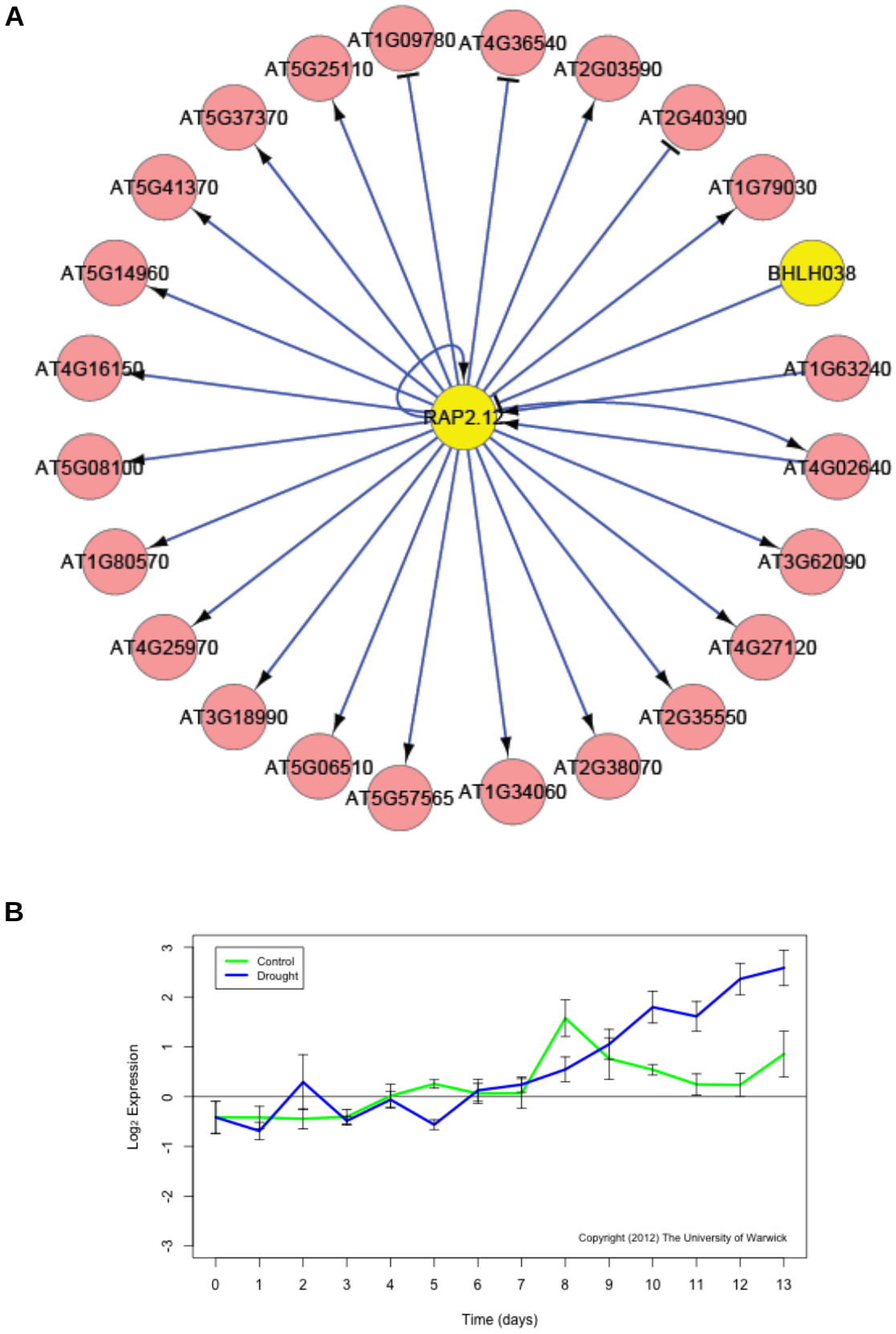


Figure 3.6. Identification of the hub gene *RAP2.12*.
(A) Gene regulatory network modelled from cluster 8 using VBSSM which identified *RAP2.12* as a hub gene. The network with a confidence level of 95% is shown here. 23 of the 40 genes in the cluster were connected to *RAP2.12*, including *BHLH038*. (B) The expression profile of *RAP2.12* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

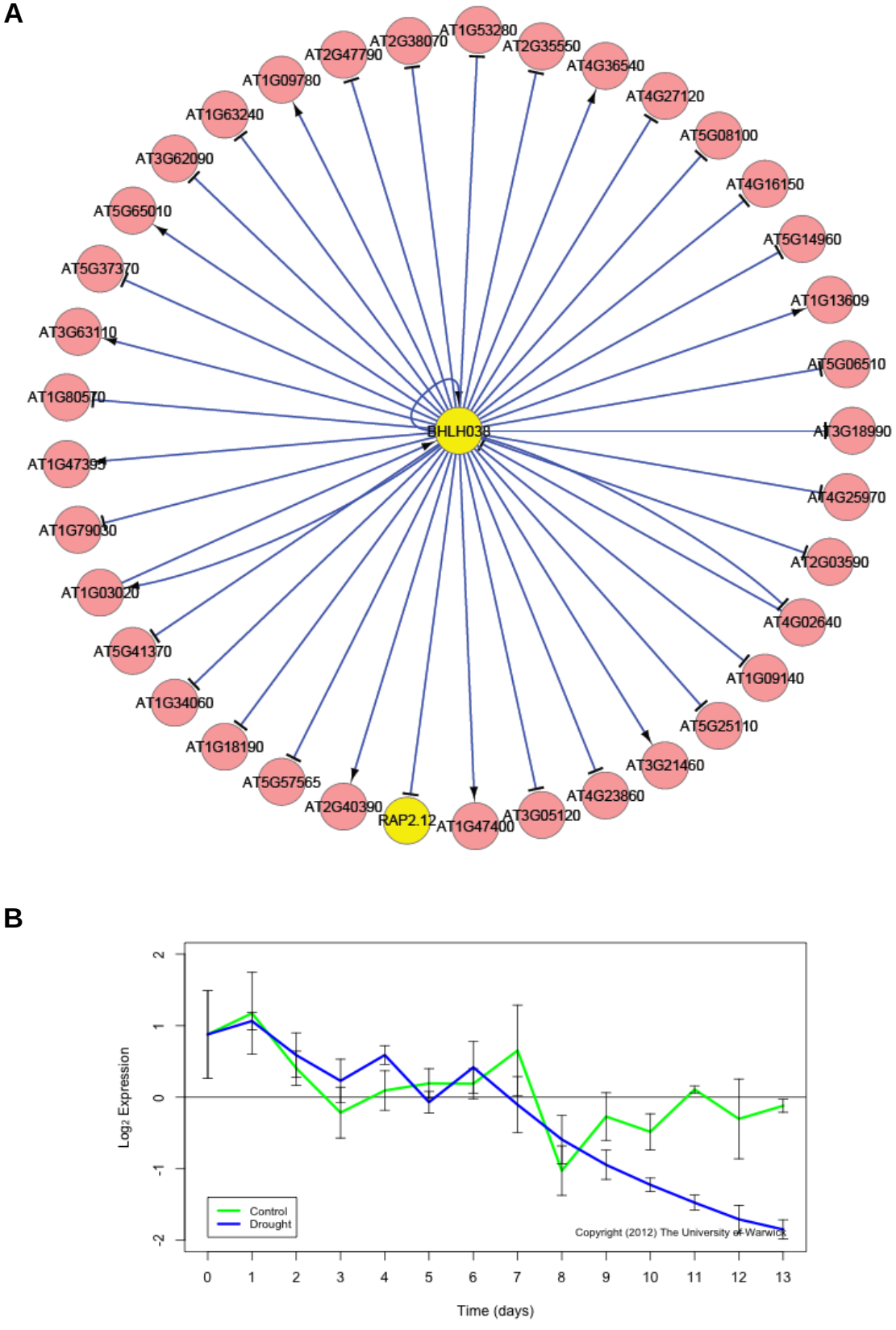


Figure 3.7. Identification of the hub gene *BHLH038*.
(A) Gene regulatory network modelled from cluster 8 using VBSSM which identified *BHLH038* as a hub gene. The network with a confidence level of 95% is shown here. 36 of the 40 genes in the cluster were connected to *BHLH038*, including *RAP2.12*. (B) The expression profile of *BHLH038* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

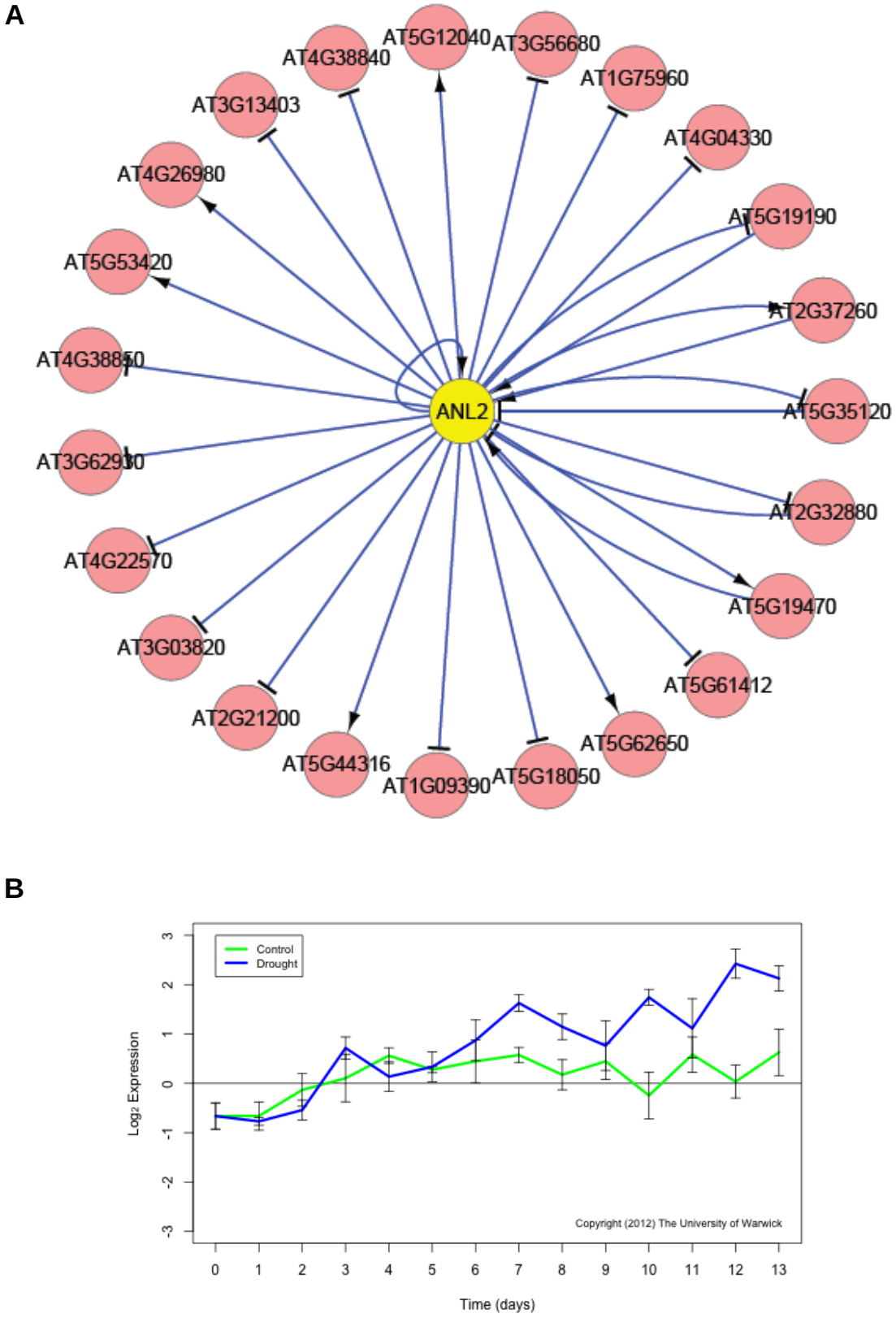


Figure 3.8. Identification of the hub gene *ANL2*.
(A) Gene regulatory network modelled from cluster 16 using VBSSM which identified *ANL2* as a hub gene. The network with a confidence level of 95% is shown here. All the genes in the cluster were connected to *ANL2*. (B) The expression profile of *ANL2* under drought-stressed (blue) and control (green) conditions over the course of the drought period as determined in Bechtold *et al.* (2016).

AGL22 and *FD* are negative (Hartmann *et al.*, 2000) and positive (Abe *et al.*, 2005) regulators of flowering time, respectively, and also affect flower development. As described in section 3.6 of this chapter, UKTF was identified as a possible MADS-Box interacting protein (MIP1). A *MIP1* gene in *Antirrhinum majus* was shown to be involved in carpel development (Causier *et al.*, 2003) and so UKTF may also be involved in carpel development. As three of the hub genes identified are potentially involved in flower development, *AGL22*, *FD* and *UKTF* were modelled with other flower development / flowering-related genes present in the list of 2190 genes (section 3.2). There were 85 genes in total and this list is given in Appendix A. Figure 3.9 shows the gene network modelled from this list. It can be seen that *AGL22* and *UKTF* are hubs in this network, while *FD* is only present as a downstream target node.

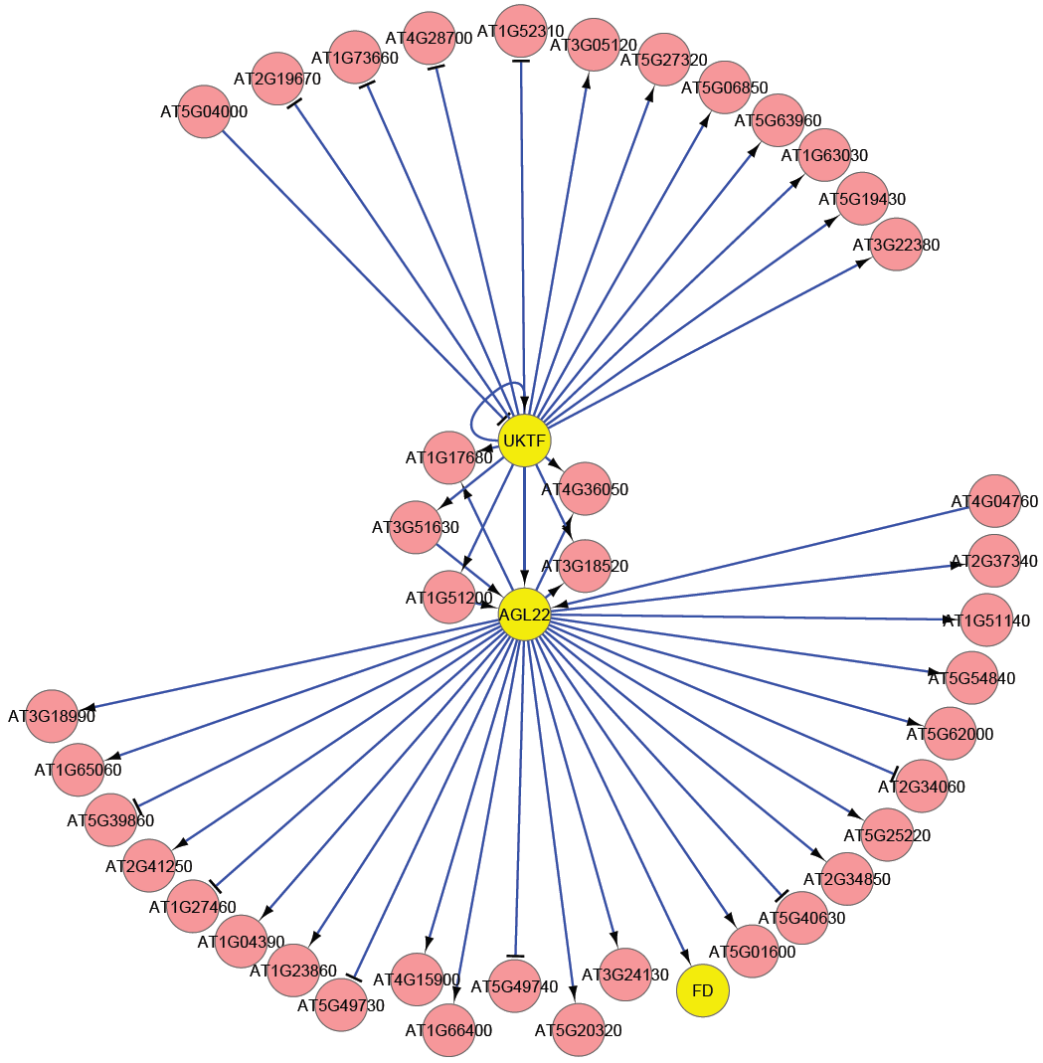


Figure 3.9. Gene regulatory network generated using genes involved in flowering, showing *AGL22* and *UKTF* as hub genes. The network with a confidence level of 95% is shown here.

An analysis of the above seven genes was performed using The Bio-Analytic Resource (BAR) for Plant Biology (http://bar.utoronto.ca/; Winter *et al.*, 2007) to identify if these genes are abiotic stress-responsive. The abiotic stresses analysed were cold, osmotic, salt, drought and heat stresses, and the data is derived from the results obtained by Killian *et al.* (2007). In their experiment using 18 day-old seedlings grown on liquid media, cold stress was applied by maintaining the temperature of the medium at 4 °C using ice and a cold chamber, while osmotic and salt stresses were applied using 300 mM mannitol and 150 mM NaCl, respectively. Drought stress was applied by exposing the plants to a stream of air for 15 minutes till 10% fresh weight was lost, and heat stress was applied by exposing the seedlings to 38 °C for 3 hours. Shoots and roots were sampled separately at 0, 0.5, 1, 3, 6, 12 and 24 hours for all stresses, except heat which also included 0.25 and 4 hours samples.

The results for six of the seven genes are shown in Figures 3.10-3.15, as no data was available for *BHLH038*. It can be seen that all the genes were differentially expressed in either shoots or roots or both under all stress conditions; however, only *AGL22*, *FD* and *ANL2* were drought-responsive, while *AGL22*, *POZ*, *FD*, *RAP2.12* were osmotic stress-responsive. The nature of the drought stress used could explain the discrepancy between this analysis and the results of the microarray analysis by Bechtold *et al.* (2016). Thus, it would appear that VBSSM is able to identify known abiotic stress-responsive genes which could be potentially important in regulating gene networks during drought. The subsequent sections describe the isolation of knockouts and overexpressors of these seven genes, as well as *in silico* characterisation of the unknown hub genes *UKTF* and *POZ*.

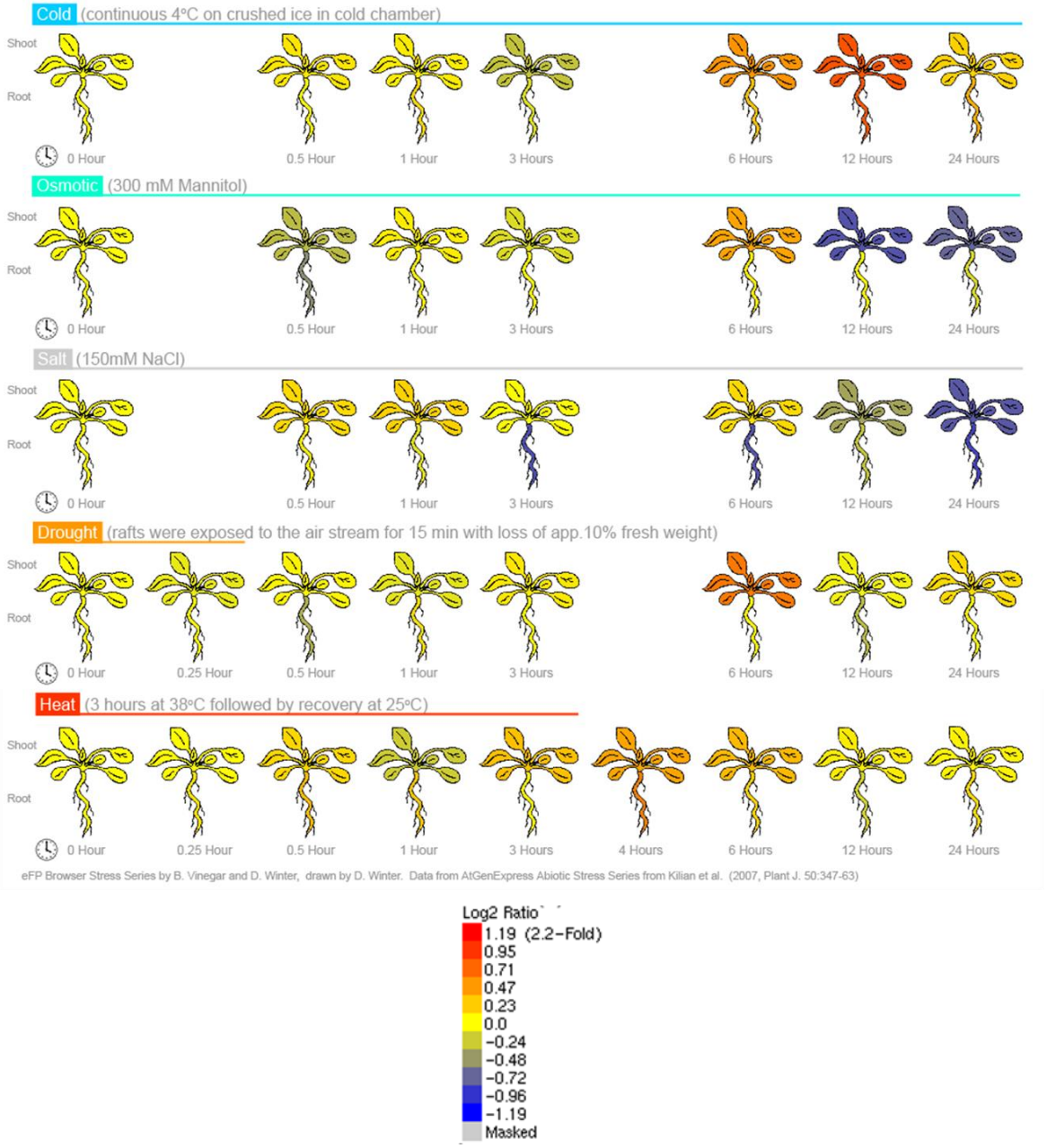


Figure 3.10. BAR analysis for *AGL22* in shoots and roots under different abiotic stresses. The gene was induced in shoots by cold, osmotic and drought stresses, and in roots by cold and heat stresses. *AGL22* was also down-regulated in shoots by osmotic and salt stresses, and in roots by salt stress.

Figure 3.11. BAR analysis for *POZ* in shoots and roots under different abiotic stresses. The gene was induced in shoots by osmotic and salt stresses. *POZ* was also down-regulated in both shoots and roots by heat stress.

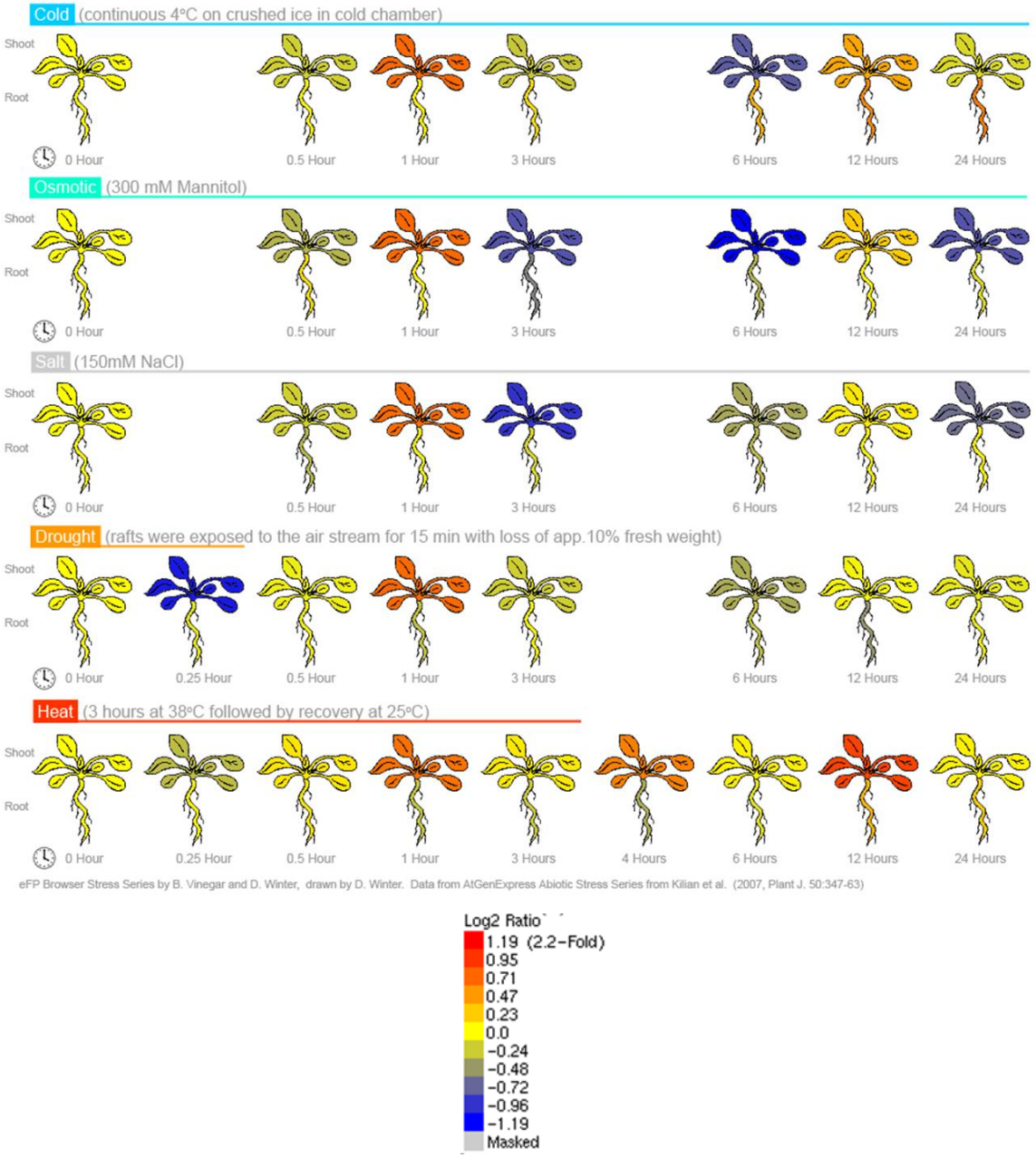


Figure 3.12. BAR analysis for *FD* in shoots and roots under different abiotic stresses. The gene was induced in shoots by cold, osmotic, salt, drought and heat stresses, and in roots by cold stress. *FD* was also down-regulated in shoots by cold, osmotic, salt and drought stresses, and in roots by osmotic stress.

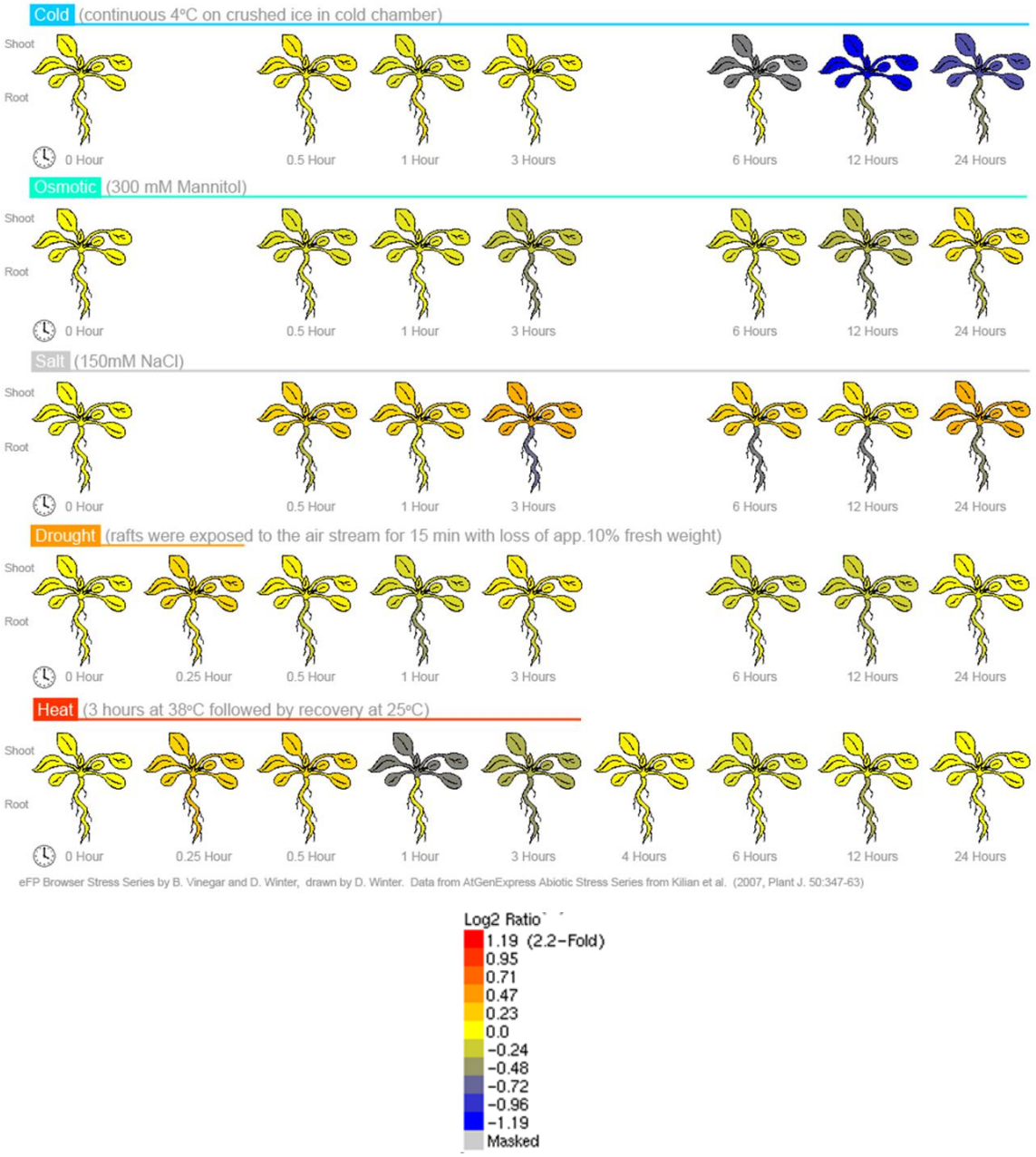


Figure 3.13. BAR analysis for *UKTF* in shoots and roots under different abiotic stresses. The gene was down-regulated in shoots by cold and heat stresses, and in roots by cold, osmotic and salt stresses.

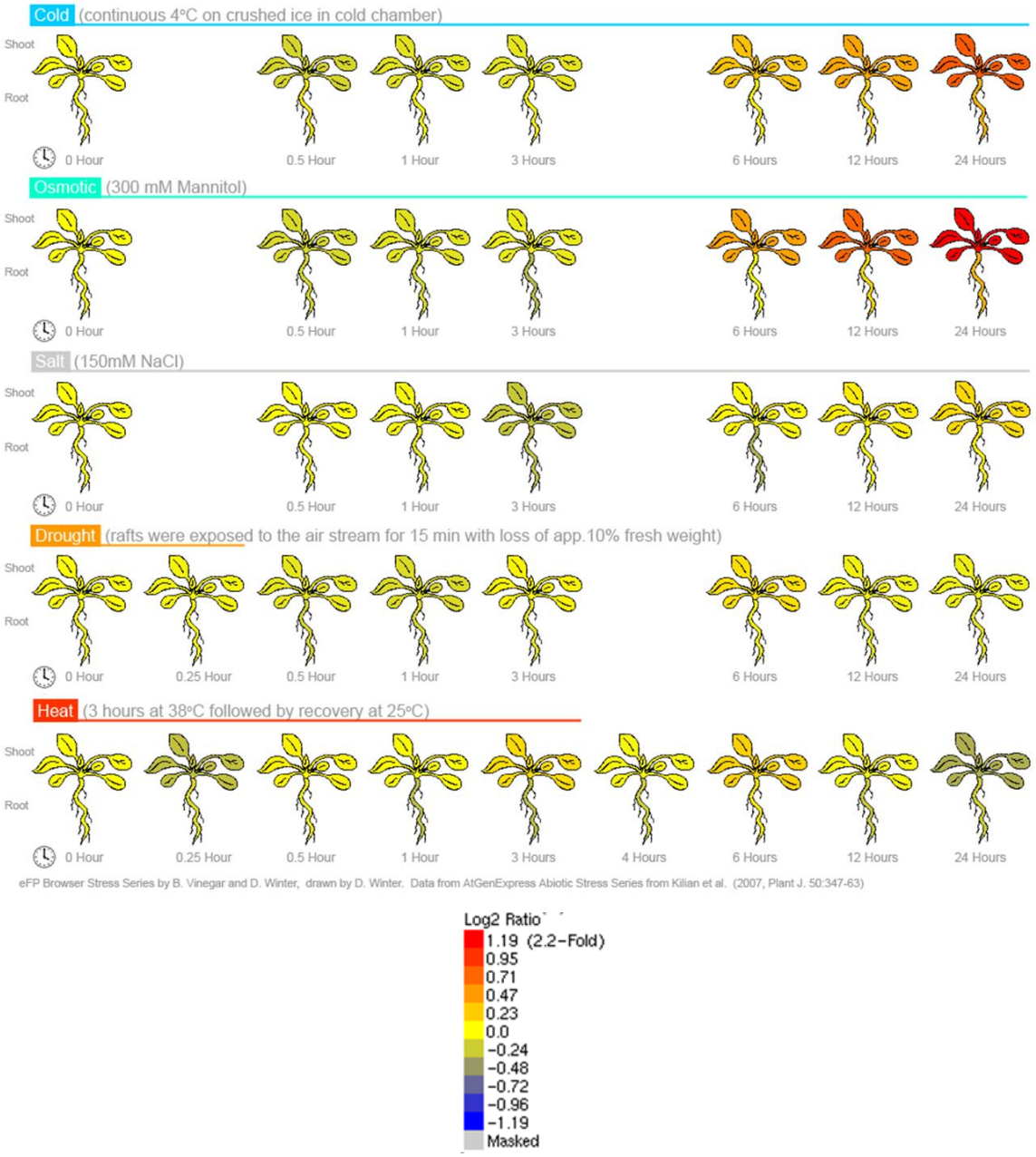


Figure 3.14. BAR analysis for *RAP2.12* in shoots and roots under different abiotic stresses. The gene was induced in shoots by cold and osmotic stresses.

Figure 3.15. BAR analysis for *ANL2* in shoots and roots under different abiotic stresses. The gene was induced in shoots by drought stress and in roots by cold stress. *ANL2* was also down-regulated in shoots by cold and heat stresses, and in roots by osmotic, salt and heat stresses.

3.4. Isolation of knockouts for hub genes

A number of T-DNA insertion lines (SALK, SAIL and GABI-Kat lines) for each of the hub genes were selected and ordered from the Nottingham Arabidopsis Stock Centre (NASC; http://arabidopsis.org.uk/). A list of all the lines ordered and screened is in Appendix B. These lines were screened by PCR to verify the presence of the T-DNA insertion, and the expression levels of the respective genes of interest were analysed using RT-PCR.

DNA was extracted from a single leaf and used in PCR to screen for T-DNA insertions in the obtained lines. As mentioned in Chapter 2, the PCR was done using gene-specific and insertion-specific primers in the combinations of: the gene-specific forward primer with the gene-specific reverse primer (F+R), the gene-specific forward primer with the insertion-specific primer (F+I) and the gene-specific reverse primer with the insertion-specific primer (R+I) as shown in Figure 2.1. At least 10 plants from each of the lines were analysed to identify at least one homozygote for the T-DNA insertion in the gene of interest. PCR products obtained in the F+I and R+I reactions were sequenced through the DNA sequencing company GATC Biotech AG (Cologne, Germany), to map the position of the insertions within the genes. Gene knockout mutants were then identified from positive, homozygous T-DNA insertion lines, by RT-PCR with primers amplifying the full-length protein-coding sequence. The sequences of all the primers used for screening and RT-PCR are available in Appendix C.

Figures 3.16 to 3.22 show the results obtained from screening knockout mutant plants. Only the lines that were subsequently used for further experiments and analyses are listed in Table 3.2 and described below. All the T-DNA insertional lines obtained from NASC and screened in this project are given in Appendix B. The figures show gel pictures confirming homozygous plants obtained from screening the T-DNA insertion lines and the position of the insertion(s) mapped within the gene by sequencing. The results of the RT-PCR showing the lack of full-length protein-coding sequence in these lines are also shown.

As shown in Figures 3.16A and C, two T-DNA insertional lines were identified for *AGL22* that were homozygous for the presence of the insertion. *agl22-3* was found to contain two insertions within the gene, the positions and orientations of which are shown in Figure 3.16B. *agl22-4* was found to contain a single T-DNA insertion in *AGL22*, as shown in Figures 3.16C and D. Figure 3.16E shows the absence of the full-length coding sequence of *AGL22* in both *agl22-3* and *agl22-4* compared to the Col-0, confirming that these two

lines are knockouts of *AGL22*. Compared to Col-0, knockouts of *AGL22* display an early flowering phenotype as shown in Figure 3.16F.

Two T-DNA insertional lines were identified for *POZ*, as shown in Figure 3.17, and both lines were found to have two insertions within the gene (Figures 3.17A and C). When the positions of the insertions were ascertained by sequencing, the insertions were found to be in exactly the same positions in both lines (Figures 3.17B and D). Despite the apparent mix-up of these two lines at NASC, they were treated as two independent lines for the remainder of this thesis, as no other T-DNA insertional lines were available for *POZ*. Both lines were confirmed as knockouts of the gene, as shown in Figure 3.17E. Under normal conditions, visually there appeared to be no difference in the phenotype of the knockouts compared to the wild-type (Figure 3.17F).

Only one T-DNA insertional mutant was identified for *FD* (*fd-5*; Figure 3.18A and B), and this was determined to be a knockdown mutant of the gene (Figure 3.18C), even though it has been published as a knockout mutant (Wigge *et al.*, 2005). There appeared to be no visual difference in phenotype between Col-0 and *fd-5* (Figure 3.18D). A knockout of *FD*, *fd-4*, was obtained from Dr. Lucio Conti, University of Milan (Wigge *et al.*, 2005; Riboni *et al.*, 2013) and this was found to have a larger rosette than Col-0 (Figure 3.18E).

Two T-DNA insertional lines of *UKTF* were identified, as shown in Figure 3.19. UKTF-3 was found to contain two T-DNA insertions within the gene (Figures 3.19A and B), while UKTF-4 had one (Figures 3.19C and D). When RT-PCR was done for these lines, it was seen that *UKTF* was overexpressed, rather than knocked out (Figure 3.19E), due to the presence of the insertions within the promoter region of the gene in both lines (Figures 3.19B and D). The up-regulation of *UKTF* in these lines was confirmed by qPCR (Figure 3.19F), showing approximately 9-fold and 12-fold overexpression of the gene in UKTF-3 and UKTF-4 compared to Col-0. However, there was no visual difference in the phenotype of these overexpressing plants, compared to the wild-type (Figure 3.19G).

For each of the genes *RAP2.12*, *BHLH038* and *ANL2*, two T-DNA insertional mutants were identified as shown in Figures 3.20, 3.21 and 3.22. *rap2.12-1* was found to contain two T-DNA insertions within the gene (Figures 3.20A and B), while *rap2.12-3* had only one insertion within the gene (Figures 3.20C and D). Both lines were confirmed as knockouts of *RAP2.12* (Figure 3.20E), but visually did not show any phenotypic differences with Col-0 (Figure 3.20F). Both *bhlh038-2* and *bhlh038-4* were found to contain two copies of the insertion within the gene (Figures 3.21A, B, C and D) and were

confirmed as knockouts (Figure 3.21E). Once again, there was no visual difference in the phenotype of these mutants compared with the wild-type (Figure 3.21F).

Two T-DNA insertion lines for *ANL2*, each carrying a single copy of the insertion within the gene, were identified as shown in Figure 3.22. Both *anl2-2* and *anl2-4* were verified to be knockouts of *ANL2* (Figure 3.22E). As seen in Figure 3.21, the two knockouts of *ANL2* appeared to produce slightly curled leaves, but otherwise appeared to be similar to the wild-type, particularly in terms of rosette area (data shown in Chapter 4).

Table 3.2. T-DNA insertional lines for the hub genes used in subsequent analyses.

Gene name	**Gene locus ID**	**T-DNA insertional lines**
AGL22	AT2G22540	*agl22-3* (SALK_141674) *agl22-4* (SAIL_583_C08)
POZ	AT1G55760	*poz-2* (SALK_127778c) *poz-3* (SALK_075267c)
FD	AT4G35900	*fd-5* (SALK_150991)
UKTF	AT1G16750	UKTF-3 (SALK_144830c) UKTF-4 (SAIL_327_D01)
RAP2.12	AT1G53910	*rap2.12-1* (GK-503A11) *rap2.12-3* (GK-137C12)
BHLH038	AT3G56970	*bhlh038-2* (SALK_108159c) *bhlh038-4* (SAIL_447_H01)
ANL2	AT4G00730	*anl2-2* (SALK_000196c) *anl2-4* (SAIL_418_C10)

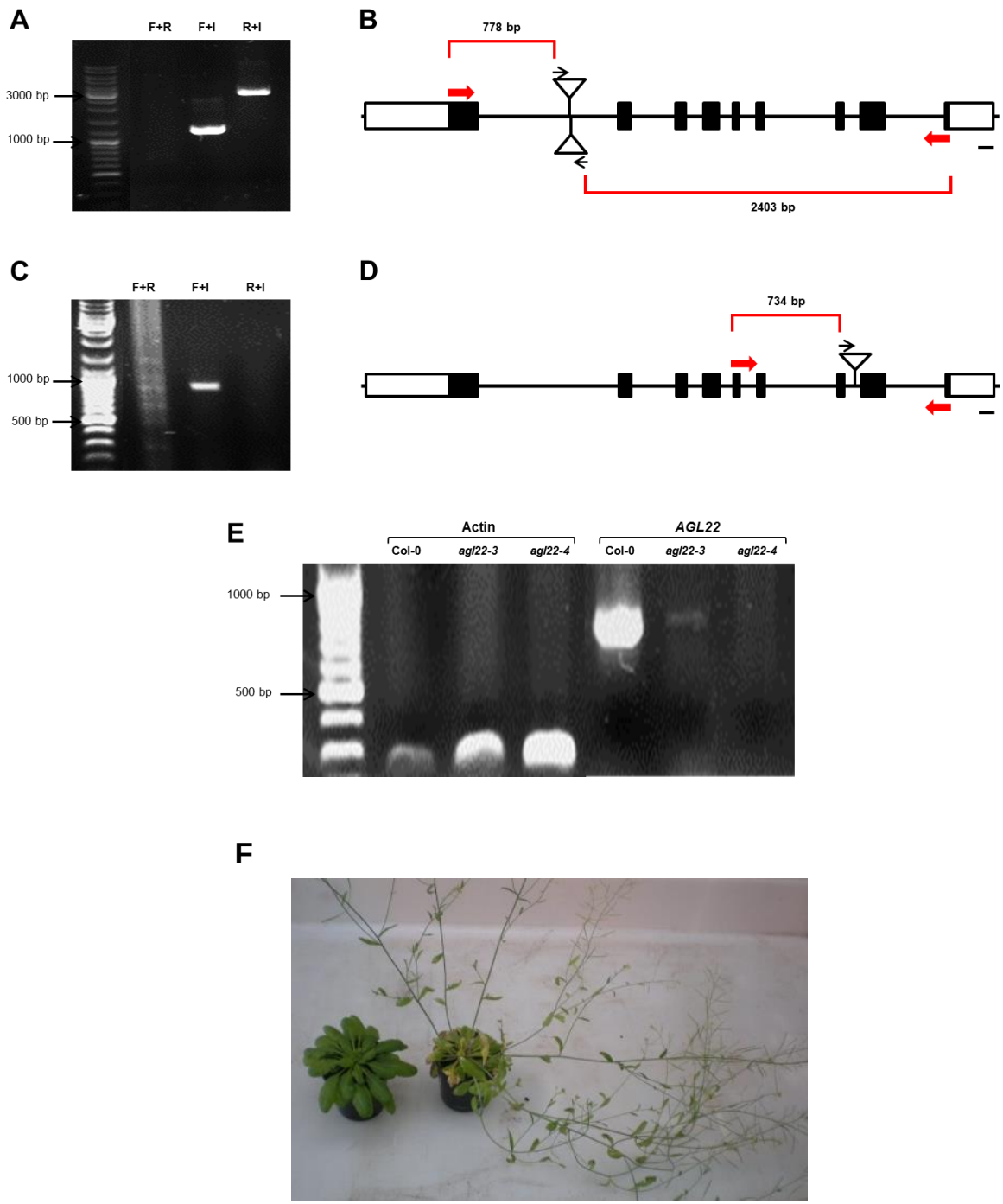


Figure. 3.16. Identification and phenotype of T-DNA insertion lines for *AGL22*.
(A) PCR showing *agl22-3* is homozygous for the presence of T-DNA insertions in *AGL22*. (B) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for *agl22-3*. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) PCR showing *agl22-4* is homozygous for the presence of a T-DNA insertion in *AGL22*. (D) Schematic diagram showing the position of the T-DNA insertion and the primer-binding sites in the gene for *agl22-4*. (E) Results of the RT-PCR for both lines showing the absence of a full-length transcript for *agl22-3* and *agl22-4* compared to the control gene, Actin. (F) Image showing the early-flowering phenotype of *agl22-3* (right) compared to a wild-type Col-0 (left).

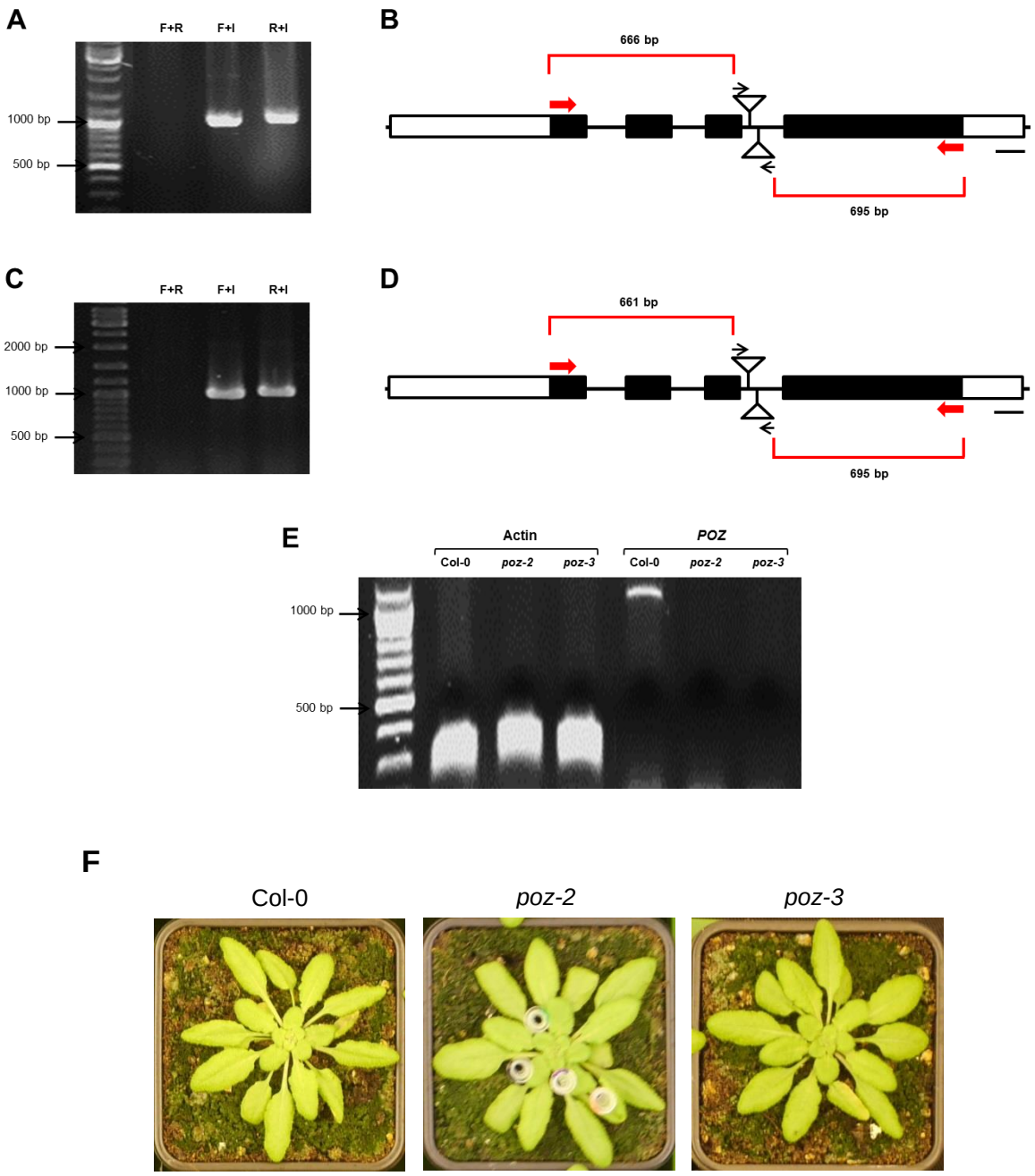


Figure. 3.17. Identification of T-DNA insertion lines for *POZ*.
(A) PCR showing *poz-2* is homozygous for the presence of T-DNA insertions in *POZ*. (B) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for *poz-2*. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) PCR showing *poz-3* is homozygous for the presence of T-DNA insertions in *POZ*. (D) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for *poz-3*. (E) Results of the RT-PCR for both lines showing the absence of a full-length transcript for *poz-2* and *poz-3* compared to the control gene, Actin. (F) Images of *poz-2* and *poz-3* compared to the wild-type Col-0.

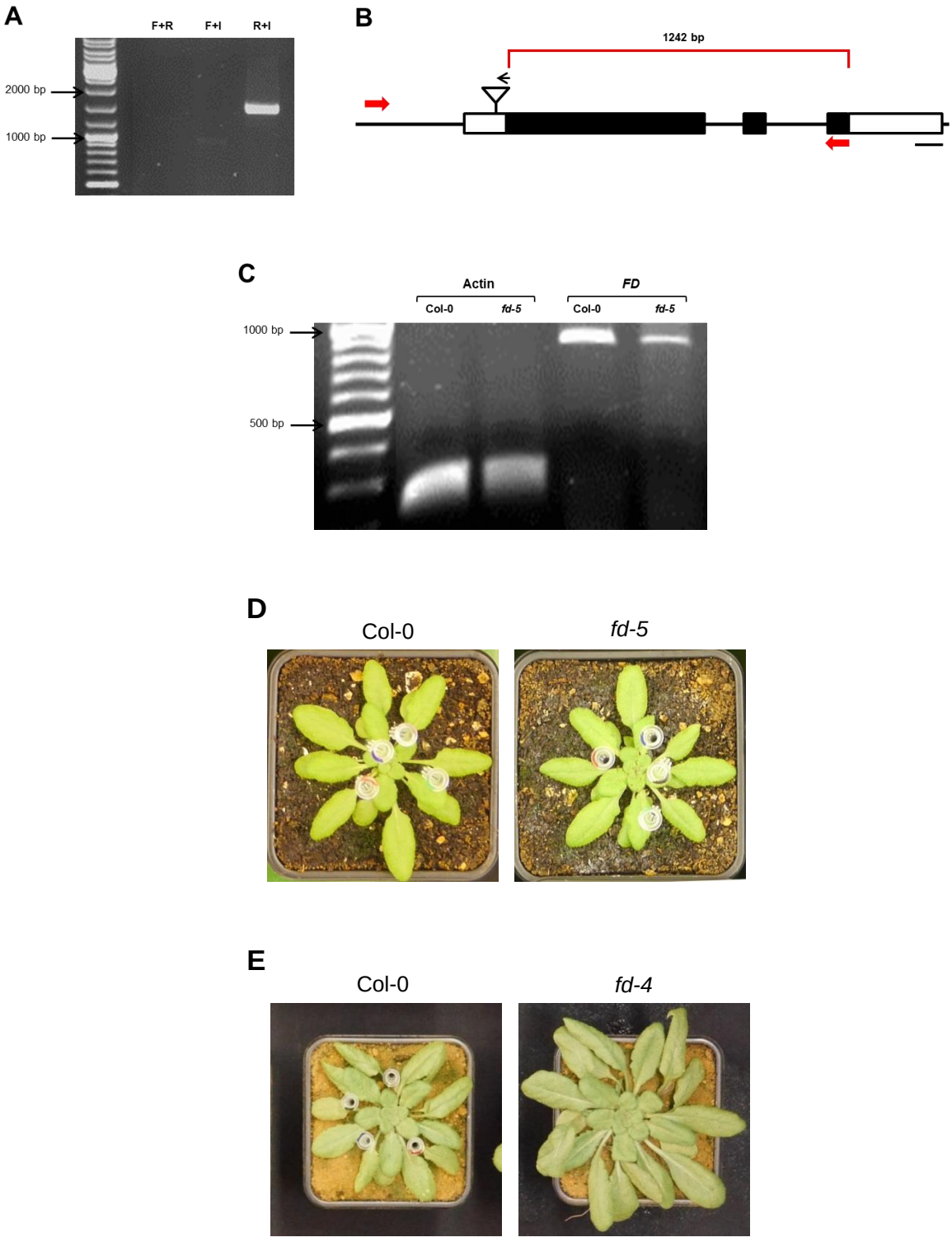


Figure. 3.18. Identification of a T-DNA insertion line for *FD*.
(A) PCR showing *fd-5* is homozygous for the presence of a T-DNA insertion in *FD*. (B) Schematic diagram showing the position of the T-DNA insertion and the primer-binding sites in the gene for fd-5. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) Results of the RT-PCR for fd-5 showing the reduced levels of the full-length transcript compared to the control gene, Actin. Images of *fd-5* (D) and *fd-4* (E) compared to the wild-type Col-0. The knockout *fd-4* has a larger rosette than Col-0, while no visual difference is seen between *fd-5* and Col-0.

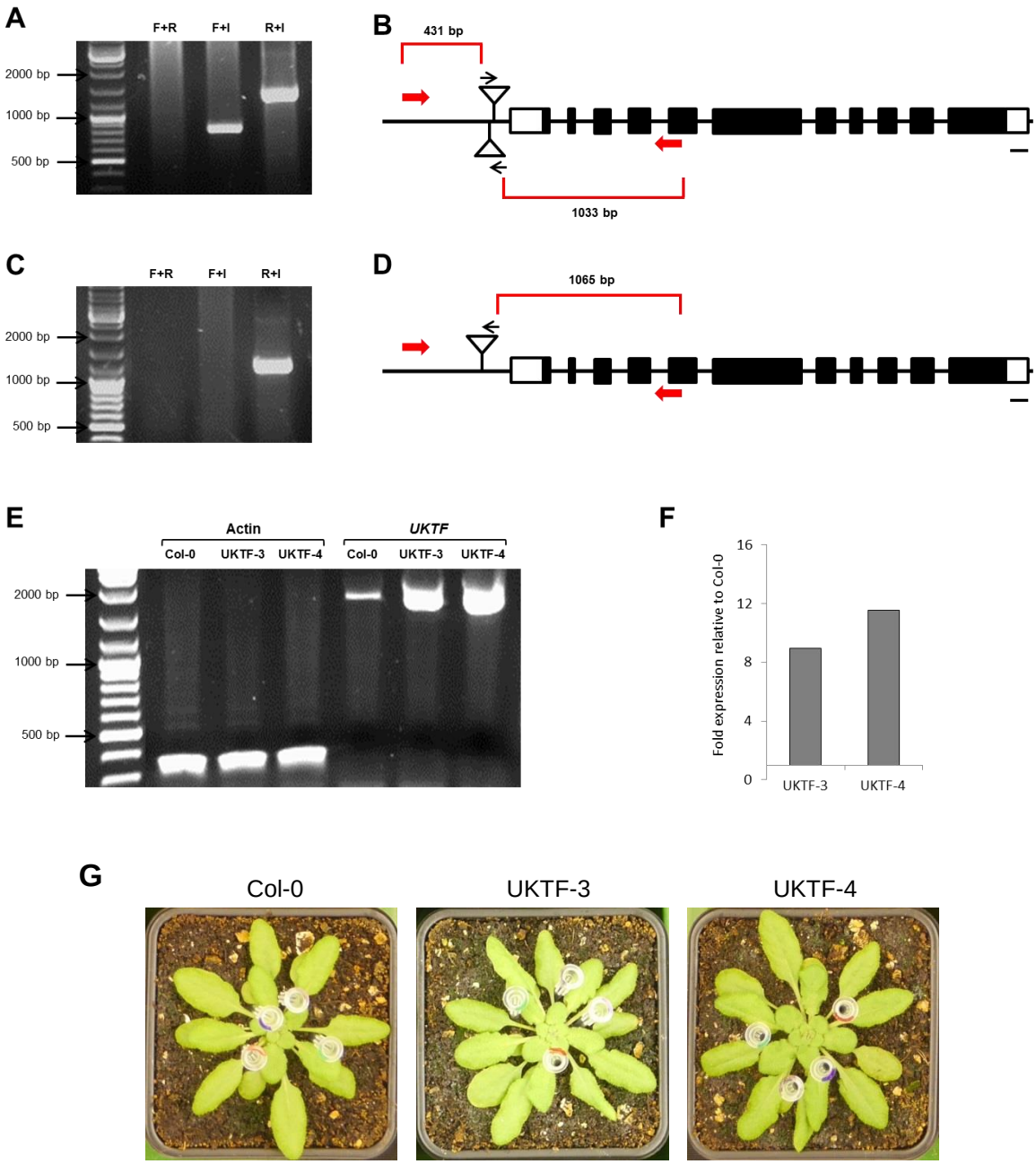


Figure. 3.19. Identification of T-DNA insertion lines for *UKTF*.
(A) PCR showing UKTF-3 is homozygous for the presence of T-DNA insertions in *UKTF*. (B) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for UKTF-3. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) PCR showing UKTF-4 is homozygous for the presence of a T-DNA insertion in *UKTF*. (D) Schematic diagram showing the position of the T-DNA insertion and the primer-binding sites in the gene for UKTF-4. (E) Results of the RT-PCR for both lines showing the absence of a full-length transcript for UKTF-3 and UKTF-4 compared to the control gene, Actin. (F) qPCR results confirming the overexpression of *UKTF* in UKTF-3 and UKTF-4. (G) Images of UKTF-3 and UKTF-4 compared to the wild-type Col-0.

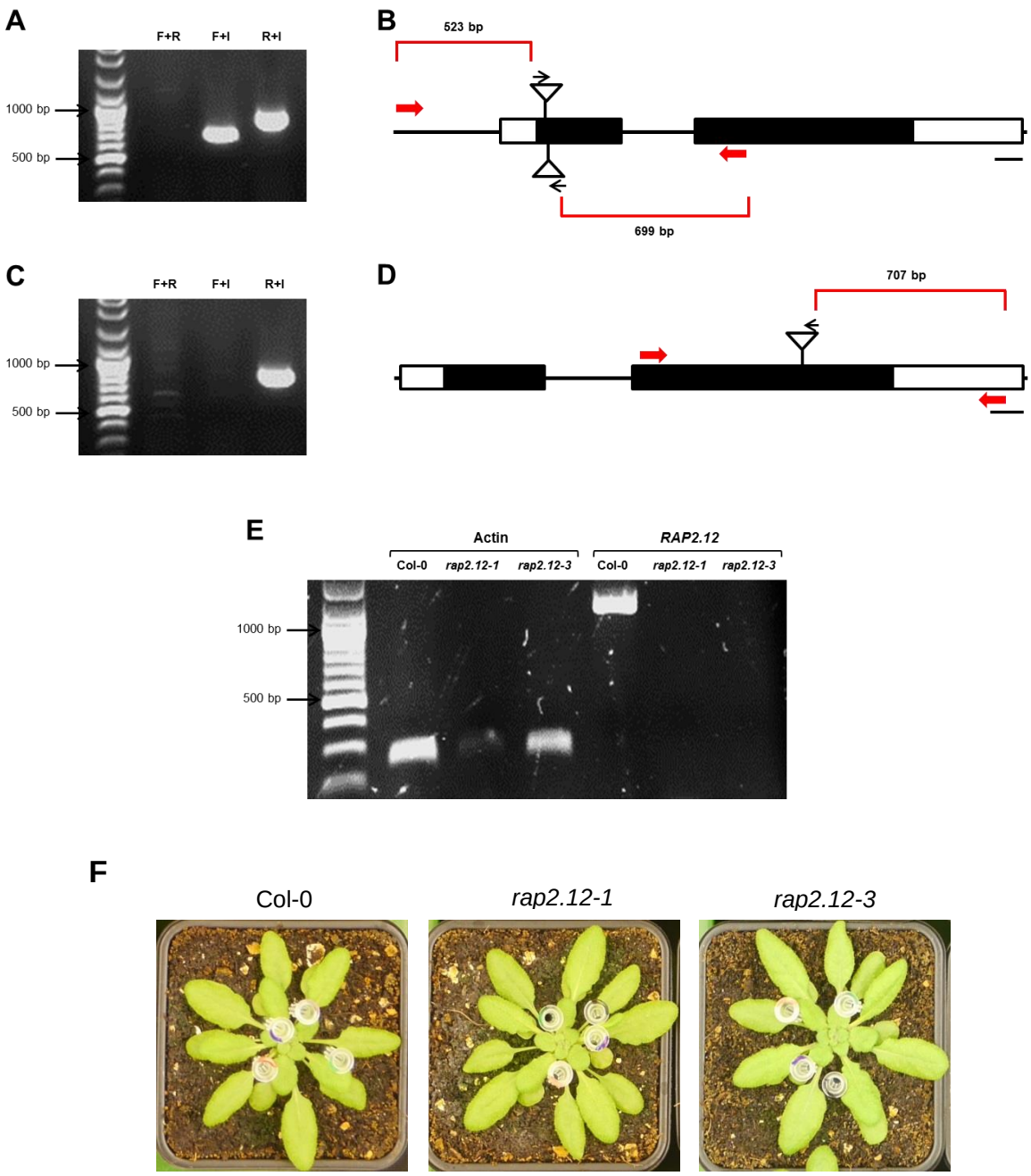


Figure. 3.20. Identification of T-DNA insertion lines for *RAP2.12*.
(A) PCR showing *rap2.12-1* is homozygous for the presence of T-DNA insertions in *RAP2.12*. (B) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for *rap2.12-1*. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) PCR showing *rap2.12-3* is homozygous for the presence of a T-DNA insertion in *RAP2.12*. (D) Schematic diagram showing the position of the T-DNA insertion and the primer-binding sites in the gene for *rap2.12-3.* (E) Results of the RT-PCR for both lines showing the absence of a full-length transcript for *rap2.12-1* and *rap2.12-3* compared to the control gene, Actin. (F) Images of *rap2.12-1* and *rap2.12-3* compared to the wild-type Col-0.

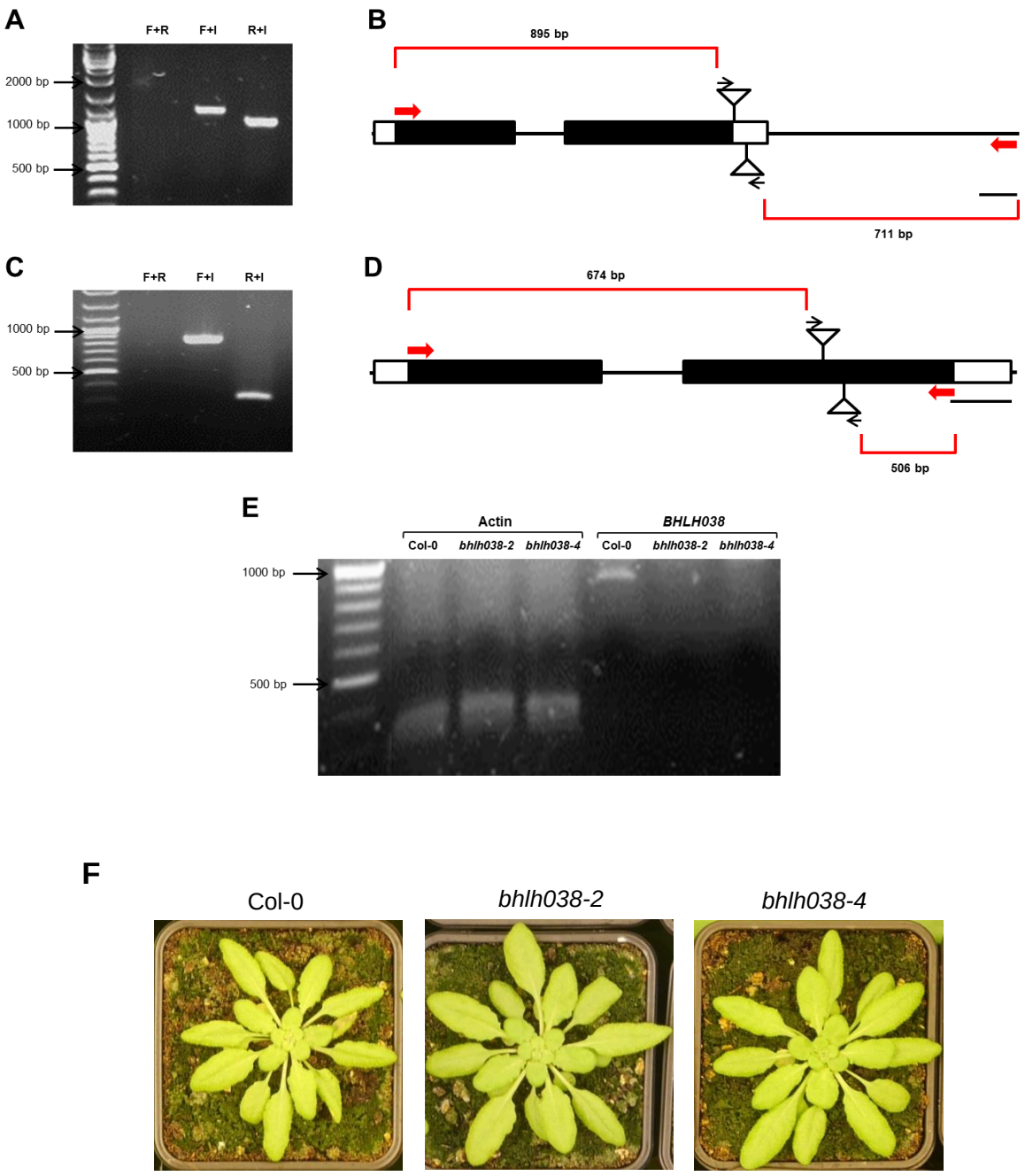


Figure. 3.21. Identification of T-DNA insertion lines for *BHLH038*.
(A) PCR showing *bhlh038-2* is homozygous for the presence of T-DNA insertions in *BHLH038*. (B) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for *bhlh038-2*. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) PCR showing bhlh038-4 is homozygous for the presence of T-DNA insertions in *BHLH038*. (D) Schematic diagram showing the position of the T-DNA insertions and the primer-binding sites in the gene for *bhlh038-4.* (E) Results of the RT-PCR for both lines showing the absence of a full-length transcript for *bhlh038-2* and *bhlh038-4* compared to the control gene, Actin. (F) Images of *bhlh038-2* and *bhlh038-4* compared to the wild-type Col-0.

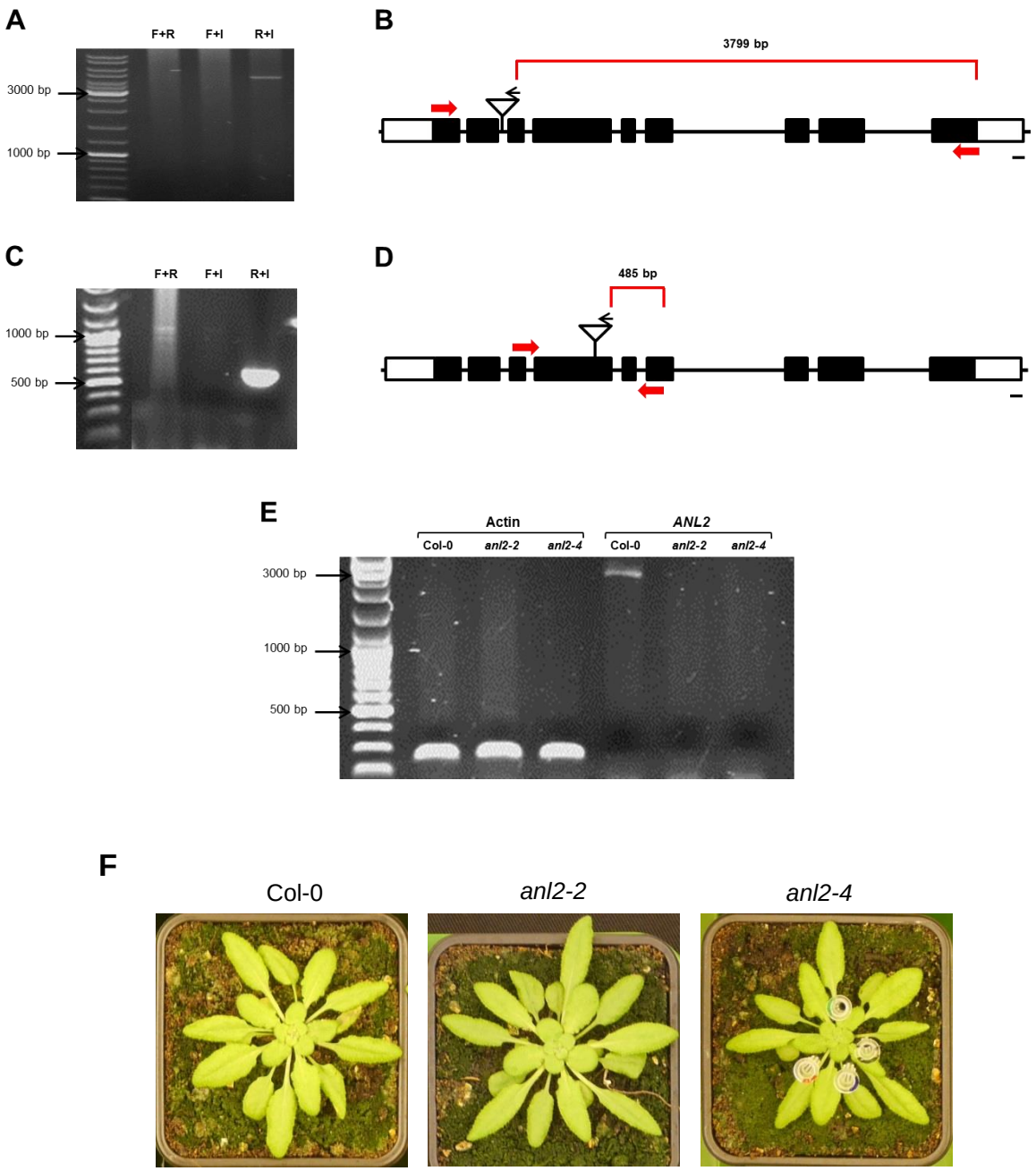


Figure. 3.22. Identification of T-DNA knockout lines for *ANL2*.
(A) PCR showing *anl2-2* is homozygous for the presence of a T-DNA insertion in *ANL2*. (B) Schematic diagram showing the position of the T-DNA insertion and the primer-binding sites in the gene for *anl2-2*. The open rectangles represent the untranslated regions (UTRs), the black rectangles represent the exons and the introns are represented by lines. The black arrows indicate the orientation of the insertions in the gene and the bar indicates 100 bases of the gene. The red arrows indicate the binding sites of the screening primers and the lengths indicate the distance between the primer and the insertion site. (C) PCR showing *anl2-4* is homozygous for the presence of a T-DNA insertion in *ANL2*. (D) Schematic diagram showing the position of the T-DNA insertion and the primer-binding sites in the gene for *anl2-4*. (E) Results of the RT-PCR for both lines showing the absence of a full-length transcript for *anl2-2* and *anl2-4* compared to the control gene, Actin. (F) Images of *anl2-2* and *anl2-4* compared to the wild-type Col-0.

3.5. Generation of overexpressing lines for the hub genes

Col-0 plants overexpressing each of the seven genes were generated using the Gateway®-compatible destination vector, pEarleyGate (Figure 3.23; Earley *et al.*, 2006). The full-length protein-coding sequence (from the translation start codon to the stop codon) of the gene of interest was amplified from Col-0 cDNA using the Phusion® Hot Start II High-Fidelity DNA polymerase (Finnzymes). The primers used to amplify the gene of interest are given in Appendix C. The forward primer contained the CACC sequence at the 5' end for directional cloning into the pENTR™/D-TOPO® vector (ThermoFisher Scientific), as per the supplier's instructions. The entry vector was then used for LR cloning of the gene of interest into the pEarleyGate vector using the LR Clonase™ II Enzyme Mix (ThermoFisher Scientific), as per the supplier's instructions. The entry and destination clones were verified by restriction digestion as described in Chapter 2.

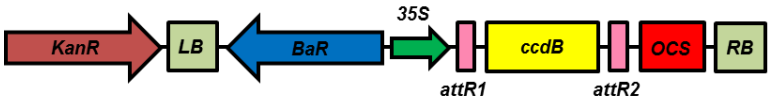


Figure. 3.23. Vector map of the destination vector pEarleyGate.

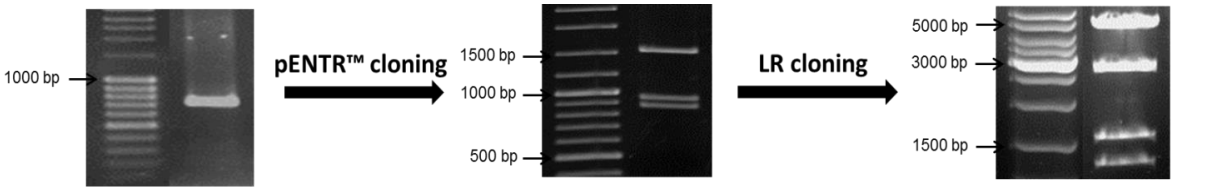
The expression vectors were transformed into Arabidopsis Col-0 plants by the floral dip method (Clough and Bent, 1998). T_1 seeds obtained from these plants were screened on soil by watering with Basta® (Bayer CropScience Ltd.) until the negative, untransformed plants appeared yellow and died out. A single leaf (approximately 100 mg of fresh weight) was taken from the positive transformants and RNA was extracted to confirm overexpression of the gene of interest using qPCR (the primer sequences used for qPCR are given in Appendix C). Figures 3.24-3.30 show the general strategy used to clone the hub genes into the pEarleyGate vector, as well as confirmation of the cloning process at each stage. The expression levels for the independent, overexpressing lines obtained for each of the genes are also shown.

Only one homozygous line each was obtained for *AGL22* and *FD* in the T_2 generation, namely AGL22ox-2 and FDox-7, and these were used for subsequent experiments

described in Chapter 4. All other available lines for these two genes and the other genes were segregating in the T_2 and T_3 generations and so were not used for phenotyping due to lack of time. Two lines of positive transformants for *RAP2.12* were identified, but it was not possible to determine the level of overexpression of the gene in these lines within the available time frame.

In addition to the overexpressors of *AGL22* generated here, another overexpressing line, henceforth known as AGL22ox-10, was obtained from Dr. Lucio Conti of the University of Milan (Masiero *et al.*, 2004; Riboni *et al.*, 2013). Both AGL22ox-10 and AGL22ox-2 plants produced the phenotype of sepaloid petals as shown in Figure 3.24F and as described in Masiero *et al.* (2004).

A



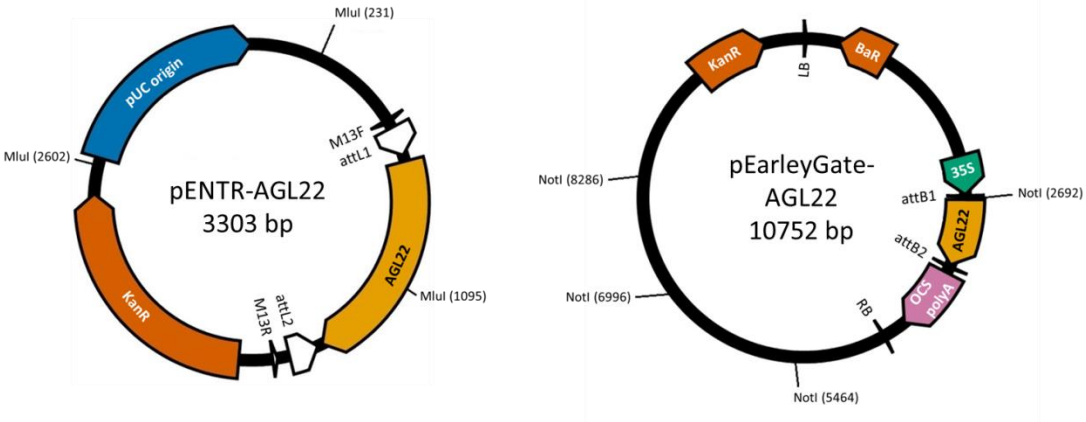
PCR with Phusion® polymerase

Digestion of pENTR™-*AGL22* plasmid

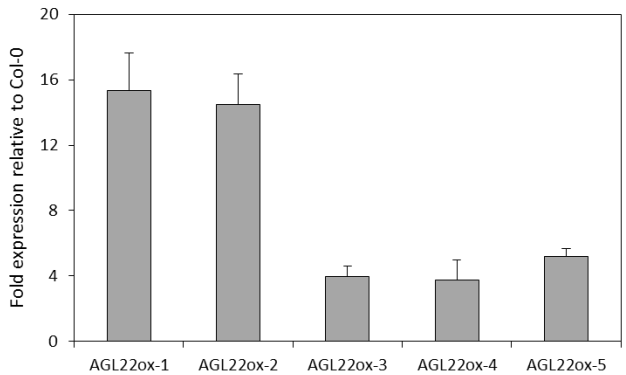
Digestion of pEarleyGate-*AGL22* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
723 bp	*MluI*	1507 bp 932 bp 846 bp	5158 bp 2772 bp 1532 bp 1290 bp

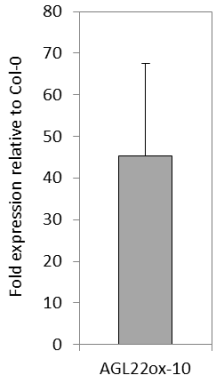
B



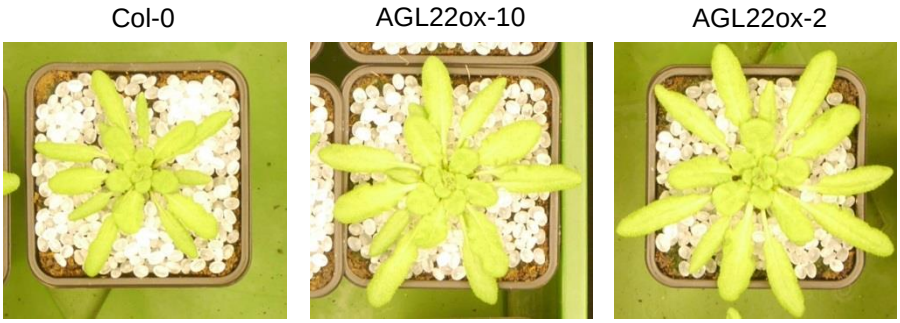
C



D



E



F

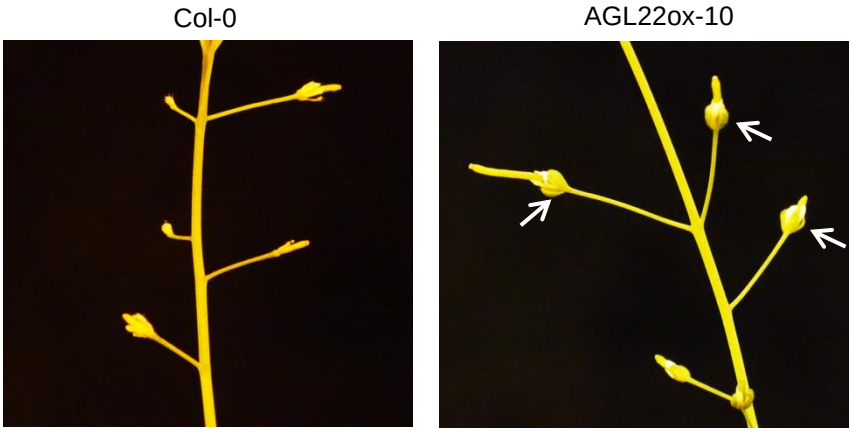
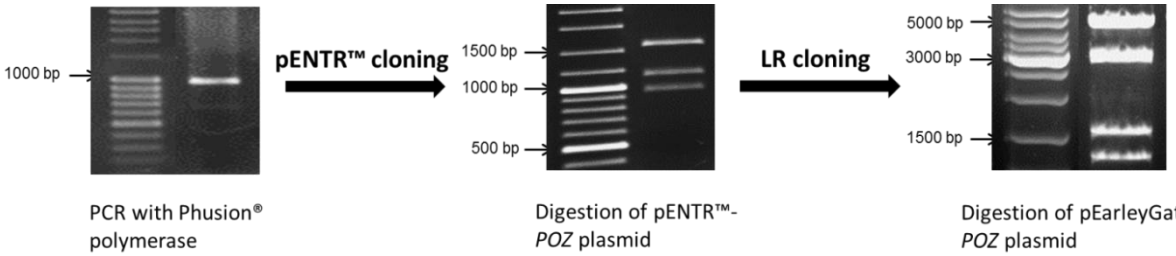


Figure 3.24. Generation of overexpressing lines of *AGL22*.
(A) Cloning strategy used to clone *AGL22* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites. (C) qPCR results measuring total transcript levels of *AGL22* in the independent plant lines relative to wild-type levels. The values are mean + standard error (n=3). (D) qPCR results measuring total transcript levels of *AGL22* in AGL22ox-10 relative to wild-type levels. The values are mean + standard error (n=3). (E) Images of AGL22ox-10 and AGL22ox-2 compared to the wild-type Col-0. Both overexpressors produced a larger rosette than the Col-0. (F) Flowers of overexpressors of *AGL22* produce sepaloid petals (indicated by arrows) compared to wild-type flowers.

A



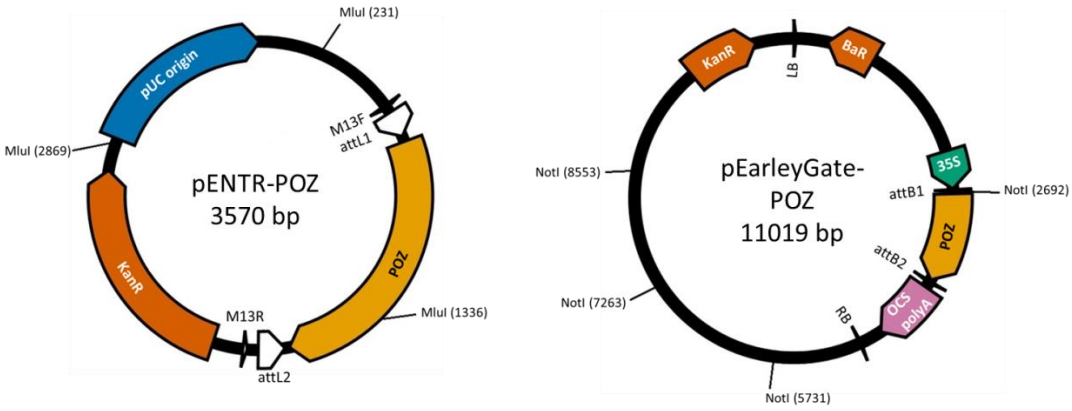
PCR with Phusion® polymerase

Digestion of pENTR™-*POZ* plasmid

Digestion of pEarleyGate-*POZ* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
990 bp	*MluI* and *HindIII*	1533 bp 1105 bp 932 bp	5158 bp 3039 bp 1532 bp 1290 bp

B



C

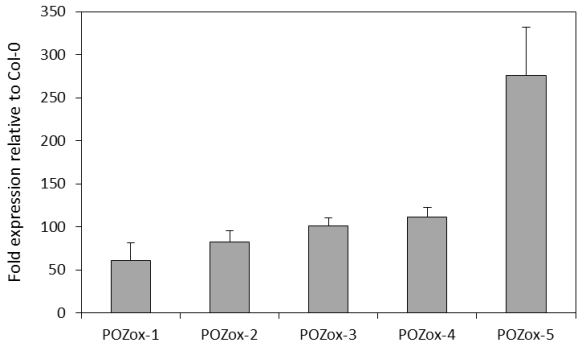
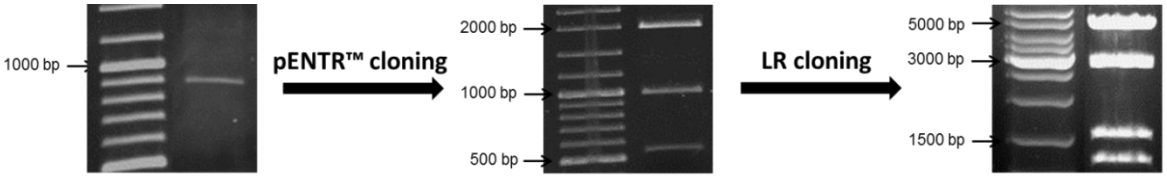


Figure 3.25. Generation of overexpressing lines of *POZ*.
(A) Cloning strategy used to clone *POZ* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites. (C) qPCR results measuring total transcript levels of *POZ* in the independent plant lines relative to wild-type levels. The values are mean + standard error (n=3).

A



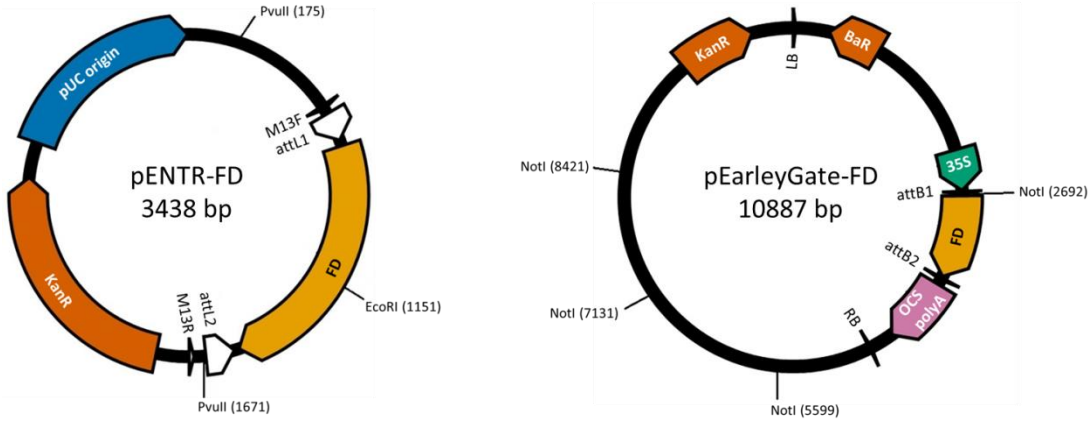
PCR with Phusion® polymerase

Digestion of pENTR™-*FD* plasmid

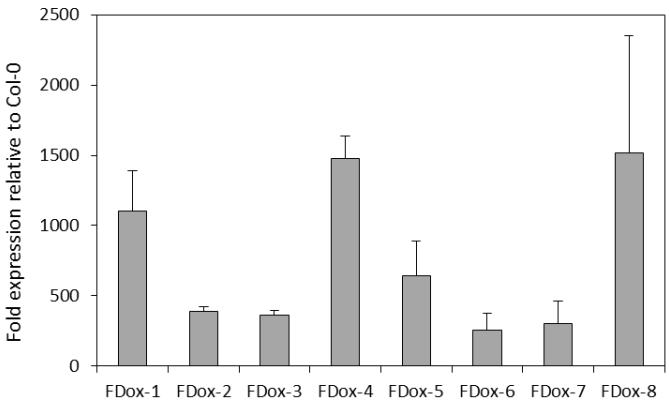
Digestion of pEarleyGate-*FD* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
858 bp	*EcoRI* and *PvuII*	1942 bp 976 bp 520 bp	5158 bp 2907 bp 1532 bp 1290 bp

B



C



D

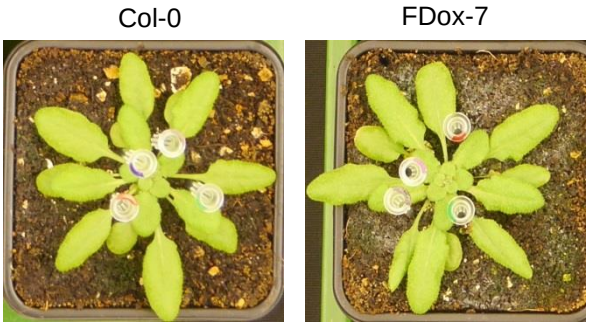
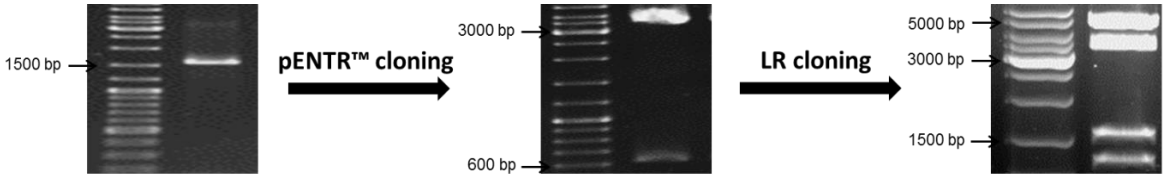


Figure 3.26. Generation of overexpressing lines of *FD*.
(A) Cloning strategy used to clone *FD* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites. (C) qPCR results measuring total transcript levels of *FD* in the independent plant lines relative to wild-type levels. The values are mean + standard error (n=3). (D) Image of FDox-7 compared to the wild-type Col-0.

A



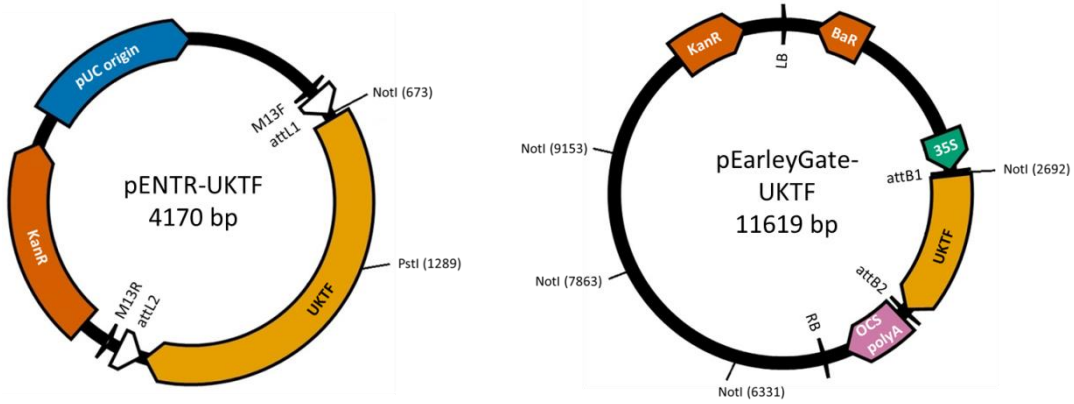
PCR with Phusion® polymerase

Digestion of pENTR™-*UKTF* plasmid

Digestion of pEarleyGate-*UKTF* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
1590 bp	*NotI* and *PstI*	3554 bp 616 bp	5158 bp 3639 bp 1532 bp 1290 bp

B



C

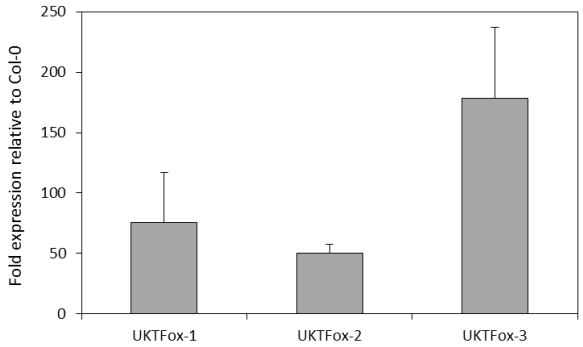
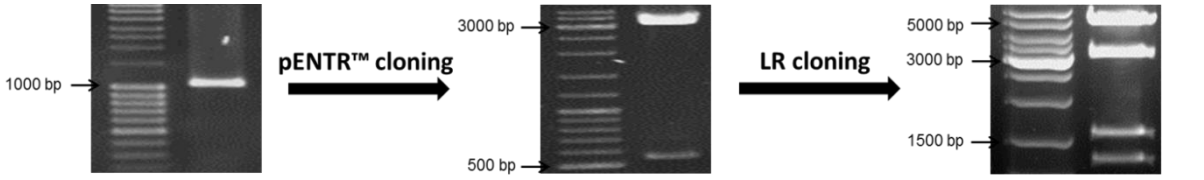


Figure 3.27. Generation of overexpressing lines of *UKTF*.
(A) Cloning strategy used to clone *UKTF* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites. (C) qPCR results measuring total transcript levels of *UKTF* in the independent plant lines relative to wild-type levels. The values are mean + standard error (n=3).

A



PCR with Phusion® polymerase

Digestion of pENTR™-*RAP2.12* plasmid

Digestion of pEarleyGate-*RAP2.12* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
1077 bp	*NotI* and *PstI*	3140 bp 517 bp	5158 bp 3126 bp 1532 bp 1290 bp

B

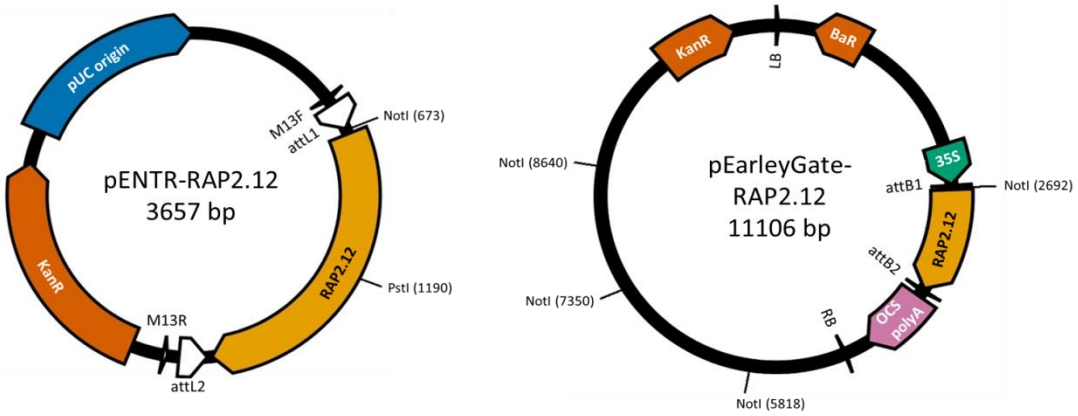
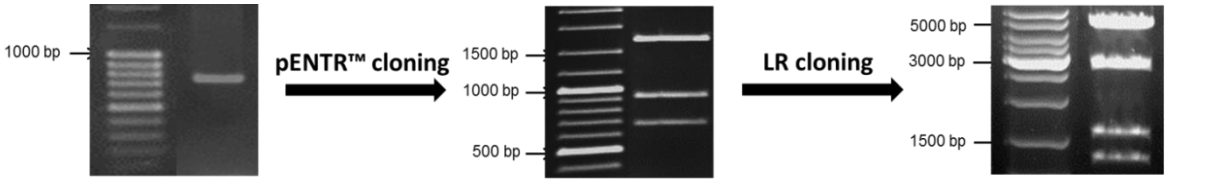


Figure 3.28. Generation of overexpressing lines of *RAP2.12*.
(A) Cloning strategy used to clone *RAP2.12* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites.

A



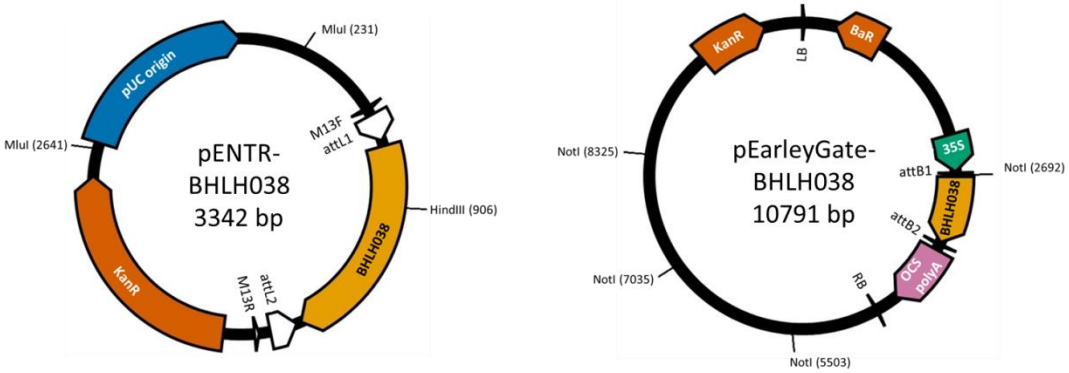
PCR with Phusion® polymerase

Digestion of pENTR™-*BHLH038* plasmid

Digestion of pEarleyGate-*BHLH038* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
762 bp	*MluI* and *HindIII*	1735 bp 932 bp 675 bp	5158 bp 2811 bp 1532 bp 1290 bp

B



C

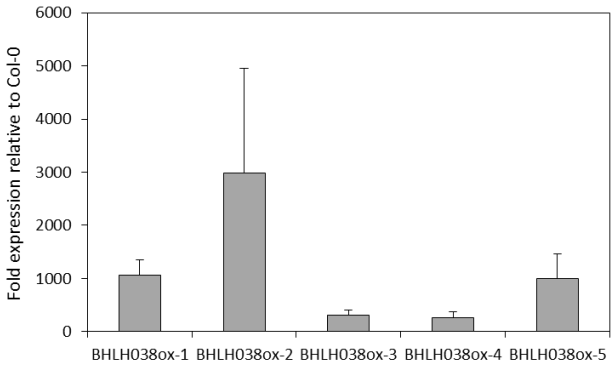
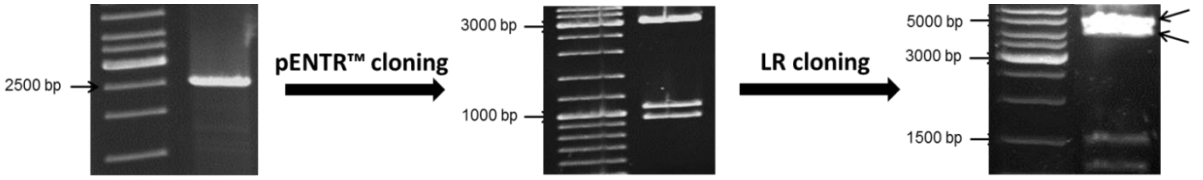


Figure 3.29. Generation of overexpressing lines of *BHLH038*.
(A) Cloning strategy used to clone *BHLH038* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites. (C) qPCR results measuring total transcript levels of *BHLH038* in the independent plant lines relative to wild-type levels. The values are mean + standard error (n=3).

A



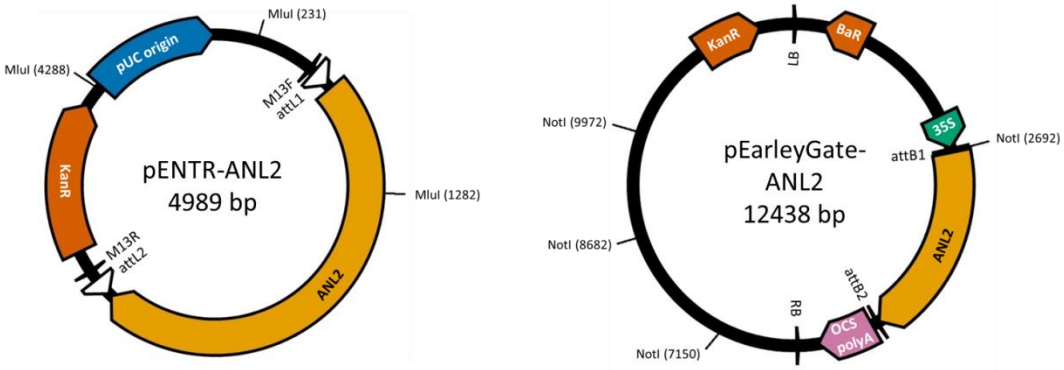
PCR with Phusion® polymerase

Digestion of pENTR™-*ANL2* plasmid

Digestion of pEarleyGate-*ANL2* plasmid

Length of coding sequence	Digestion of pENTR™ plasmid		Expected length of fragments after digestion of pEarleyGate plasmid with *NotI*
	Restriction enzymes used	Expected length of fragments	
2409 bp	*MluI*	3006 bp 1051 bp 932 bp	5158 bp 4458 bp 1532 bp 1290 bp

B



C

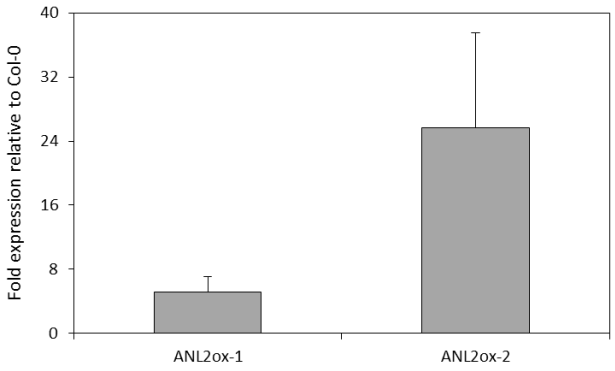


Figure 3.30. Generation of overexpressing lines of *ANL2*.
(A) Cloning strategy used to clone *ANL2* into the pEarleyGate vector using the Gateway® technology. The gel images show confirmation of the expected PCR and restriction digestion products at each stage of the cloning process, as given in the table. (B) The vector maps for the entry and destination vectors showing the restriction sites. (C) qPCR results measuring total transcript levels of *ANL2* in the independent plant lines relative to wild-type levels. The values are mean + standard error (n=3).

3.6. Characterisation of the unknown protein UKTF

UKTF is annotated in The Arabidopsis Information Resource (TAIR) to be a protein of unknown function having the domain of unknown function, DUF547 (InterPro: IPR006869). To further characterise UKTF, the protein sequence was obtained from TAIR and a protein Basic Local Alignment Search Tool (BLASTP; Altschul *et al.*, 1990) search was done using this sequence. BLASTP was accessed through the link: http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins.

As part of the result, a graphic is displayed from NCBI's Conserved Domains Database (CDD; Marchler-Bauer *et al.*, 2015) that indicates conserved domains in the query protein (Figure 3.31). UKTF has two domains, MIP1 (MADS-box interacting protein1) which contains a leucine zipper region, and DUF547. It also has the multi-domain region, Mitotic Arrest Deficient (MAD) domain, which is characteristic of the mitotic checkpoint protein, MAD1. The list of proteins generated from the BLASTP search are annotated to contain the DUF547 and MIP1 domains and all of them are plant proteins of unknown function or hypothetical proteins.

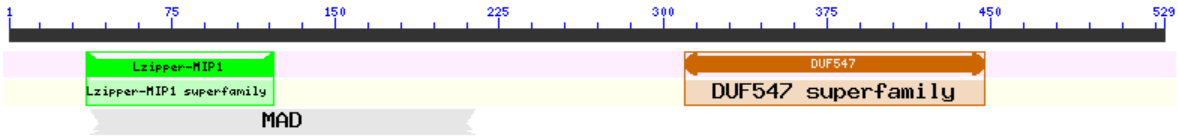


Figure 3.31. Conserved domains in UKTF determined using the Conserved Domains Database. The domains are the MADS-box interacting protein1 leucine zipper region (Lzipper-MIP1), the Domain of Unknown Function (DUF547) and the multi-domain region Mitotic Arrest Deficient (MAD). The numbers on the scale indicate amino acids.

In *A. thaliana*, there are 13 proteins including UKTF that contain both the DUF547 and MIP1 domains, and all of these proteins are uncharacterised according to TAIR. Using these protein sequences, a phylogenetic tree was created by the Maximum-Likelihood method with 1000 bootstrap replicates using MEGA 6.06 (Tamura *et al.*, 2013; Figure 3.32). It shows that all 13 proteins are closely related to UKTF (AT1G16750) and it is likely that they are a protein family, with possible functional redundancy.

MIP1 from *Antirrhinum majus* was also included in the phylogenetic tree and it can be seen that it is most similar to AT5G66600 and AT2G23700 from *A. thaliana*. Causier *et al.* (2003) identified a MIP1 protein in *Antirrhinum* with a sequence similarity of 44% and 42% to AT5G66600 and AT2G23700, respectively. They also found that this protein interacts *in vitro* with flower development C-class and E-class proteins, and thus potentially acts as a ternary complex factor with these two classes of proteins to control carpel development. MIP1 was also found to bind non-MADS-Box transcription factors and it is believed that MIP1 may act in a complex between MADS-Box proteins and transcription factors to promote transcription of target genes. UKTF has 31% similarity with the *Antirrhinum* MIP1 and so could similarly be involved in carpel development and transcriptional regulation in Arabidopsis.

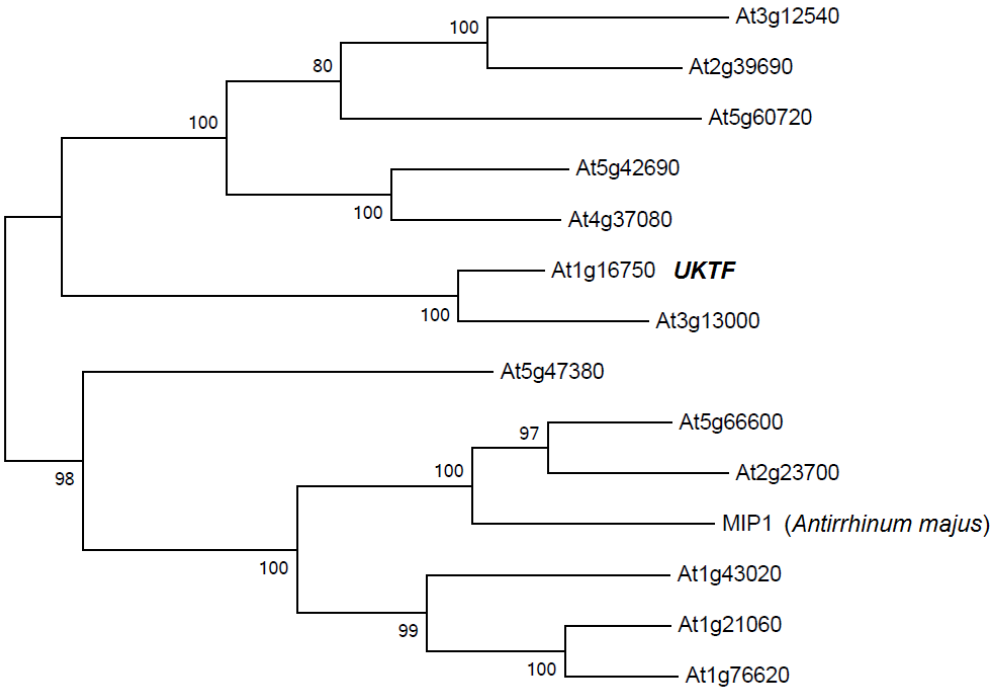


Figure 3.32. Phylogenetic tree of *A. thaliana* proteins that contain both the MIP1 and DUF547 domains within their sequence, and also the MIP1 protein from *Antirrhinum majus*. The numbers indicate node statistics. The tree was created using the Maximum-Likelihood method with 1000

None of the above 13 proteins have the MAD domain indicated in Figure 3.31 and it appears to be exclusive to UKTF in this protein family. Figure 3.33 shows the alignment of the MAD domain of UKTF with a consensus sequence of MAD1 proteins obtained from: http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=pfam05557. The consensus sequence was obtained from CDD and uses sequences from *Xenopus laevis*, *Drosophila melanogaster*, *Mus musculus*, *Arabidopsis thaliana* and *Schizosaccharomyces pombe*. The MAD domain of UKTF aligns with residues 191 to 355 of the consensus sequence, which is the N-terminal region of the 722-residue consensus sequence. The N-terminal region of the human MAD1 is involved in the homodimerisation of the protein, but otherwise has no functional role (Kim *et al.*, 2012). This region also overlaps with the leucine zipper MIP1 domain. Leucine zippers are domains that also form dimers and complexes with other proteins (Hakoshima, 2005). Thus it appears that the N-terminal region of UKTF may be involved in forming homo- and heterodimers, and is most likely involved in protein complex formation, though why a MAD domain is present only in UKTF is unclear.

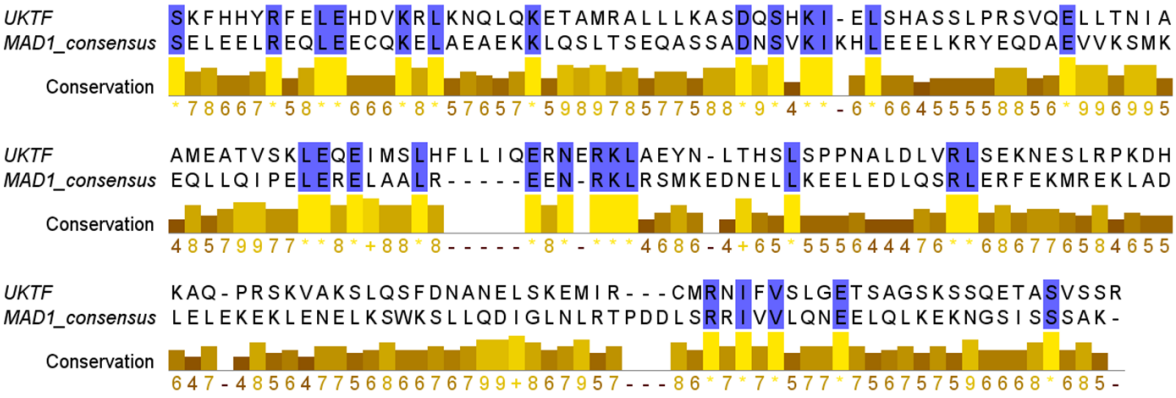


Figure 3.33. Alignment of the MAD domain of UKTF with the consensus MAD1 sequence using Jalview (Version 2.8).

To verify the subcellular localisation of UKTF in the cell, the server SUBA3 (http://suba3.plantenergy.uwa.edu.au/; Tanz *et al.*, 2013) was used. The consensus result given by SUBA3 indicated that UKTF localises to the nucleus as shown in Figure 3.34, and thus could be involved in transcriptional regulation.

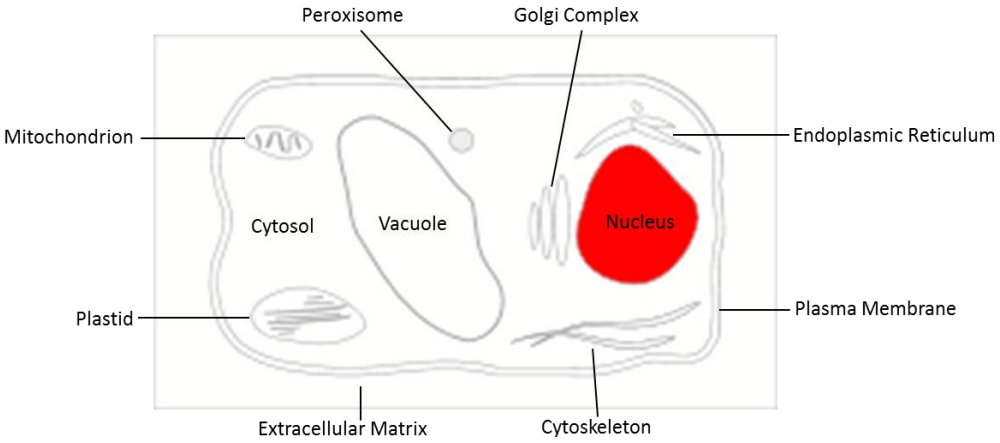


Figure 3.34. Subcellular localisation of UKTF within the nucleus as determined by SUBA3.

To try to identify the secondary structure of UKTF, the protein sequence of UKTF was submitted to Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index), which models the structure of proteins using homology-based modelling (Kelley and Sternberg, 2009). The output PDB file obtained was opened in Swiss-PdbViewer (version 4.1.0) and is shown in Figure 3.35. Only 15% of the protein was modelled, and is based on the NMR structure of the tc10 and cdc42 interacting domain2 of cip4 from humans (the template sequence). The query sequence (UKTF) had a percentage identity of 25% with the template sequence, and a confidence of 95.8%. The confidence gives the probability that the modelled region of the query and template sequences are homologous. The protein structure model presented in Figure 3.35 corresponds to residues 42 to 118 of UKTF, and this region corresponds with the leucine zipper MIP1 domain in the protein. A prediction by Phyre2 for the complete sequence and disorder prediction is shown in Figure 3.36 along with the confidence values. The whole protein disorder prediction shows that 42% of the protein is disordered.

Protein sequences submitted to Phyre2 are automatically submitted to 3DLigandSite, which predicts ligand-binding sites (Wass *et al.*, 2010). The PDB file obtained from the output was opened in Swiss-PdbViewer. Figure 3.34 shows the ligands that bind the leucine zipper domain of UKTF: 9 molecules of Heme, 15 atoms of Iron and 40 atoms of Zinc.

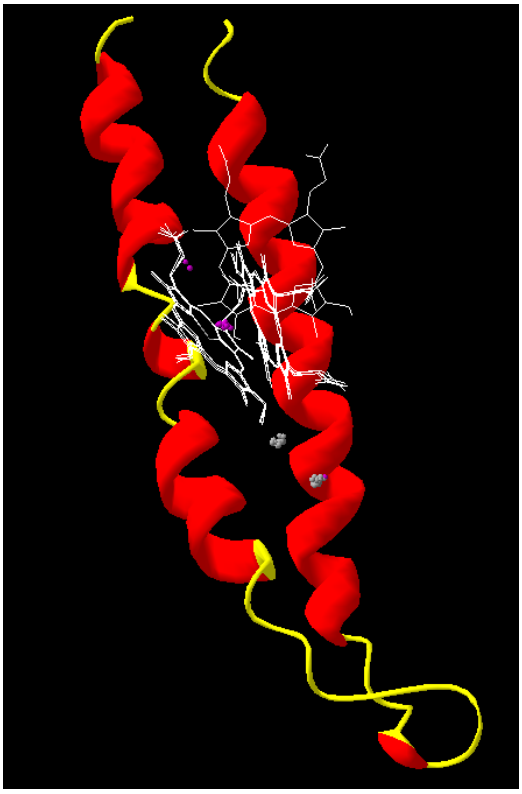
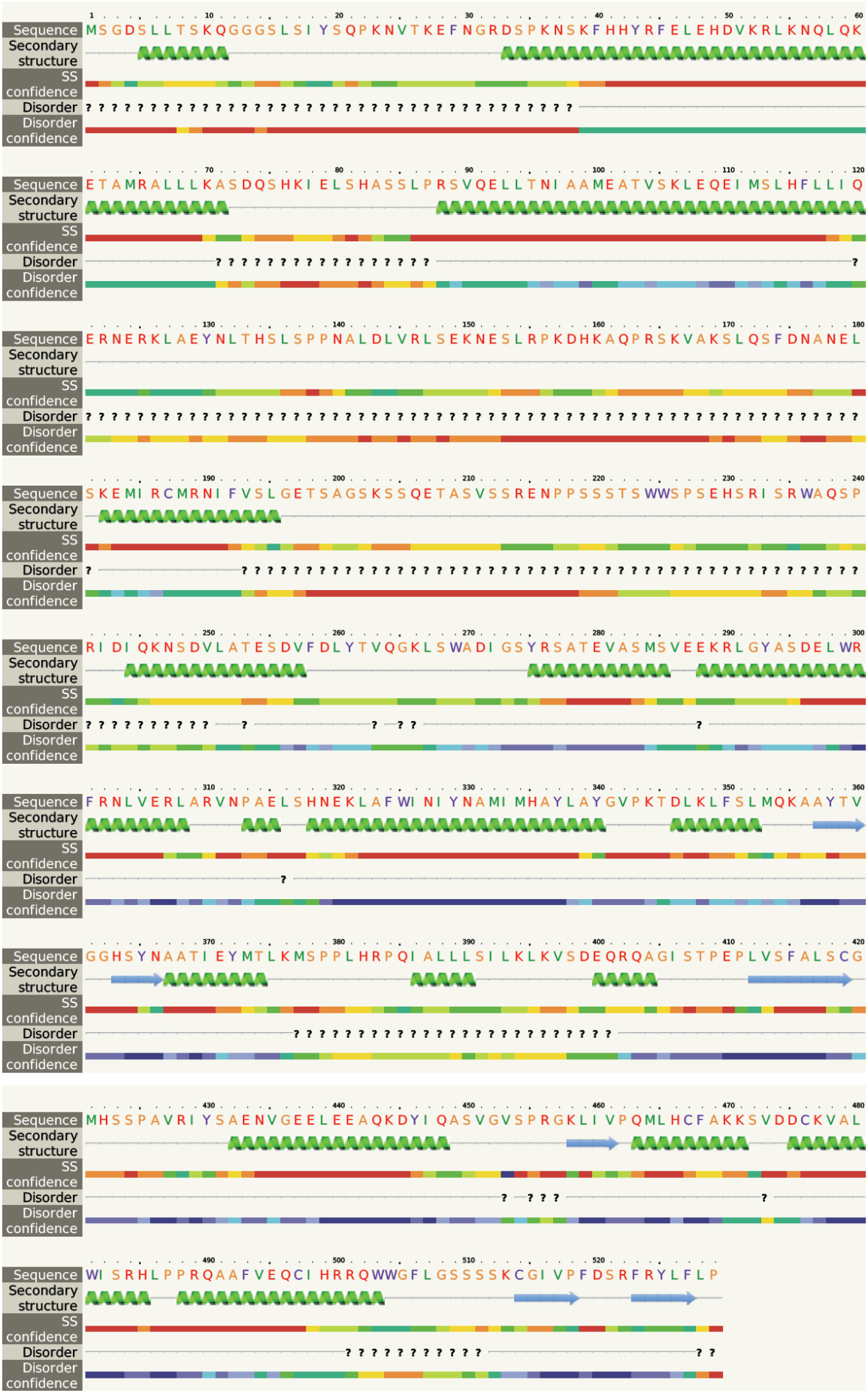


Figure 3.35. The predicted protein structure for UKTF generated using Phyre2. The alpha helices are in red and the coils are in yellow. The haem ligand is portrayed in white, with Zn ions in grey and Fe ions in purple.


Sequence
Secondary structure
SS confidence
Disorder
Disorder confidence

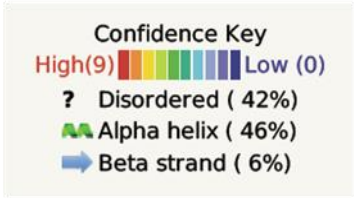


Figure 3.36. The prediction for the complete protein secondary structure and disorder for UKTF generated using Phyre2. The confidence values for the secondary structure and disorder predictions are also shown.

Only a partial structure prediction was obtained for UKTF and as it is not homologous to any known protein, the *ab-intio* protein structure prediction server, RaptorX (Källberg *et al.*, 2014; available from the link: http://raptorx.uchicago.edu/StructurePrediction/predict/) was used. RaptorX has the ability to model proteins that are not homologous to known structures. The protein structure predicted by RaptorX is shown in Figure 3.37. This prediction also indicates that a large portion of the protein is disordered.

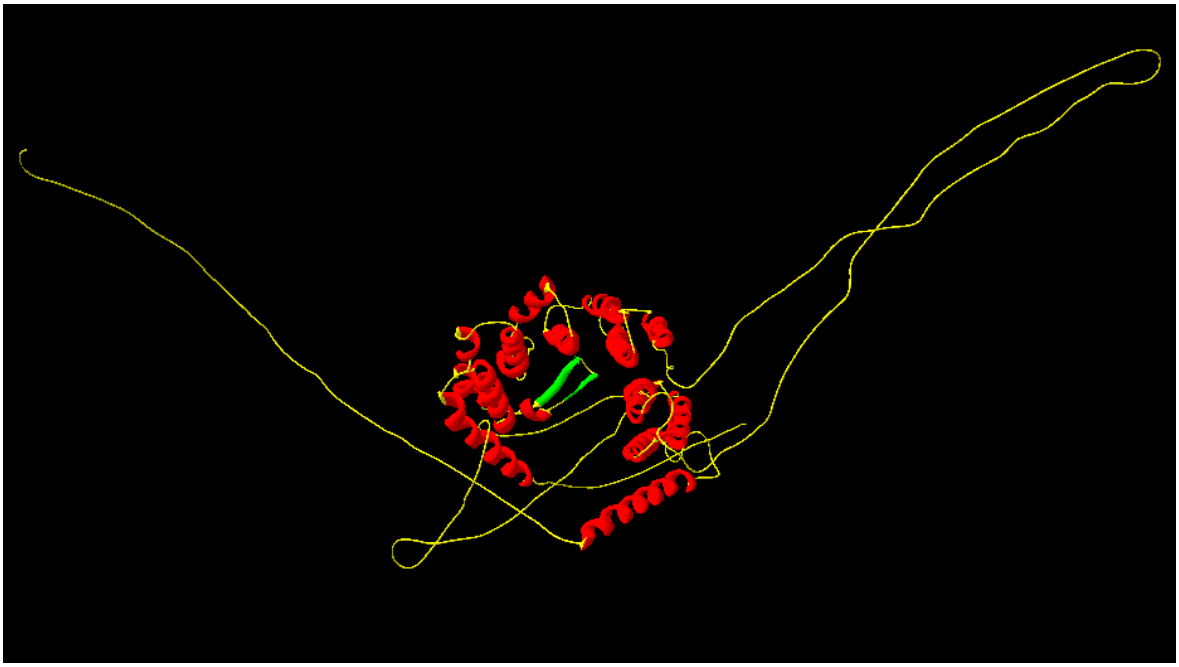


Figure 3.37. The prediction for the complete protein secondary structure for UKTF generated using RaptorX. It predicts that a large part of the protein is disordered. The alpha helices are in red, the beta sheets are in green and the coils are in yellow.

The protein sequence was then submitted to DISOPRED (http://bioinf.cs.ucl.ac.uk/psipred/) which identifies disordered regions in proteins. DISOPRED predicted that the N-terminal half of the protein was highly disordered (Figure 3.38). This is significant as disordered proteins are associated with cellular signalling and transcription regulation, and disordered regions are involved in protein binding (Pazos *et al.*, 2013). Thus, this indicates that UKTF could be similarly involved.

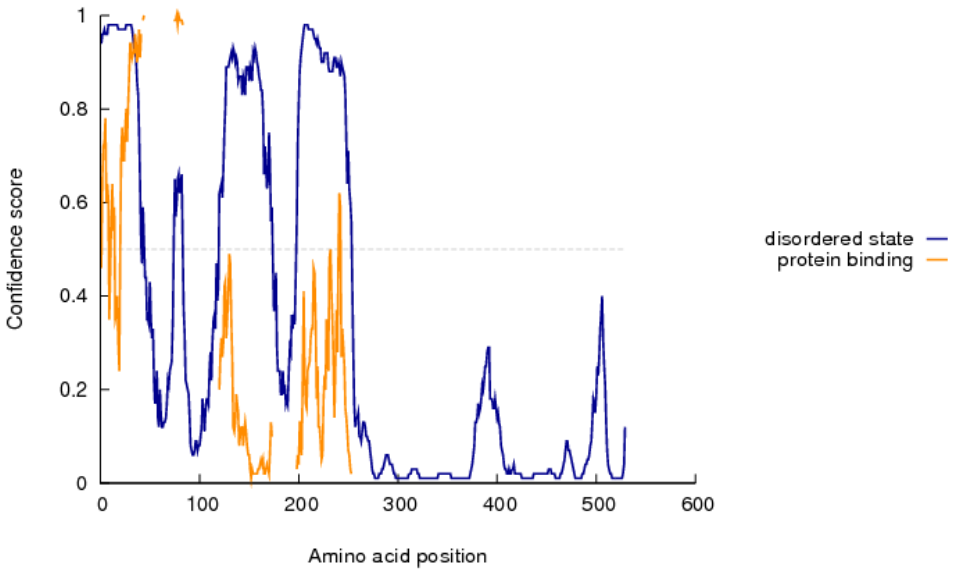


Figure 3.38. The results of DISOPRED showing that the N-terminal region of UKTF is disordered. The disordered region is more likely to be involved in protein binding.

In order to identify a possible function of UKTF, the protein sequence was submitted to CombFunc (http://www.sbg.bio.ic.ac.uk/~mwass/combfunc/; Wass *et al.*, 2012) to find GO terms associated with the Molecular Function and Biological Processes of the protein. CombFunc collates information about the query sequence from other analyses, such as BLAST, Interpro, Pfam, along with gene expression and protein-protein interaction data, and makes predictions of the function of the query protein by combining the information using the machine learning algorithm, Support Vector Machines (SVMs). The GO terms predicted by CombFunc for UKTF is shown for Molecular Function terms in Table 3.3 and for Biological Processes in Table 3.4. The results in Table 3.3 predict that UKTF is possibly involved in binding phosphatidylinositol, metal ions and Rho GTPases. Phosphatidylinositol binding implies that UKTF is involved in cellular signalling, and could possibly be involved in stress signalling (Munnik and Vermeer, 2010), while the results of Table 3.4 indicate that UKTF may be a transcription factor.

Table 3.3. CombFunc results for Molecular Functions UKTF are involved in.

GO Term	Description	SVM Probability
GO:0035091	Phosphatidylinositol binding	0.625
GO:0046872	Metal ion binding	0.503
GO:0017048	Rho GTPase binding	0.392
GO:0005100	Rho GTPase activator activity	0.346
GO:0016301	Kinase activity	0.334

Table 3.4. CombFunc results for Biological Processes UKTF are involved in.

GO Term	Description	SVM Probability
GO:0006355	Regulation of transcription, DNA-templated	0.323

As UKTF has a possible cellular signalling role, the protein sequence was analysed with MEMSAT3 (http://bioinf.cs.ucl.ac.uk/psipred/?memsatsvm=1; Jones, 2007) to identify transmembrane residues. Residues 324 to 339 of UKTF were identified as pore-lining residues by the server. To predict the orientation of the protein in the membrane, the protein sequence was analysed with MEMEMBED (Nugent and Jones, 2013) and the output of MEMEMBED is given in Figure 3.39. Only a part of the protein is shown and includes the region containing the pore-lining residues, as determined by MEMSAT3. These residues lie just after the N-terminal region of the protein containing the MIP1 domain. Since UKTF is localised to the nucleus, this result indicates that the protein may be anchored to the nuclear membrane. This feature was also seen in MIP1 from *Antirrhinum majus* and in its closest homologue in *A. thaliana*, AT5G66600 (data not shown), and thus it appears that it is a feature of this protein family.

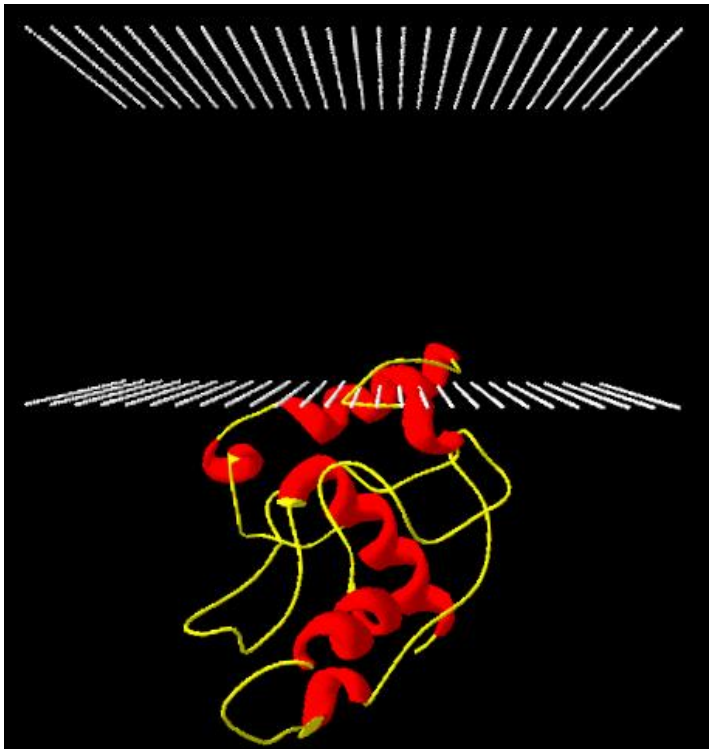


Figure 3.39. Prediction of the orientation of UKTF within the lipid bilayer as determined by MEMEMBED. The alpha helices are in red, the coils are in yellow and the lipid bilayer is shown in white. The protein shown here represents only the part of the protein that contains the pore-lining residues.

To further investigate the function of UKTF, gene expression data was obtained for *UKTF* in different tissues and under different developmental stages, using Genevestigator (Hruz *et al.*, 2008). Figure 3.40 shows the relative gene expression data for *UKTF* in the top 10 most expressing tissues, which predominantly include guard cell and reproductive tissues. This result ties in well with the notion that UKTF is involved in carpel development. Figure 3.41 shows the relative gene expression of *UKTF* throughout the development of the plant, which appears to remain constant until it decreases at senescence.

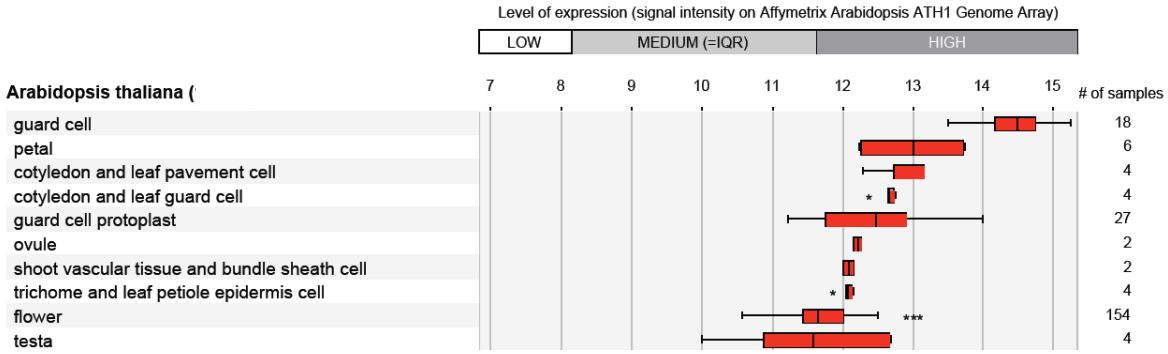


Figure 3.40. The expression of *UKTF* within the top 10 most expressing tissues (from Genevestigator). The numbers indicate $\log_2$ values.

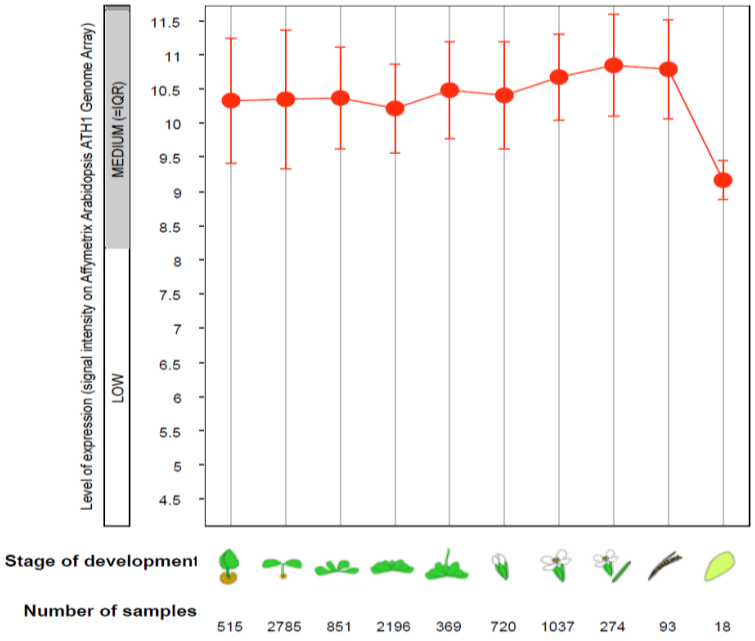


Figure 3.41. The expression of *UKTF* during plant development (from Genevestigator). The numbers indicate $\log_2$ values.

3.7. Characterisation of the unknown protein POZ

POZ is annotated in TAIR as a protein containing a Bric-a-Brac, Tramtrack, Broad-complex / Pox virus and Zinc finger (BTB/POZ) domain. Proteins containing these domains are involved in protein binding (Bardwell and Treisman, 1994; Zollman *et al.*, 1994). The protein sequence from POZ was submitted to BLASTP to identify potential homologues. The CDD graphic (Figure 3.42) obtained from the BLASTP search shows that the protein contains the BTB/POZ domain (IPR013069) and also the SPOP-C-like (C-terminal of the speckle-type POZ protein) which is a BACK-like (BTB And C-terminal Kelch) domain (IPR011705). The protein also has a hypothetical multidomain, PHA02713. Proteins containing the BTB and BACK domains are involved in forming a complex with Cullin3 to form E3 ligases for protein ubiquitination (Stogios and Privé, 2004; Gingerich *et al.*, 2005). Accordingly, CDD predicts four Cullin-binding sites near residue 280 within the BACK-like domain of POZ. The results obtained from the BLAST search also bring up a list of other BTB/POZ containing proteins.

Results obtained from CombFunc indicate that POZ possibly binds an RNA polymerase II transcription factor and also binds with E3 ubiquitin ligase enzymes, and may possibly have kinase activity (Table 3.5). Table 3.6 shows that in addition to protein ubiquitination, POZ may also be involved in response to salt stress and developmental processes. Using SUBA3, the subcellular localisation of POZ was predicted to be predominantly in the cytoplasm (probability value of 0.799) with a lower probability prediction (of 0.199) of the protein being in the mitochondria (Figure 3.43). Althogether, these results indicate that POZ is likely to be involved in protein ubiquitination by acting as a scaffold protein between E3 ligases and target proteins (Gingerich *et al.*, 2005).

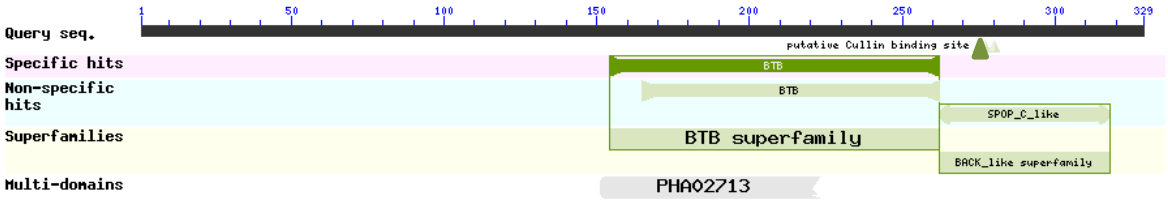


Figure 3.42. Conserved domains in POZ determined from NCBI's CDD. The domains are BTB and SPOP_C_like, and the multi-domain PHA02713. Putative Cullin binding sites are indicated by the triangle. The numbers on the scale indicate amino acids.

Table 3.5. CombFunc results for Molecular Functions POZ are involved in.

GO Term	Description	SVM Probability
GO:0001085	RNA polymerase II transcription factor binding	0.341
GO:0031625	Ubiquitin protein ligase binding	0.339
GO:0016301	Kinase activity	0.264

Table 3.6. CombFunc results for Biological Processes POZ are involved in.

GO Term	Description	SVM Probability
GO:0016567	Protein ubiquitination	0.380
GO:0071472	Cellular response to salt stress	0.318
GO:0032502	Developmental process	0.310
GO:0032501	Multicellular organismal process	0.310

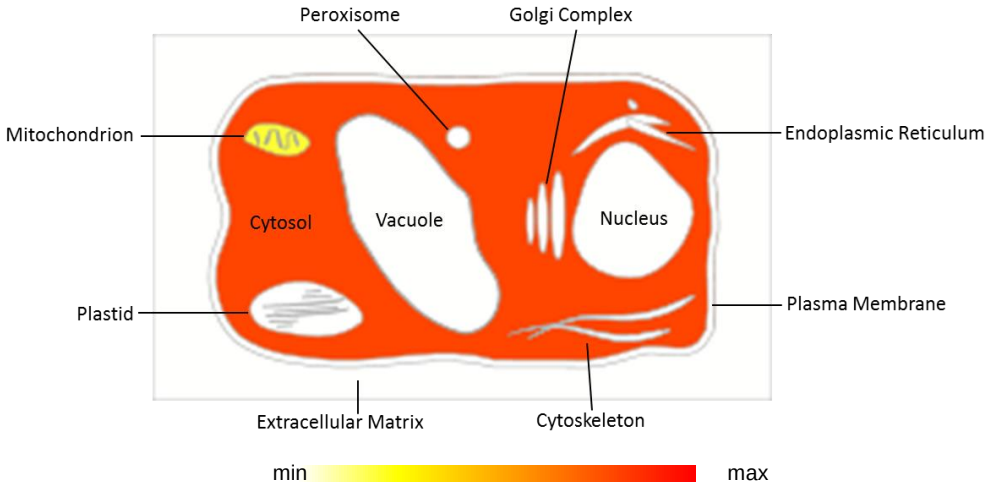


Figure 3.43. Subcellular localisation of POZ within the cytoplasm as determined by SUBA3. The heat map indicates that the protein is more likely to be present in the cytoplasm, with a reduced probability of being localised within the mitochondria.

There are 85 BTB-containing proteins in *A. thaliana* and these were used to create a phylogenetic tree, (Figure 3.44). The POZ protein (AT1G55760.1) clusters on the same branch as AT1G21780.1, which has been shown experimentally to form Cullin3-based E3 ligases (Gingerich *et al.*, 2005). It is also seen from Figure 3.43 that POZ is closely related to the 6 BTB-MATH (Meprin And Traf Homology) domain-containing proteins (BPMs; open circles) in Arabidopsis, despite not having a MATH domain. The BTB-MATH proteins have been shown to bind AP2/ERF-containing proteins and target them for degradation (Weber and Hellman, 2009; Chen *et al.*, 2013).

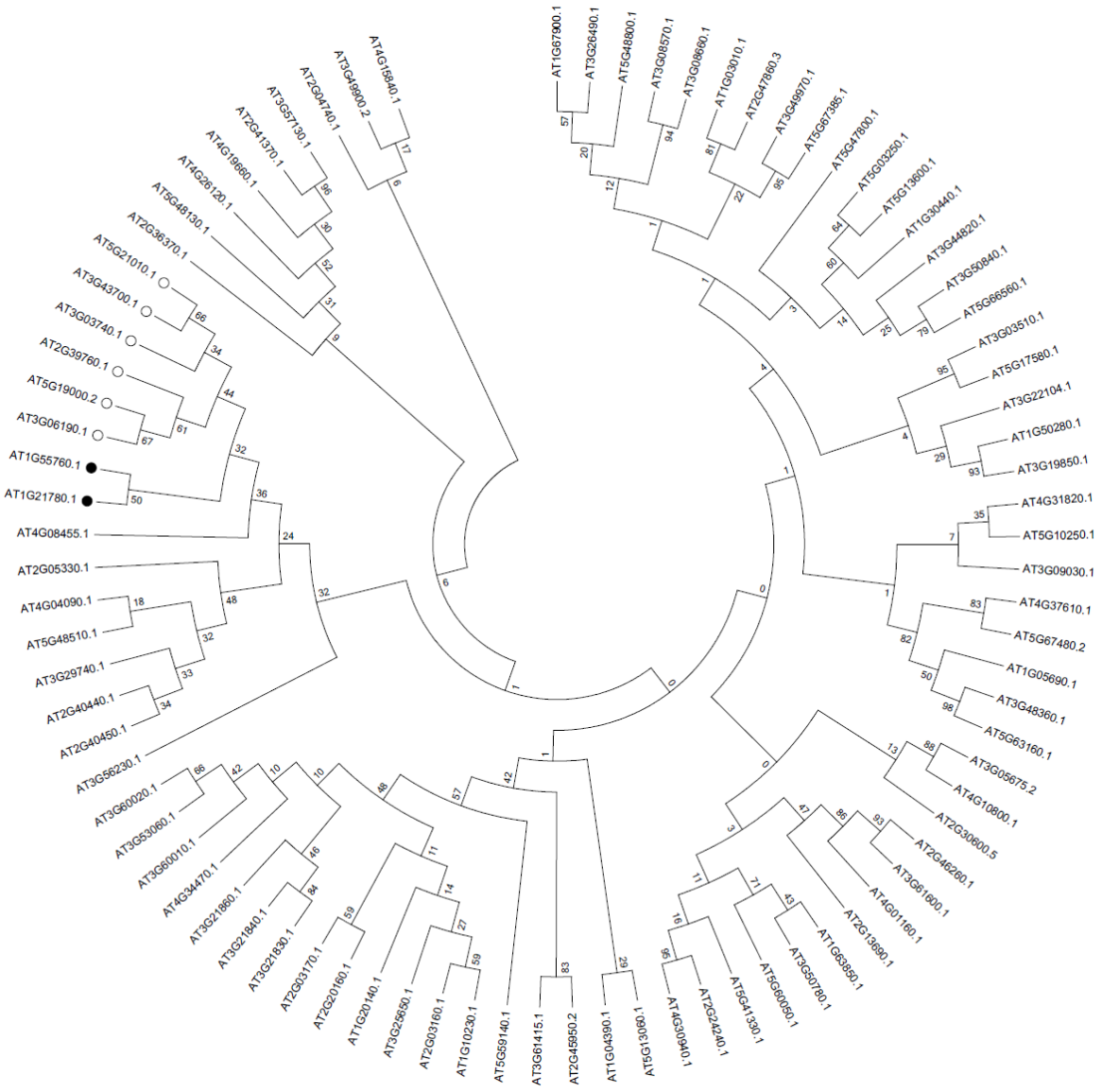


Figure 3.44. Phylogenetic tree for BTB-containing proteins in *A. thaliana*. POZ and AT1G21780.1 are labelled with black closed circles, while the 6 BTB-MATH domain-containing genes are labelled with open circles.

In order to identifiy if POZ has a MATH domain, the protein sequence of POZ was submitted to Phyre2 to model its 3D structure (Figure 3.45) – MATH domains have an 8-strand beta barrel (Sunnerhagen *et al.*, 2002). Using Phyre2, POZ was modelled on a human speckle-type POZ protein with 81% coverage and 22% protein identity, and a confidence of 100% between the protein sequences. The N-terminal region of the protein is predicted to have a beta barrel, while the C-terminal region is dominated by alpha helices. According to the output of 3DLigandSite, the protein also binds five Zn ions. The complete prediction of the structure of POZ is shown in Figure 3.46, along with putative Cullin binding sites and residues predicted to be involved in dimerisation of POZ. The beta barrel predicted at the N-terminal region of POZ, shown in Figures 3.45 and 3.46, is similar to that of MATH domains.

To check the similarity of POZ with BPMs, a structural alignment was done using SALIGN (Braberg *et al.*, 2012) with the predicted 3D structure of POZ and one of the BPMs, BPM1 (AT5G19000), as shown in Figure 3.47. The output files obtained from the alignment was viewed using Chimera (Petterson *et al.*, 2004). This was compared against the alignment of POZ with a BTB-only-containing protein, AT2G40440 (Figure 3.48). It can be seen in Figure 3.47 that the beta barrel region of POZ (in brown) aligns with the MATH domain of BPM1 (in blue), while in Figure 3.48 there is no structure in AT2G40440 (in green) that aligns with the beta barrel region of POZ. This indicates that POZ may have a MATH domain in its structure, and thus could be involved in targeting AP2/ERF proteins for degradation.

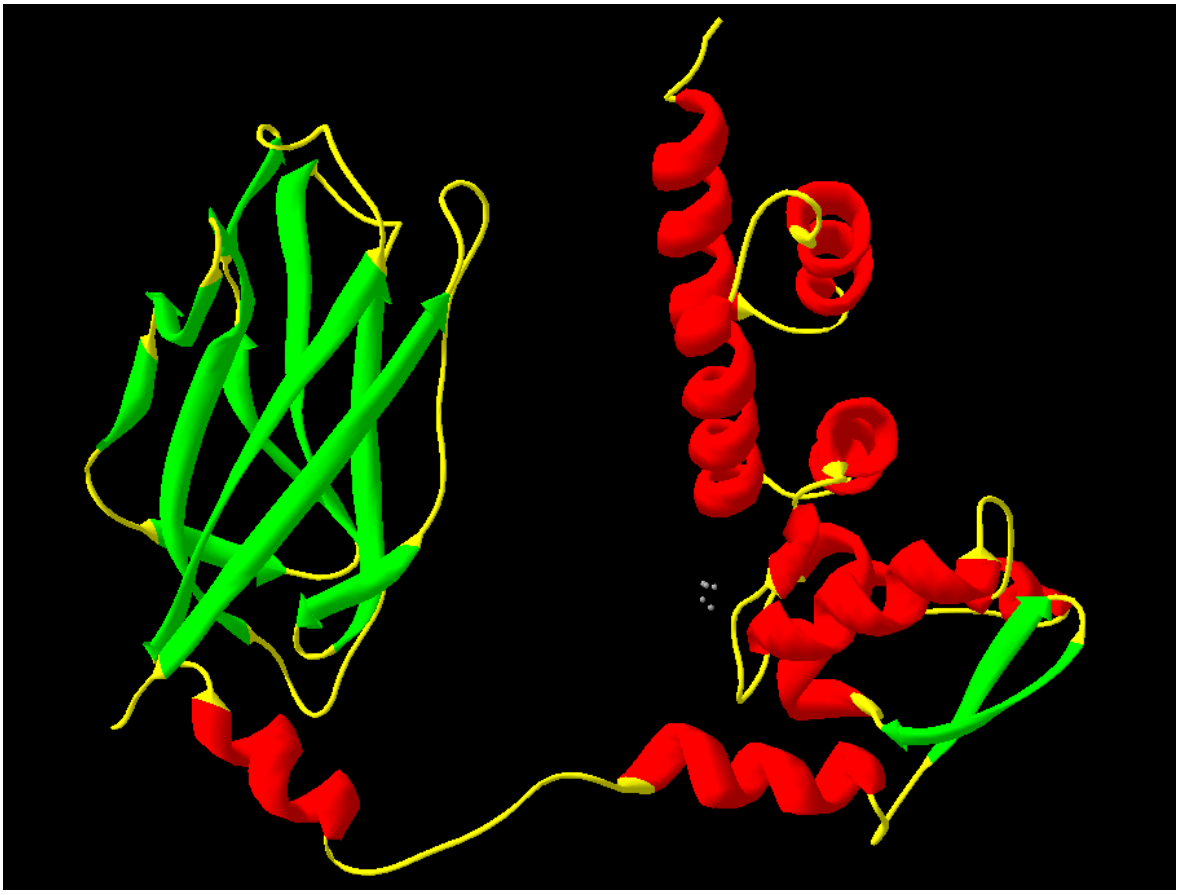


Figure 3.45. The predicted protein structure for POZ generated using Phyre2. The alpha helices are in red, the beta sheets are in green and the coils are in yellow. The Zn ions are shown in grey.

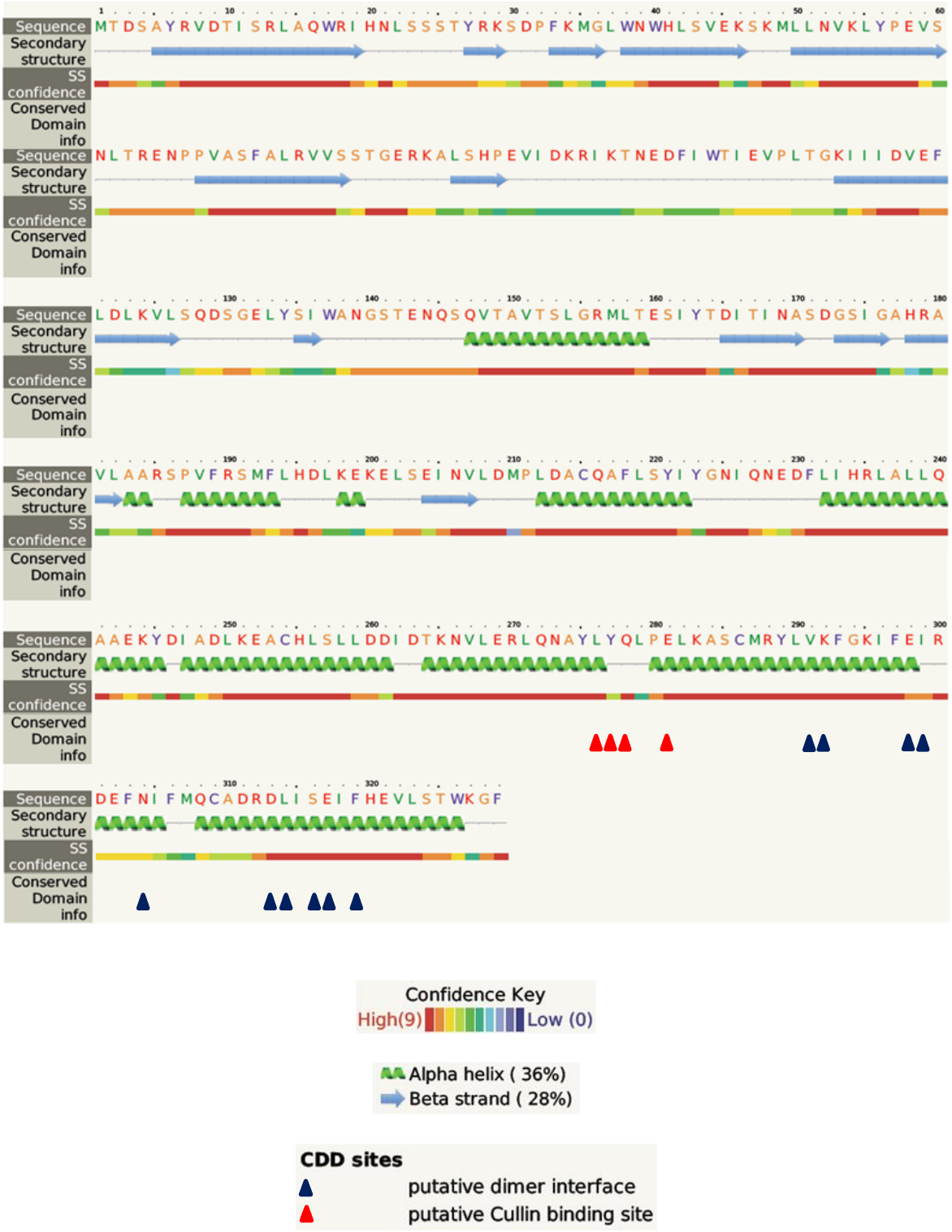


Figure 3.46. The complete protein structure prediction for POZ generated using Phyre2. The confidence values are shown along with putative binding sites for Cullin and putative sites for dimerization of the protein.

Figure 3.47. Structural alignment of POZ with BPM1. POZ is shown in brown and BPM1 is shown in blue. The MATH domain of BPM1 aligns with the beta barrel structure of POZ.

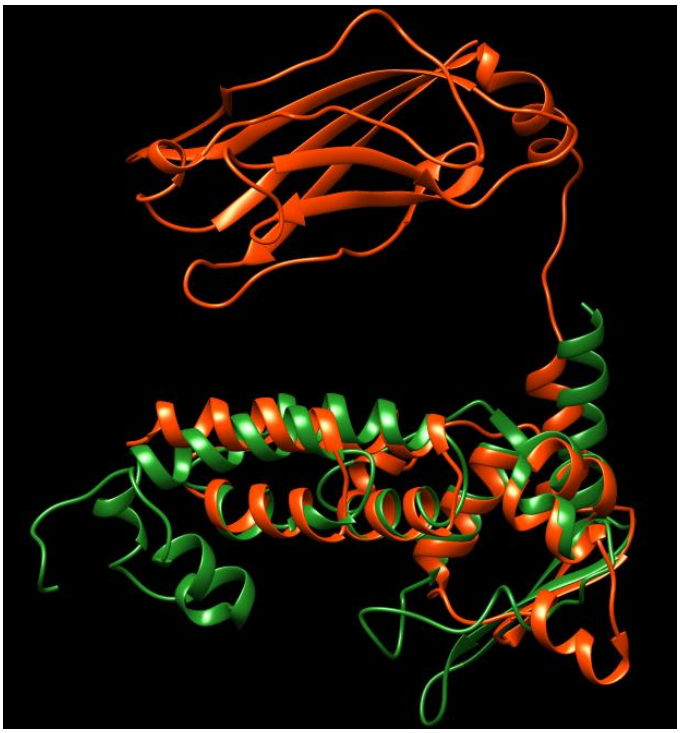


Figure 3.48. Structural alignment of POZ with a BTB-only-containing protein, AT2G40440. POZ is shown in brown and AT2G40440 is shown in green.

To identify tissues and developmental stages in which *POZ* is expressed, the gene expression levels of *POZ* in the top 10 highly expressed tissues, was obtained using Genevestigator (Figure 3.49). It can be seen that *POZ* appears to be expressed in both aerial and below-ground parts to a high level. During the development of the plant, the expression of *POZ* appears to remain relatively constant, but appears to increase towards senescence when the siliques have matured, before a reduction in expression levels at senescence (Figure 3.50).

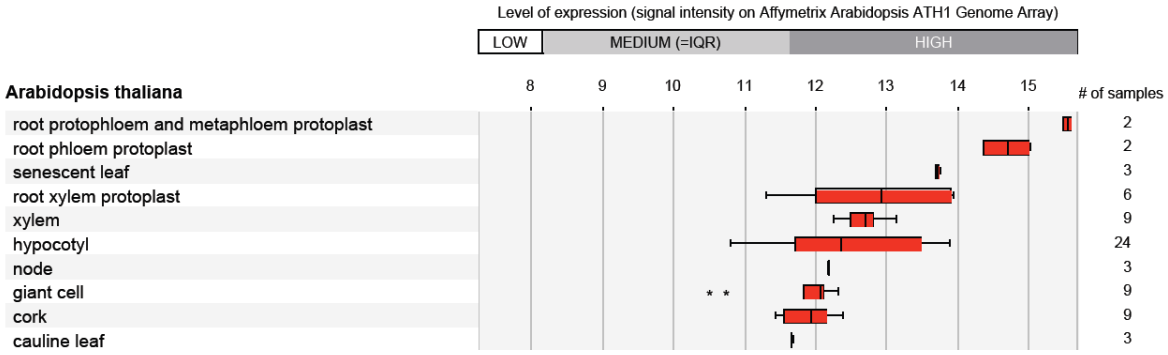


Figure 3.49. The expression of *POZ* within the top 10 most expressing tissues (from Genevestigator). The numbers indicate $\log_2$ values.

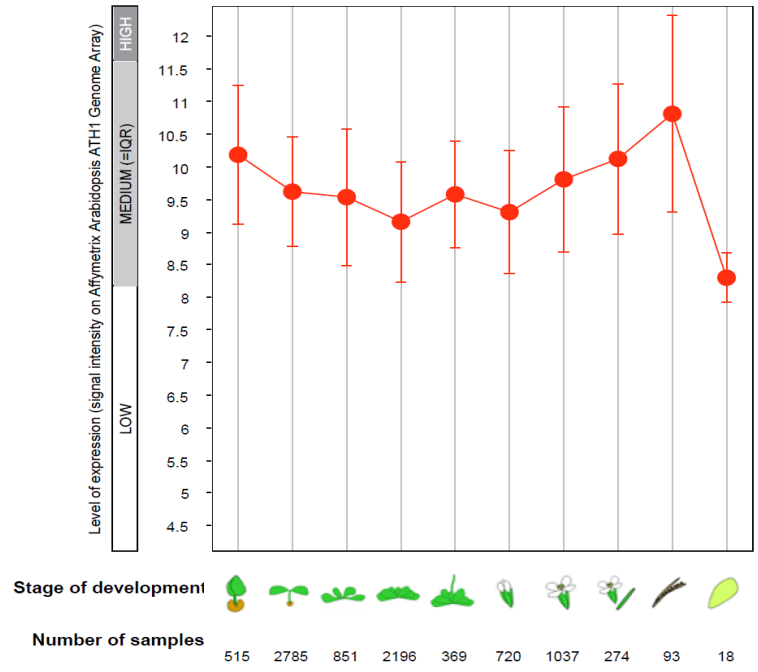


Figure 3.50. The expression of *POZ* during plant development (from Genevestigator). The numbers indicate $\log_2$ values.

3.8. Discussion

3.8.1. Six hub genes were identified using VBSSM

This work is based on a time-series transcriptome analysis of slow-drying Arabidopsis Col-4 plants (Bechtold *et al.*, 2016). A list of differentially expressed genes was obtained from the transcriptome analysis. These were clustered using TCAP to create smaller lists of genes to input into VBSSM to make it computationally less intensive. These smaller gene lists were used to construct GRNs using VBSSM from which network hub genes were identified as potentially important in the drought response.

In addition to *Agamous-like22* (*AGL22*) which was identified by Bechtold *et al.* (2016), 6 more genes were identified from modelling GRNs: an unknown gene whose protein contains the BTB/POZ domain, and is designated as *POZ*; *Flowering Locus D* (*FD*), a bZIP (Basic Leucine Zipper) protein involved in the positive regulation of flowering (Abe *et al.*, 2005); another unknown gene designated as *UKTF*; *Related to AP2.12* (*RAP2.12*), a member of the ERF (Ethylene Response Factor) subfamily of the ERF/AP2 transcription factor family, and has been found to be stress-responsive (Licausi *et al.*, 2011; Papdi *et al.*, 2015); *Basic helix-loop-helix038* (*BHLH038*), a protein that regulates iron homeostasis (Wang *et al.*, 2007) and *Anthocyaninless2* (*ANL2*), a protein involved in the accumulation of anthocyanin and in root development (Kubo *et al.*, 1999).

RAP2.12 is the only gene of the seven hub genes that has previously been implicated in any stress response (Licausi *et al.*, 2011; Papdi *et al.*, 2015). In the work of Licausi *et al.* (2011), plants overexpressing *RAP2.12* were found to recover better from submergence-induced hypoxia due to the activation by *RAP2.12* of hypoxia-responsive genes. Papdi *et al.* (2015) showed that an inducible overexpressor of *RAP2.12* was found to be tolerant to oxidative and osmotic stresses. The other six hub genes have not been implicated during drought, but a Genevestigator analysis showed they are all responsive to one or more of cold, osmotic, salt, drought and heat stresses.

3.8.2. Characterisation of the unknown proteins, UKTF and POZ

The two unknown proteins UKTF and POZ were analysed using bioinformatics. The UKTF protein contains the MIP1 domain and thus possibly binds to and regulates MADS-box proteins. A MIP1 in *Antirrhinum majus* was found to bind C- and E-class genes

involved in carpel development. Thus it is possible that UKTF is similarly involved and regulates the development of carpels during water-deficit situations.

POZ contains the BTB/POZ and BACK-like domains which are characteristic of proteins involved in forming Cullin3-based E3 ligases (Gingerich *et al.*, 2005). The most similar protein to POZ in *A. thaliana* is AT1g21780, which has been shown to bind Cullin3 by Yeast-2-hybrid (Gingerich *et al.*, 2005). The E3 ligase brings together the target protein and the E2-Ubiquitin intermediate to enable the transfer of ubiquitin to the target protein, which leads to protein degradation by the 26S proteasome (Gingerich *et al.*, 2005). Thus, it is likely that POZ is involved in the same process.

Of the 85 BTB-containing proteins in Arabidopsis, only about 10 of these are annotated to have only a single domain. POZ is one of the few proteins that have only a single domain in their sequence. Most BTB-containing proteins in Arabidopsis possess another domain in their sequence, such as a MATH, an NPH3, a TAZ or an Ankyrin domain (Gingerich *et al.*, 2007). From the phylogenetic analysis in Figure 3.44, it can be seen that POZ is closely related to the BTB-MATH (BPM) proteins. A structural alignment of POZ with one of the BPMs, BPM1, shows a good overlap between the two, while the BTB-only protein, AT2G40440, aligns to only part of POZ. Thus, it is possible that POZ also contains the MATH domain in addition to the BTB domain. This is significant because BPM proteins have been shown to bind AP2/ERF proteins, and possibly target them for degradation (Weber and Hellman, 2009; Chen *et al.*, 2013). Hence, POZ may be involved in targeting similar proteins or other transcription factors to regulate the protein levels of these genes during drought and also promote protein turnover.

This chapter also describes the isolation of T-DNA knockout/knockdown lines for the above seven genes except for *UKTF*. No knockout for *UKTF* could be identified instead two T-DNA insertional lines were identified that were overexpressing the gene. Overexpressing lines for each of the genes were also made in the Col-0 background. Both loss-of-function and gain-of-function mutants were phenotyped to identify a role of these genes during drought, and this is described in chapter 4.

3.9. Conclusion

- Based on the work of Bechtold *et al.* (2016), a list of differentially expressed genes was analysed to identify gene regulatory networks involved in drought using VBSSM.

- Six genes were identified as ‘hub’ genes: *FD*, *RAP2.12*, *BHLH038* and *ANL2*, and the unknown genes *POZ* and *UKTF*. These genes were analysed in this thesis along with *AGL22* which was identified by Bechtold *et al.* (2016).
- Knockouts T-DNA lines of these genes were screened and identified and overexpressors were generated by Gateway® cloning.
- UKTF was identified *in silico* as a putative MADS-box interacting protein that may be involved in carpel development.
- POZ may possibly be involved in forming Cullin3-based E3 ligases.

Chapter 4

Phenotyping mutant lines of hub genes under progressive drought

4.1. Introduction

In the previous chapter, gene regulatory networks were modelled using time-series transcriptomics data obtained from a progressive drought experiment in Arabidopsis (Bechtold *et al.*, 2016). The gene networks generated identified six 'hub' genes that could potentially have an important role in plants' drought stress response and regulate drought-specific gene networks: *POZ*, *FD*, *UKTF*, *RAP2.12*, *BHLH038* and *ANL2*. Loss-of-function and gain-of-function mutant lines of these genes, as well as for *AGL22* which was identified by Bechtold *et al.* (2016), were obtained to study the role of these genes during drought.

This chapter describes the phenotyping of these lines under drought conditions and elucidates a possible role for these genes under water stress. Where possible, phenotyping was done to assess water use, photosynthetic performance, stress status under drought and plant growth and development in well-watered and drought-stressed plants. Drying rate and relative leaf water content were used to measure water use, while gas exchange and thermography were carried out to follow the progression of drought stress as well as give indications about photosynthetic performance. Electrolyte leakage and hydrogen peroxide levels were measured along with chlorophyll and anthocyanin content, as an indication of the stress status of the plants, while general growth parameters were taken as an indication of the impact of drought stress on plant performance.

AGL22 is a repressor of flowering that interacts with *Flowering Locus C* (*FLC)* and subsequently binds at the promoters of *Flowering Locus T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*) to repress floral transition (Lee *et al.*, 2007; Li *et al.*, 2008). *AGL22* integrates signals from the autonomous, gibberellic acid and thermosensory pathways to control flowering, by regulating the expressions of the above two genes. Though *AGL22* has not been reported to be involved in any abiotic or biotic stress response, it has been shown to detect changes in ambient temperature as part of its thermosensory regulation of flowering (Lee *et al.*, 2007). In that study, flowering time was found to be delayed in wild-type plants grown at 16 °C compared to at 23 °C, as expression levels of *AGL22* increased and *FT* greatly decreased, at 16 °C. Thus, there appears to be a temperature component to the regulation of the expression of *AGL22*.

FD is another gene that is involved in flowering time that was detected by the VBSSM algorithm. It is a protein present in the shoot apex that interacts with *FT* to promote

flowering (Abe *et al.*, 2005; Wigge *et al.*, 2005). Overexpression of *FD* in Arabidopsis leads to early flowering and knockouts of the gene have a delayed flowering phenotype, but like *AGL22*, it has not been shown to have a role in any stress response.

RAP2.12 has been found to be involved in the response of plants to hypoxia (Licausi *et al.*, 2011). Wild-type plants and plants overexpressing *RAP2.12* were subjected to flooding and submergence, resulting in a hypoxic situation. The plants overexpressing *RAP2.12* were found to recover better after normal oxygen levels returned, due to the activation of hypoxia-responsive genes by *RAP2.12*. It was also found that in rice, another AP2/ERF protein called *SUBMERGENCE1A* (*SUB1A*) has been found to mediate tolerance to drought and submergence in accession lines introgressed with the gene or overexpressing the gene (Fukao *et al.*, 2011). It was shown that *SUB1A* did this by controlling the expression of stress-responsive genes and by regulating the hormones abscisic acid, ethylene and gibberellic acid.

BHLH038 is involved in maintaining iron homeostasis and uptake, and along with *BHLH039* interacts with *FIT (Fe deficiency Induced Transcription factor)* to induce the expression of *FRO2* (a ferric chelate reductase) and *IRT1* (Ferrous iron transporter; Yuan *et al.*, 2008). Overexpressors of *BHLH038* and *BHLH039* are tolerant to iron deficiency and have higher accumulation of iron in shoots, due to constitutive expression of *FRO2* and *IRT1* (Yuan *et al.*, 2008).

ANL2 is involved in the accumulation of anthocyanin in the sub-epidermal layer of the leaf and the cellular organisation of the roots in Arabidopsis (Kubo *et al.*, 1999). Anthocyanins are produced when plants are subjected to stress (Chalker-Scott, 1999). They appear capable of absorbing light incident on chloroplasts, and thus reducing chlorophyll bleaching and superoxide production, and they are also capable of scavenging superoxide radicals (Neill and Gould, 2003). Thus, accumulation of anthocyanins would be beneficial during periods of stress, to reduce the effect of oxidative stress.

Other than *RAP2.12*, none of the other genes identified using VBSSM have been documented as having a stress-responsive role in plants. Using knockouts and overexpressors of these genes, it will be possible to study them in the context of drought, and test the potential role of these genes in the plant response to drought. This chapter describes the effect of a progressive (slow-drying) drought on mutants of these genes, their response to the effects of oxidative stress due to the drought and the yield capability of these mutants under normally watered and drought conditions.

4.2. Phenotyping and analysis of knockouts and overexpressors of *AGL22*

4.2.1. Measurement of water use under drought

To ascertain whether *AGL22* is involved in drought stress response, two knockouts of *AGL22* (*agl22-3* and *agl22-4*) and two overexpressors (AGL22ox-10 and AGL22ox-2) were subjected to a progressive drought. AGL22ox-10 was a highly overexpressing line (approximately 45-fold) and AGL22ox-2 was a mild overexpressor (approximately 15-fold). These were compared against the wild-type, Col-0. A segregating wild-type from the *agl22-4* line, named here as AGL22 WT, was also subjected to a progressive drought to compare it against Col-0. The drought was started 3 weeks after sowing due to the early flowering phenotypes of the knockouts. The drought was continued until the relative soil water content (rSWC) reached 20%, at which point the plants began to wilt. Figure 4.1 shows the drying rates calculated for each of the above-mentioned wild-type and mutant plants.

The two knockouts, *agl22-3* and *agl22-4*, showed a significantly higher drying rate than the wild-type plants of Col-0 (Figure 4.1A). As expected, the segregating wild-type, AGL22 WT, showed no difference in drying rate compared to Col-0 (Figure 4.1B). With regard to the overexpressors, AGL22ox-10 showed no difference in drying rate compared to the Col-0 plants (Figure 4.1C), while AGL22ox-2 had a significantly higher drying rate than the Col-0 (Figure 4.1D). The difference between these two lines could be due to the difference in the expression levels of *AGL22* (Figure 3.24).

To understand the reason for the difference in drying rates between the different lines, the rosette area of the plants were measured. Images of the plants were processed using the image processing software, ImageJ, to measure rosette area. Figure 4.2 shows the rosette areas measured for each of the different lines. The rosette areas were measured at a point during the drought period when a difference in rosette size was observed, so as to correlate the drying rate phenotype with the rosette area.

The rosette areas of the two knockout lines were measured on day 10 of the drought and found to be significantly larger than the rosette area of Col-0 (Figure 4.2A), and this corresponds to the observed higher drying rates (Figure 4.1A). The rosette area of the segregating wild-type, AGL22 WT, also measured on day 10, was found to be similar to the rosette area of Col-0 plants (Figure 4.2B). The rosette area of the overexpressor, AGL22ox-10, measured on day 14, showed a no difference compared to Col-0 (Figure

4.2C). On the other hand, on day 14 a significantly larger rosette area was measured in the overexpressor, AGL22ox-2, compared to Col-0 (Figure 4.2D) and this corresponds to the higher drying rate in AGL22ox-2 as observed in Figure 4.1D.

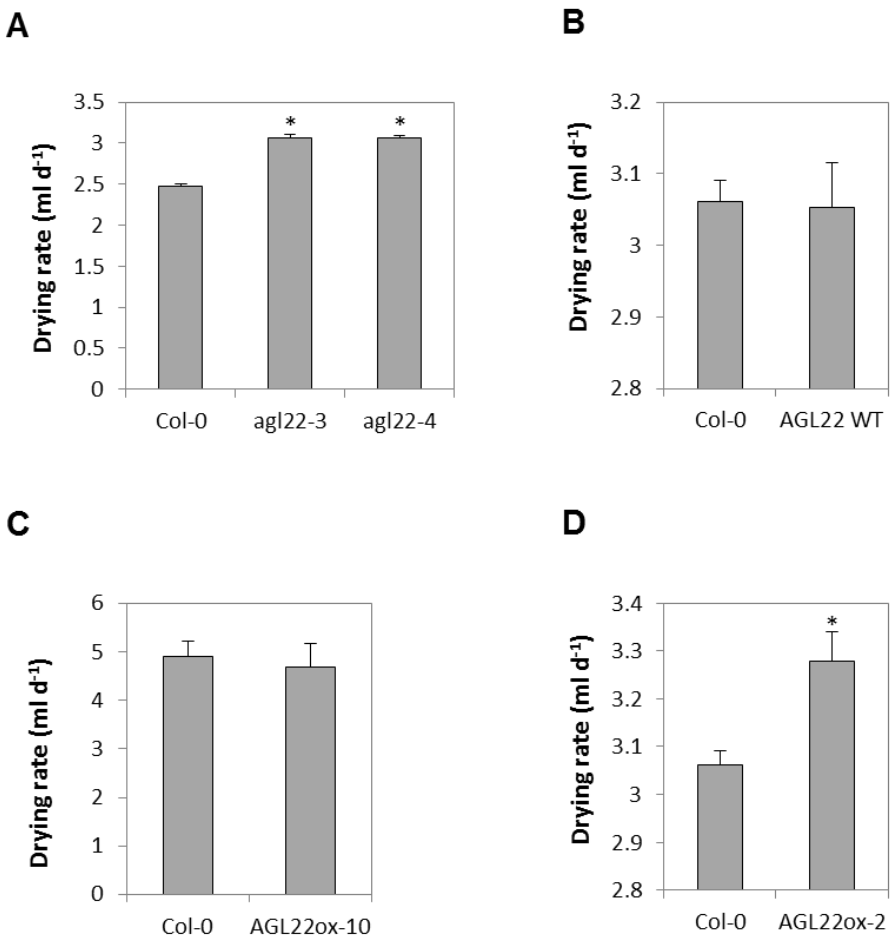


Figure 4.1. Drying rates for *AGL22* knockouts, overexpressors and a segregating wild-type compared to the wild-type, Col-0. Drying rates for (A) the knockouts, *agl22-3* and *agl22-4*, (B) a segregating wild-type, AGL22 WT, (C) the overexpressor, AGL22ox-10 and (D) the overexpressor, AGL22ox-2. The values are mean + standard error (n=10) and * indicates significant difference relative to Col-0 ($P < 0.05$, Student's t-test). The experiment was performed twice.

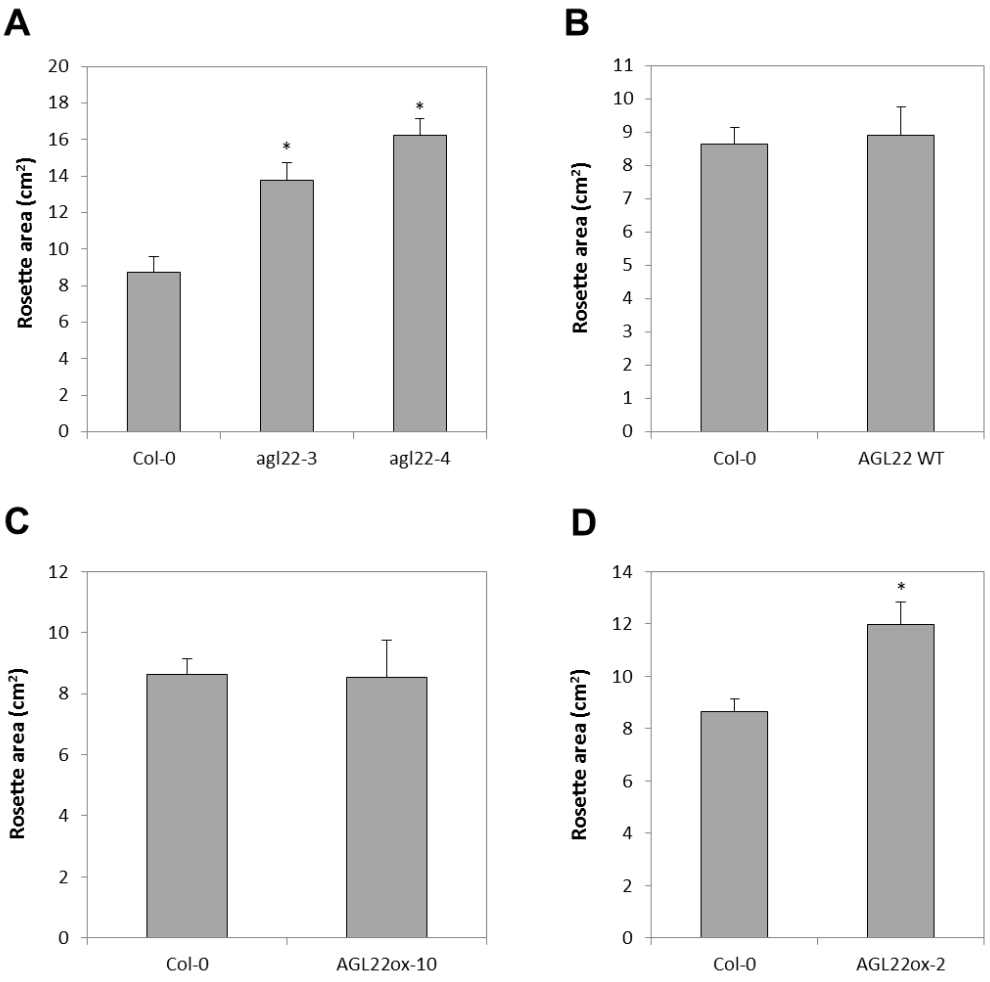


Figure 4.2. Rosette areas for *AGL22* knockouts and the overexpressor compared to the wild-type, Col-0. Rosette areas for (A) the knockouts, *agl22-3* and *agl22-4*, (B) a segregating wild-type, AGL22 WT, (C) the overexpressor, AGL22ox-10 and (D) the overexpressor, AGL22ox-2. The rosette area in figures A and B were estimated on day 10 of the drought, the rosette area in C was measured on day 0 and the rosette area in D was measured on day 14. The values are mean + standard error (n=6) and * indicates significant difference relative to Col-0 ($P < 0.05$, Student's t-test). The experiment was performed twice.

4.2.2. Measurement of photosynthetic performance

Due to the small size of the leaves for the majority of the drought period, the average temperature of the Col-0 and *agl22* knockout lines were measured as an approximate gauge of stomatal conductance. The average temperatures of the Col-0 and *AGL22* knockouts were measured over the course of the drought, using the infrared thermal camera TH7100 Thermal Tracer (NEC Avio Infra-red Technologies Co. Ltd), and these were compared against watered control plants of the same age (Figure 4.3). Figure 4.3A shows the average temperatures of Col-0, *agl22-3* and *agl22-4* at regular intervals over

the course of the drought period, with the corresponding rSWC shown in Figure 4.3C. No significant difference in the average temperatures of the droughted plants was seen between the three lines until day 23, when the average temperature of *agl22-4* was significantly higher than that of Col-0 (Figure 4.3A and D). The average temperature of *agl22-3* was also higher than that of Col-0, but this was not significant. Figure 4.3D compares the average temperature of the three lines on day 23 for both droughted and watered plants. It shows that there was a significantly higher average temperature for all the lines under drought conditions as opposed to the watered plants.

The average temperature of AGL22ox-10 plants was measured at different rSWC, namely 75%, 50%, 30%, 25% and 20% (Figure 4.4). It was only possible to measure the average temperature of AGL22ox-10 plants as the second overexpressing line was not available. An increase in temperature was observed over the course of the drought, and this was higher than the average temperature of the watered plants. However, there was no observed difference in temperature between the overexpressor and the wild-type over the course of the drought (Figure 4.4A).

The stomatal conductance (g_s) and carbon assimilation rate (*A*) of AGL22ox-10 plants were also measured in both droughted and watered plants using a CIRAS-1 (PP Systems; Figure 4.5). A decrease in g_s was observed over the course of the drought (Figure 4.5A), while g_s remained the same in the watered plants over the same period (Figure 4.5B). However, there was no significant difference in g_s between AGL22ox-10 and Col-0 under both conditions. No difference in *A* was observed in the droughted plants or between the two lines over the course of the drought (Figure 4.5C and D).

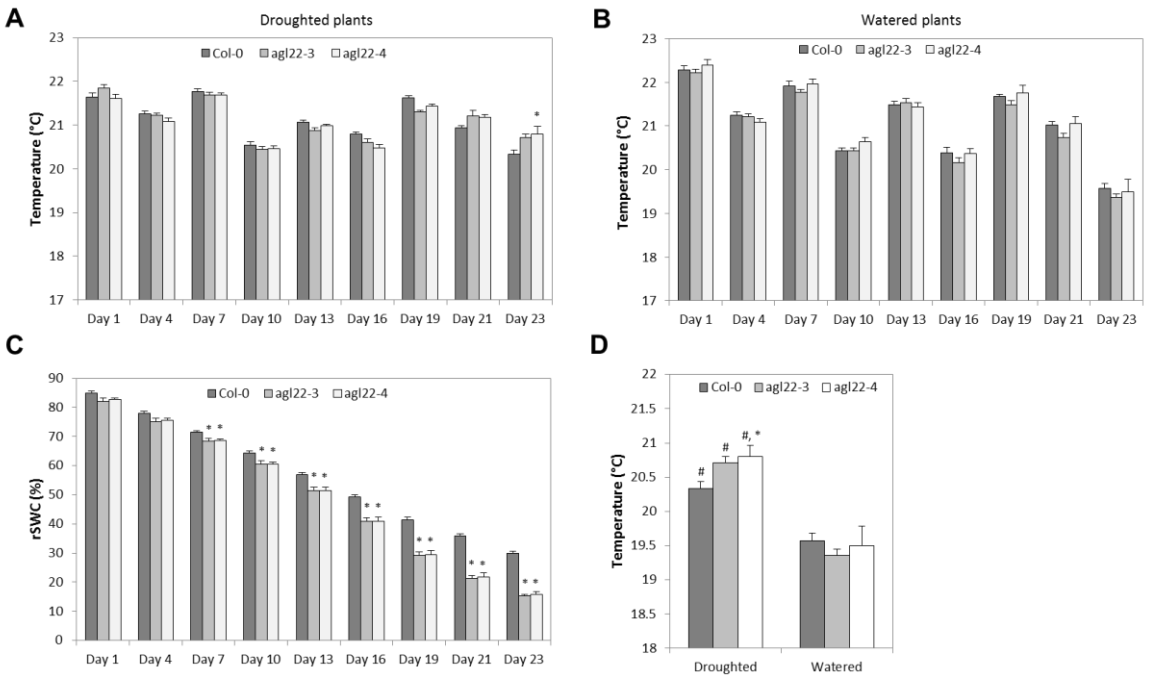


Figure 4.3. Temperature of the *agl22* knockouts compared to the Col-0 wild-type, measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought (n=10). (B) The temperature of the corresponding watered control plants (n=10). (C) The relative soil water content (rSWC) of the plants subjected to drought whose temperature was measured (n=10). (D) The temperature of the droughted and watered plants on day 23 (n=10). All the values are mean + standard error. * indicates significant difference relative to Col-0 of the same treatment ($P < 0.05$, Student's t-test); # indicates significant difference relative to watered control plants ($P < 0.05$, Student's t-test). The experiment was performed once.

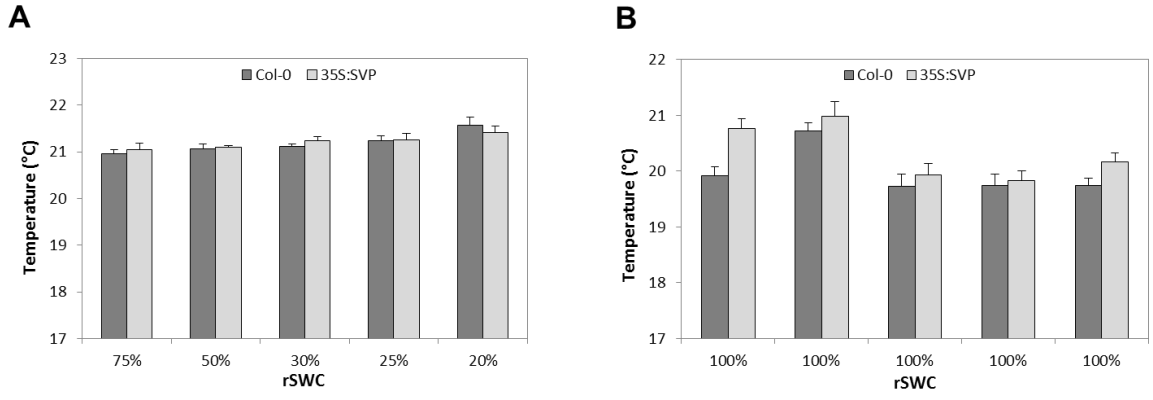


Figure 4.4. Temperature of AGL22ox-10 compared to the Col-0 wild-type, measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

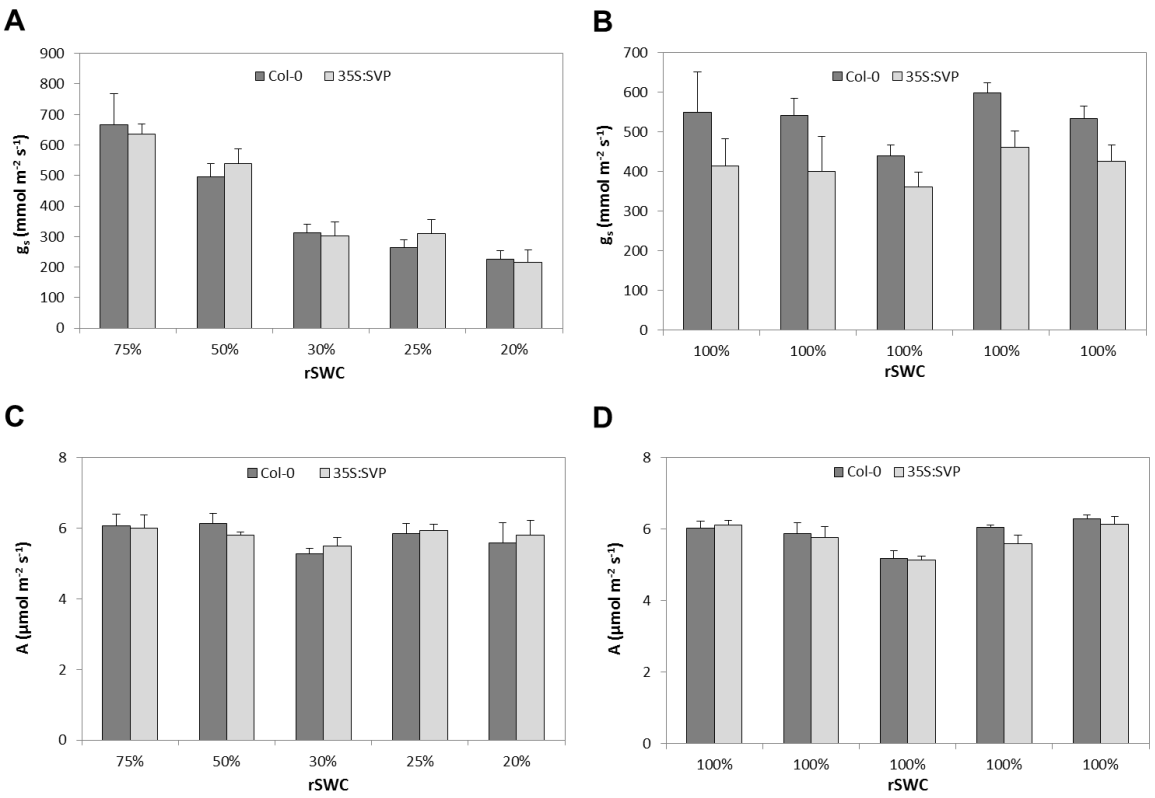


Figure 4.5. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in AGL22ox-10 and the Col-0 wild-type, at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

4.2.3. Measurement of the stress status of drought-stressed plants

As a measure of the effect of the stress on the plants, the levels of the reactive oxygen species (ROS), hydrogen peroxide (H_2O_2), produced in Arabidopsis leaves were measured using the fluorescent dye homovanillic acid (Guilbault *et al.*, 1967). Figure 4.5 shows the H_2O_2 levels measured in Col-0, the segregating wild-type, AGL22 WT, the two knockouts and the two overexpressors, under drought and watered conditions. As shown in the figure, higher levels of H_2O_2 was found in the plants subjected to drought than in the control plants across the wild-type and overexpressor lines. However, the knockouts did not show any difference in the levels of H_2O_2 between the drought and watered plants. Also, the level of H_2O_2 produced in the knockouts under drought conditions was less than the amounts produced in the wild-type lines and overexpressors. It can also be

seen that the overexpressors produced more H_2O_2 than the wild-type plants, and the greatest difference in the amounts of H_2O_2 produced between the droughted and watered plants was observed in the overexpressors.

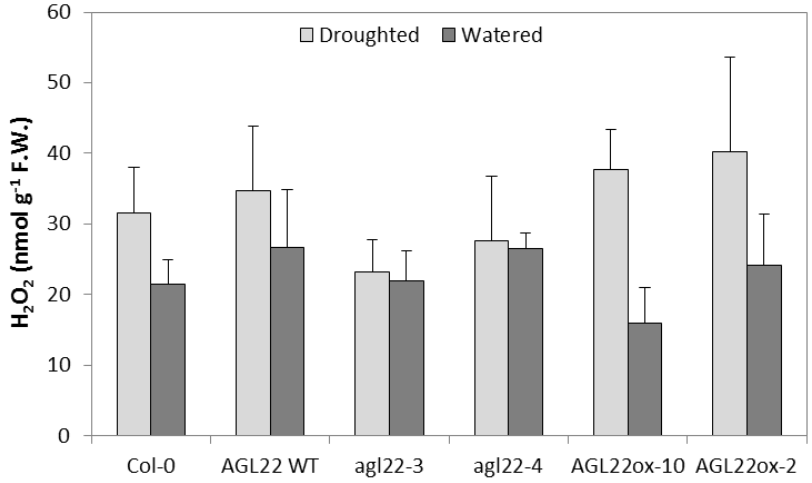


Figure 4.6. Measurement of H_2O_2 levels in two wild-types, two knockouts and two overexpressors of *AGL22* under droughted and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

4.2.4. Measurement of growth and development

4.2.4.1. Flowering time

The flowering time of the various lines under droughted and watered conditions was also studied as there is a known correlation between drought and flowering time (Juenger *et al.*, 2005), and also because *AGL22* is involved in the regulation of flowering. Flowering time was measured in terms of number of leaves at floral transition and the number of days to floral transition from sowing. As shown in Figure 4.7A, there appeared to be no difference in flowering time between the droughted and watered plants in both wild-type and overexpressor lines. On the other hand, a difference was observed in both knockouts in which the droughted plants flowered earlier than the watered plants, and this was particularly significant for *agl22-4*. Also, the number of leaves at flowering for the *agl22*

knockouts was also found to be significantly fewer than, and almost half the number of, those of the respective treatments of the Col-0, AGL22 WT, AGL22ox-10 and AGL22ox-2 lines. There was no significant difference in the number of leaves at flowering between the droughted and watered plants of AGL22ox-2. However, AGL22ox-2 was significantly different to all other lines: AGL22ox-2 had more leaves at flowering than the two knockouts, but fewer than the two wild-type lines and the second overexpressing line, AGL22ox-10.

No differences were observed in the number of days to flowering between the droughted and watered plants in all the lines (Figure 4.7B). However, flowering occurred significantly earlier in the knockouts subjected to both treatments than in the two wild-types and overexpressors of the respective treatments. The AGL22ox-2 plants subjected to drought flowered significantly later than the droughted knockouts, and significantly earlier than the Col-0. There was no difference in the flowering time between the droughted AGL22 WT and AGL22ox-2; a difference was seen between the two overexpressors but this was not significant. The watered AGL22ox-2 plants flowered significantly later than the two knockouts and significantly earlier than Col-0 and AGL22ox-10.

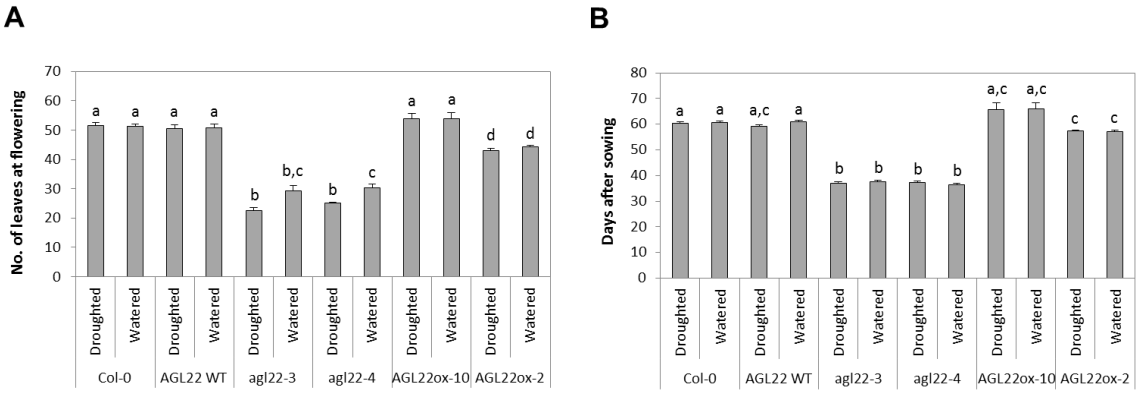


Figure 4.7. Flowering time measured in two wild-types, two knockouts and two overexpressors of *AGL22* under droughted and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=10). Means with different letters indicate significant difference ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.2.4.2. Biomass and harvest index

The yield of the plants subjected to drought were compared against watered plants to ascertain if there was any penalty to yield due to the stress and whether the knockouts and overexpressors of the flowering time gene, *AGL22*, showed any difference in yield. The dried rosette weight, the weight of dried stalks and siliques, seed yield and total biomass were measured as parameters to determine the yield of the plants. Figure 4.8 shows the average yield measured for the lines Col-0, AGL22 WT, *agl22-3*, *agl22-4*, AGL22ox-10 and AGL22ox-2, under droughted and watered conditions.

As seen in Figure 4.8A, no difference in dried rosette weight was seen between the droughted and watered Col-0 and plants. However, the dried rosette mass obtained from droughted AGL22 WT plants was less than the watered plants, though this was not significant. Also, no difference in rosette mass was seen between droughted AGL22 WT and Col-0. The dried weight of the watered AGL22 WT plants, on the other hand, was higher than that of the corresponding Col-0. The dried weight of the two knockouts was found to be significantly less than the two wild-type and overexpressor lines, in both treatments, but there was no difference in dried rosette mass between the droughted and watered plants of both lines. A difference between the two treatments was also seen in the overexpressor AGL22ox-10, but this was not significant and not different to the Col-0. However, it appeared to be higher than AGL22ox-2, even though this difference was not significant. Also, no difference was seen between the droughted and watered AGL22ox-2.

The reproductive structures (stalks and siliques) were also weighed and compared between droughted and watered plants of the various lines (Figure 4.8B) Surprisingly, significantly more mass of reproductive structures was produced in the Col-0 plants subjected to drought, than the watered plants. Similarly, more biomass was produced in AGL22 WT plants subjected to drought than the control plants, though this was not significant. Droughted *agl22-3* and *agl22-4* plants produced fewer stalks and siliques than the respective watered plants, though this difference was also not significant. The droughted plant of the two knockouts produced significantly fewer stalks and siliques than the droughted plants for the two wild-types and AGL22ox-2, while the watered *agl22-3* also produced significantly fewer stalks and siliques than watered AGL22ox-2 plants.

Figure 4.8C shows the seed yield obtained for the droughted and watered plants for the two wild-types, the two knockouts and the two overexpressors. A general trend of higher

seed yield in the droughted plants than in the watered plants can be seen. The two knockouts, *agl22-3* and *agl22-4*, had produced significantly less seed (almost one-fifth less) than the two wild-types. The droughted and watered AGL22ox-2 also produced less seed than the two wild-types. Less seed was also produced by AGL22ox-2 than by the other overexpressor, AGL22ox-10, but more than the two knockouts.

Total biomass produced by the different lines was measured under drought and watered conditions (Figure 4.8D). It appears that the droughted Col-0 plants produce significantly more biomass than the watered plants. There appears to be no difference in the biomass produced by droughted and watered AGL22 WT, AGL22ox-10 and AGL22ox-2 plants. In the case of the knockouts, the droughted plants of both lines produced less total biomass than the control plants, though this difference was not significant in both. It can also be seen that the knockouts produced less biomass than the two wild-types and two overexpressors. The droughted knockouts produced significantly less biomass than the droughted wild-types and overexpressors. *agl22-3* watered plants produced less biomass than the corresponding AGL22 WT and AGL22ox-2, and *agl22-4* watered plants produced less total biomass than AGL22ox-2.

The harvest indexes (amount of seed yield produced relative to the total above-ground biomass produced) of the various lines under drought and watered conditions are shown in Figure 4.9. The harvest indexes for the two wild-types were found to be very similar to each other and also showed no difference between the droughted and watered plants of the two lines. The two knockouts had lower harvest index than the wild-types and the overexpressing lines, and droughted *agl22-3* appeared to have a better harvest index than the watered plants, while there was no difference in the *agl22-4* plants. A clear difference was seen between the two overexpressors, wherein AGL22ox-10 produced a higher harvest index than AGL22ox-2 under drought and watered conditions and in both conditions, the droughted plants appeared to outperform the watered plants.

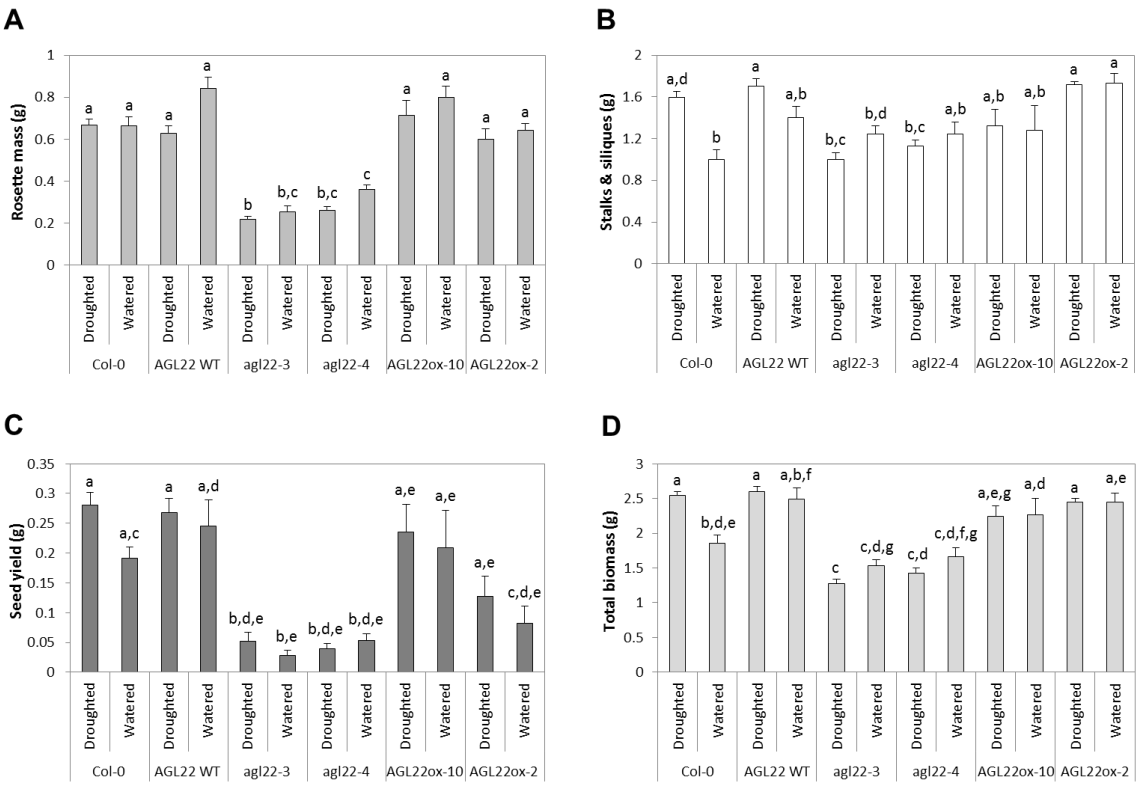


Figure 4.8. Measurement of the vegetative and reproductive biomass of two wild-types, two knockouts and two overexpressors of *AGL22* under droughted and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). Means with different letters indicate significant difference ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

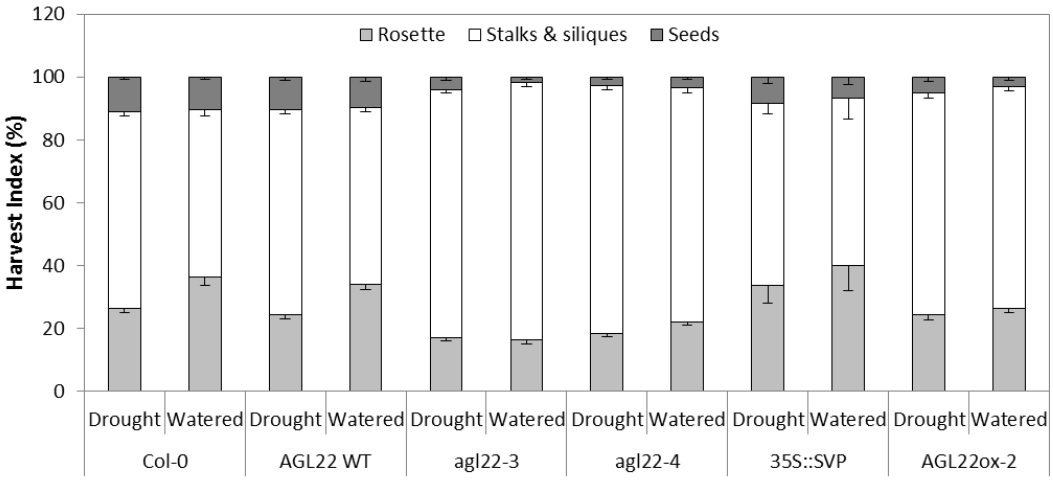


Figure 4.9. Harvest index estimated in two wild-types, two knockouts and two overexpressors of *AGL22* under droughted and watered conditions. All the values are mean - standard error (n=7).

4.3. Phenotyping and analysis of knockouts and overexpressor of *FD*

4.3.1. Measurement of water use under drought

The flowering time gene *FD* was identified as a hub gene while modelling time-series transcriptomics data obtained from a drought experiment. Using the knockout line, *fd-4*, the knockdown, *fd-5*, and an overexpressor of *FD*, FDox-7, a possible role of this gene during drought was studied. Six-week old plants for each of the three lines and Col-0 plants were subjected to a progressive drought experiment for two weeks until 20% rSWC was reached. *fd-4* was obtained quite late during the project and so it was not possible to do a comprehensive analysis with this line, as was done with *fd-5* and FDox-7, and thus this data is missing from many of the experiments below.

Figure 4.10 shows the drying rate calculated for the mutants and the wild-type subjected to drought. There was no difference in the drying rate between Col-0 and the knockdown, *fd-5*, or between Col-0 and the overexpressor, FDox-7. However, the knockout, *fd-4*, had a significantly higher drying rate compared to the wild-type. An analysis of the rosette area of the various lines was done on day 0 of the drought (Figure 4.11). This showed that there was no difference in rosette area between Col-0 and either *fd-5* or FDox-7. *fd-4* appeared to produce a larger rosette area compared to the wild-type, though this was not significant.

Relative leaf water content (rLWC) was measured in Col-0, *fd-5* and FDox-7 plants subjected to drought and watered conditions. Figure 4.12 shows the rLWC calculated for the above. The rLWC of the droughted plants was significantly lower than the rLWC of the respective watered plants. The rLWC of the droughted wild-type was also significantly lower than that of *fd-5* subjected to drought, and there was no difference in rLWC between the two *FD* lines.

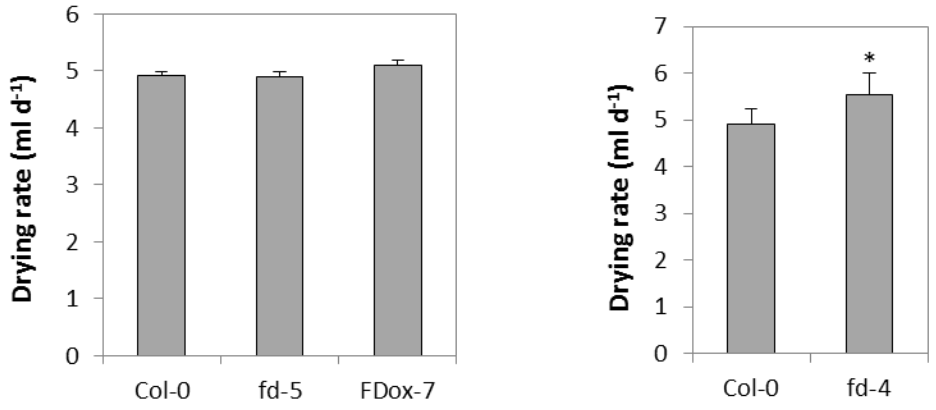


Figure 4.10. Drying rates for *fd* loss-of-function mutants and the overexpressor compared to the wild-type, Col-0. The values are mean + standard error (n=10) and * indicates significant difference relative to Col-0 ($P < 0.05$, Student's t-test). The experiment was performed twice.

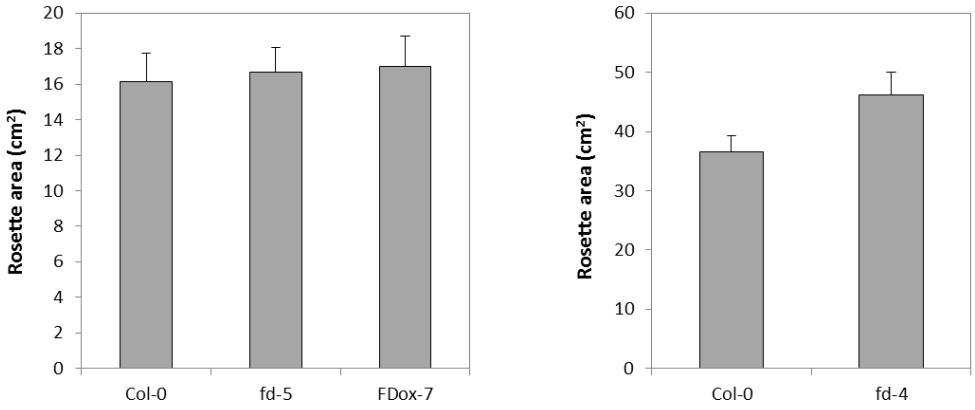


Figure 4.11. Rosette areas for *fd* loss-of-function mutants and the overexpressor compared to the wild-type, Col-0. The values are mean + standard error (n=6). No significant difference was observed ($P > 0.05$, Student's t-test). The experiment was performed once.

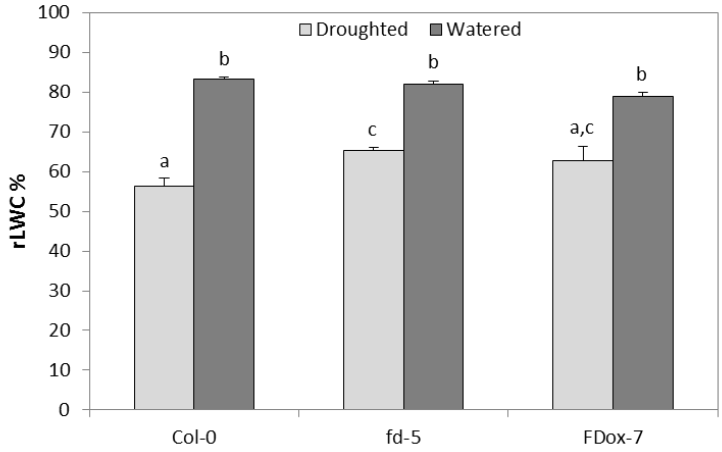


Figure 4.12. Measurement of relative leaf water content (rLWC) in wild-type, the knockdown and the overexpressor of *FD* under droughted and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey HSD test). The experiment was performed once.

4.3.2. Measurement of photosynthetic performance

Stomatal conductance and photosynthetic carbon assimilation of Col-0 and *fd-4* were measured under well-watered and drought stress conditions (Figure 4.13). A decrease in g_s was observed over the course of the drought in both lines (Figure 4.13A), which was not observed in the corresponding control plants (Figure 4.13B). However, there was no difference in g_s between the two lines at each point measured. No change was seen in *A* between the start and the end of the drought experiment, and no difference in *A* was seen between the two lines at any point during the drought (Figure 4.13C).

The average temperature of droughted Col-0 plants compared to *fd-4* plants was also analysed at different rSWC (Figure 4.14). An increase in the average rosette temperature was observed in both Col-0 and *fd-4* plants over the course of the drought, however, no difference was observed between the two lines throughtout the drought.

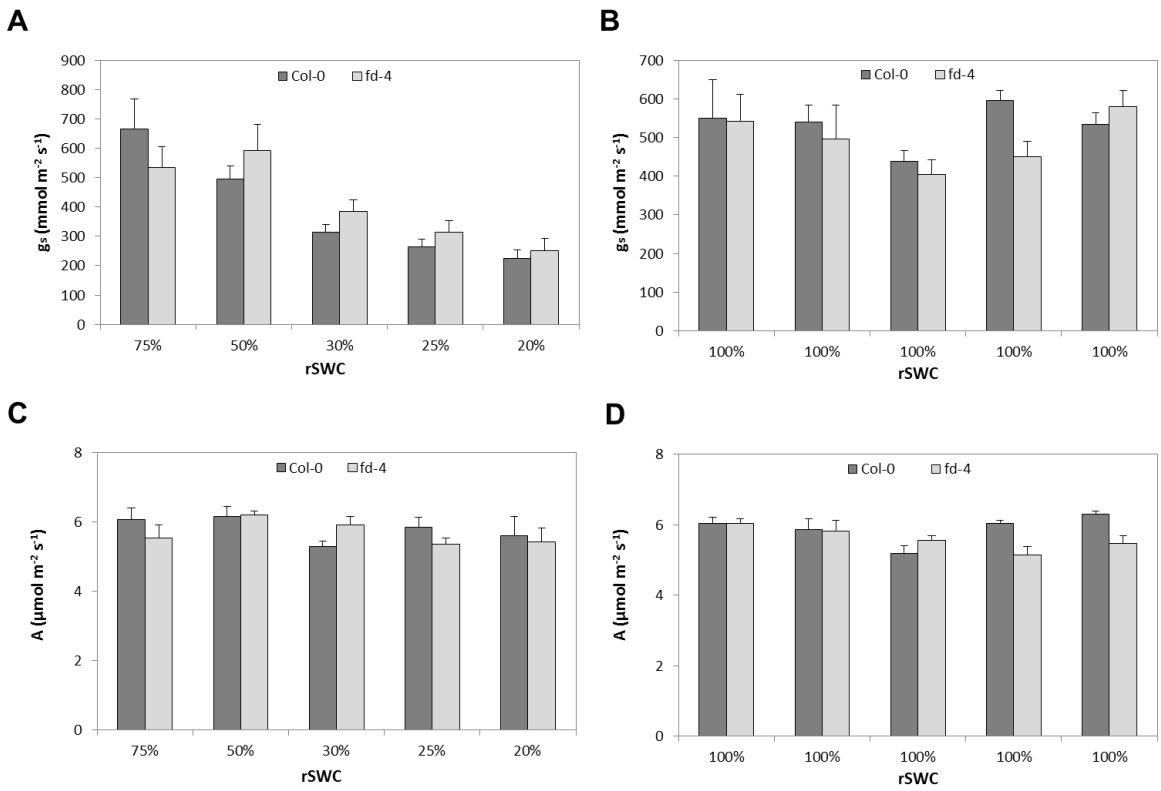


Figure 4.13. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in the *fd* knockout and Col-0 wild-type at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

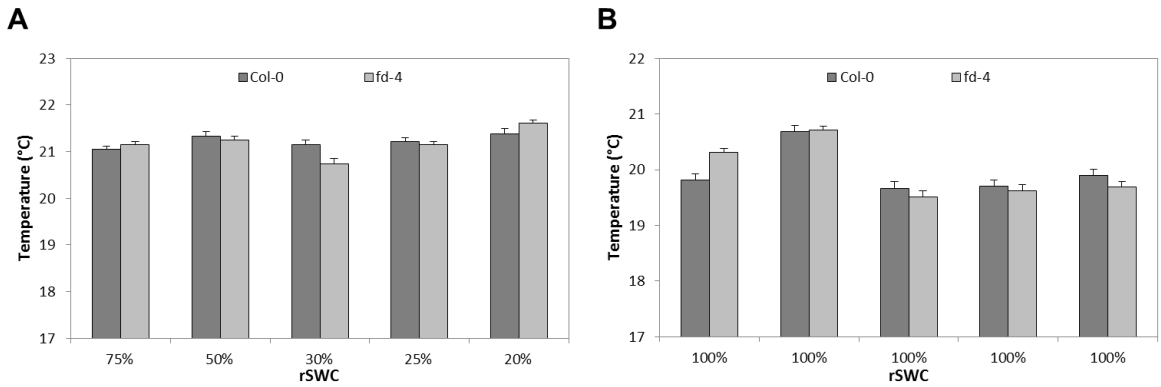


Figure 4.14. Temperature of the *fd* knockout and Col-0 wild-type measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

4.3.3. Measurement of growth and development

4.3.3.1. Flowering time

The flowering time of the *FD* lines was also measured under drought and watered conditions because *FD* is a positive regulator of flowering. Figure 4.15A shows the number of leaves at flowering for Col-0, *fd-5* and FDox-7 subjected to drought or left watered. No difference in flowering time was seen between the droughted and watered plants for *fd-5*, whereas the droughted Col-0 plants appeared to flower earlier than the watered Col-0 plants, though not significantly. On the other hand, the droughted FDox-7 plants appeared to flower later than the watered overexpressors. Between the droughted plants, the overexpressor produced fewer leaves at flowering than *fd-5*, which itself produced fewer leaves at flowering than the Col-0. A similar trend was also seen in the watered plants, and overall the difference in flowering time between the watered lines was greater than that of the droughted plants.

Figure 4.15B shows the flowering time of the three lines under droughted and watered conditions measured in terms of the number of days to flowering from sowing. The droughted plants of all three lines were found to flower earlier than the respective watered plants, though it was not significant for Col-0 and *fd-5*. There appeared to be no difference in flowering time between *fd-5* plants and wild-type plants of the same treatment, while the *FD* overexpressing plants flowered significantly earlier than both the Col-0 and *fd-5* plants of the corresponding treatment.

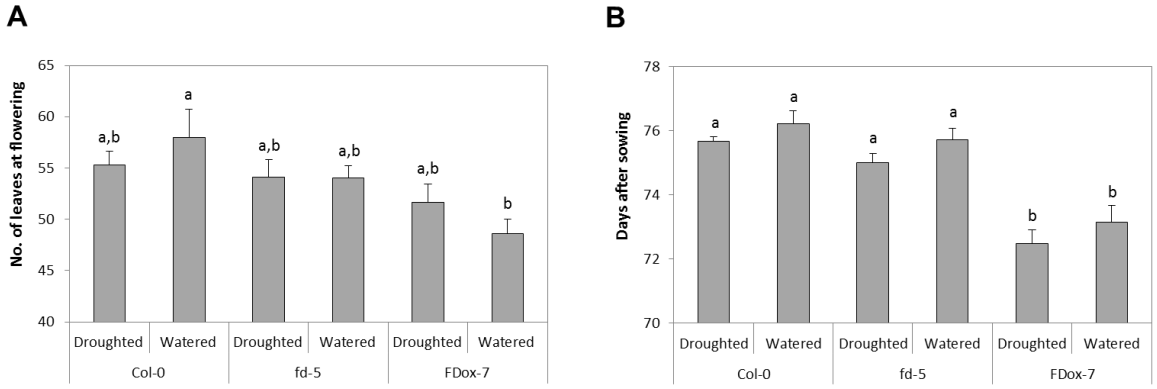


Figure 4.15. Flowering time measured in the wild-type, the knockdown and the overexpressor of *FD* under droughted and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=10). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Gabriel test). The experiment was performed once.

4.3.3.2. Biomass and harvest index

The yield of the three *FD* lines compared to the wild-type under drought and watered conditions was also analysed (Figures 4.16 and 4.18). Rosette weight, the weight of the reproductive structures and seed yield were measured. Figure 4.16 shows the yield of *fd-5* and FDox-7 compared to Col-0. The rosette mass measured for the droughted plants was found to be less than that of the watered plants in all three lines (Figure 4.16A). Though there was no difference in rosette weight between the three lines under watered conditions, both *fd-5* and FDox-7 plants produced less rosette mass than the Col-0 plants under droughted conditions. However, these differences were not significant.

Fewer stalks and siliques were also produced in the plants subjected to drought, compared to the watered plants, though this difference was not significant (Figure 4.16B). No difference was observed in the mass of reproductive structures produced in Col-0 and *fd-5* plants under both drought and watered conditions. However, the overexpressor, FDox-7, produced significantly fewer stalks and siliques compared to Col-0 and *fd-5* under both treatments.

In terms of seed yield, less seed was produced in droughted Col-0 compared to their respective watered controls (Figure 4.16C). No difference in seed yield was observed in both *fd-5* and FDox-7 between droughted and watered plants. Droughted *fd-5* plants produced more seed than the Col-0 plants and significantly more seed than FDox-7 plants. However, under watered conditions, Col-0 plants produced more seed than *fd-5* and FDox-7 plants.

Looking at the total above-ground biomass produced, less biomass was produced in the droughted plants than in watered plants, and this difference was significant for FDox-7. No difference was seen in the amount of biomass produced in the Col-0 and *fd-5* lines under both conditions. FDox-7 plants produced less biomass than Col-0 and *fd-5* plants under both conditions, and this difference was particularly significant under watered conditions for both lines and between droughted Col-0 and FDox-7 plants.

Figure 4.18 shows the yield measured for Col-0 and the knockout, *fd-4*. Unexpectedly, drought-stressed Col-0 plants produced a significantly higher rosette mass than the watered plants, while no difference in rosette weight was seen between differently treated *fd-4* plants (Figure 4.18A). Also, *fd-4* plants appeared to produce more rosette mass than that of Col-0 plants, and this difference was significant between the watered plants.

Figure 4.18B shows that droughted Col-0 plants produced more stalks and siliques than the watered plants, while droughted *fd-4* plants produced fewer reproductive structures than the watered plants. Overall, *fd-4* plants produced less biomass than the wild-type plants under both drought and watered conditions, and this difference was significant in the droughted plants.

Significantly more seed was produced in the Col-0 plants subjected to drought than in the watered Col-0 plants (Figure 4.18C). No difference was seen in the seed yield between the differently treated *fd-4* plants. As with the case of the reproductive structures, less seed was produced in *fd-4* plants than in Col-0 plants, and this difference was significant between the droughted plants.

In Figure 4.18D, it can be seen that the total biomass produced by Col-0 plants subjected to drought was higher than the amount produced by the watered plants, and was also higher than droughted *fd-4* plants. No difference was seen between the droughted and watered *fd-4* plants or between the watered plants of both Col-0 and *fd-4*.

An analysis of the harvest index of the wild-type and *FD* mutant lines under drought and watered conditions is shown in Figures 4.17 and 4.19. In Figure 4.17, the harvest index of droughted Col-0 plants was less than that of the watered plants, while the opposite was true for *fd-5* plants, in which the droughted plants had a higher harvest index than the watered plants. However, no difference in harvest index was seen between the droughted and watered FDox-7, or between the three lines, under both conditions. It also appeared that under droughted conditions *fd-5* plants had a better harvest index, while under control conditions the wild-type had a better harvest index.

In Figure 4.19, it can be seen that the harvest index of Col-0 under droughted conditions was better than under watered conditions, and the same was seen in *fd-4* plants. On the whole, *fd-4* plants had a lower harvest index than Col-0 plants regardless of the conditions.

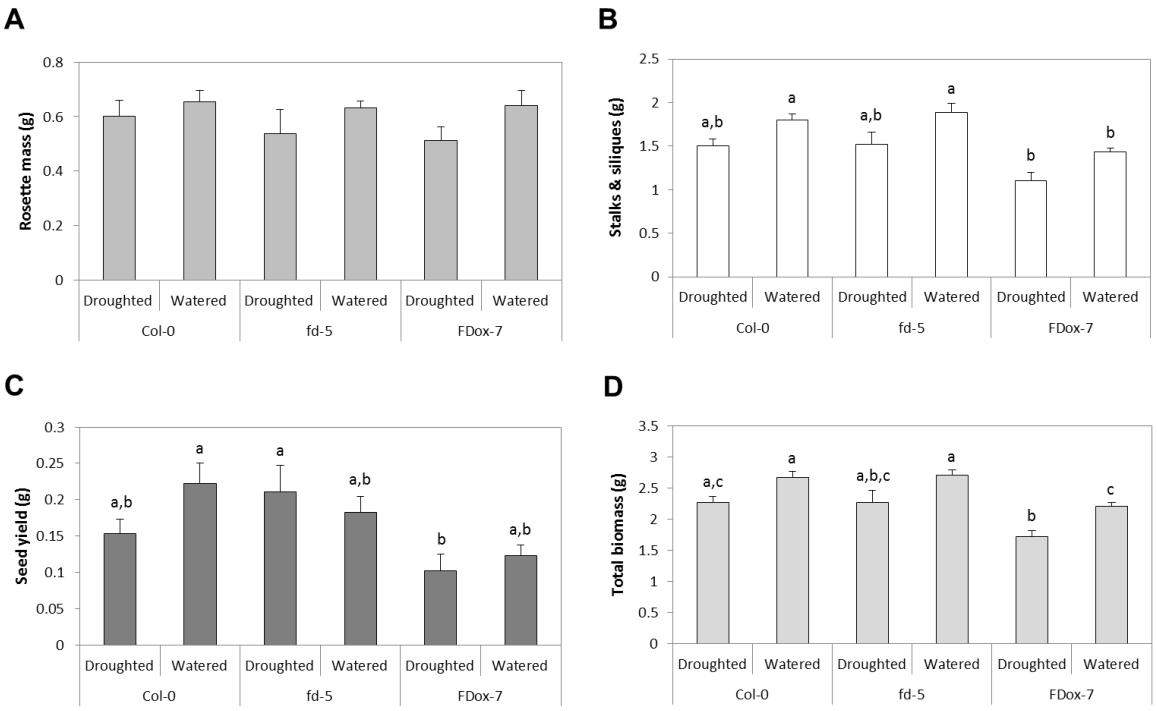


Figure 4.16. Measurement of the vegetative and reproductive biomass of the wild-type, the knockdown and the overexpressor of *FD* under droughted and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). No significant difference was observed in (A) ($P > 0.05$, One-way ANOVA). Means with different letters in (B), (C) and (D) indicate significant difference – for stalks and siliques ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test); for seed yield ($P < 0.05$, One-way ANOVA, post-hoc Tukey HSD test); for total above-ground biomass ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

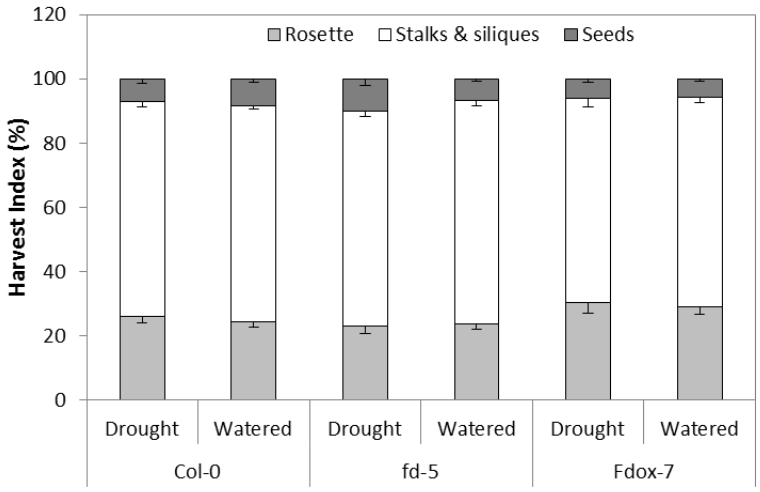


Figure 4.17. Harvest index estimated in the wild-type, the knockdown and the overexpressor of *FD* under drought and watered conditions. All the values are mean - standard error (n=7).

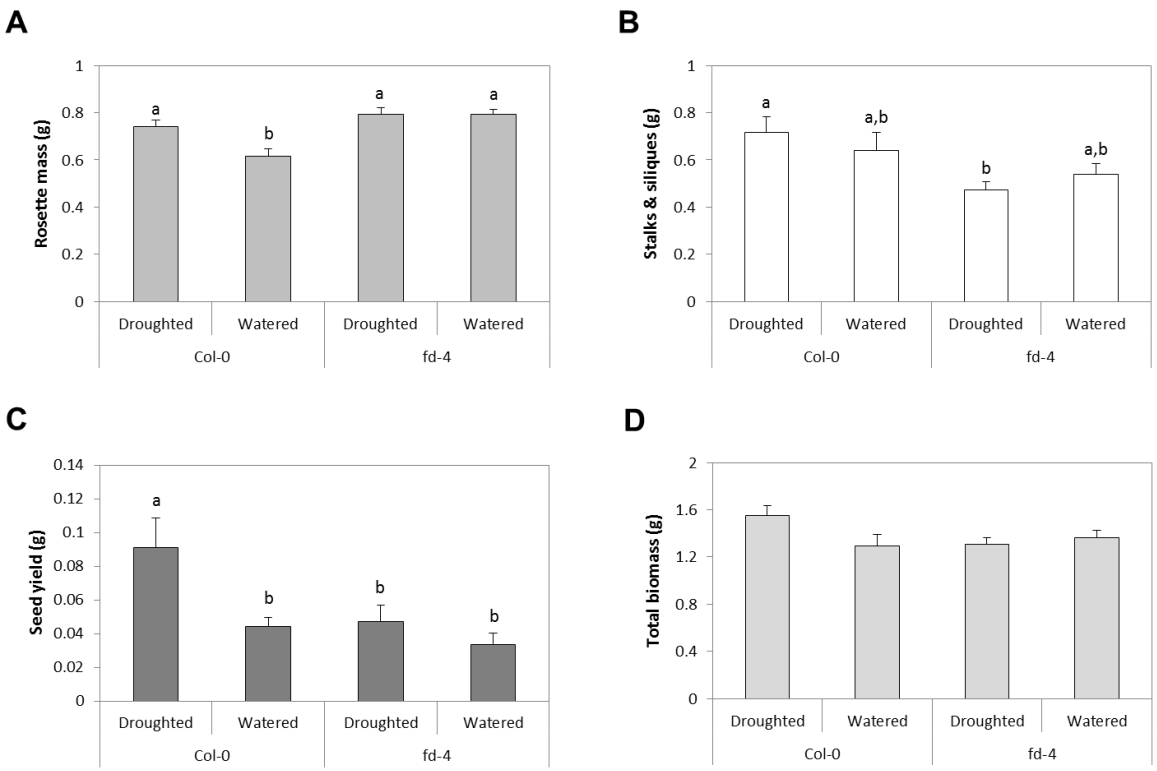


Figure 4.18. Measurement of the vegetative and reproductive biomass for the wild-type and the *fd* knockout under droughted and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). Means with different letters in (A), (B) and (C) indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey HSD test). No significant difference was observed in (D) ($P > 0.05$, One-way ANOVA). The experiment was performed once.

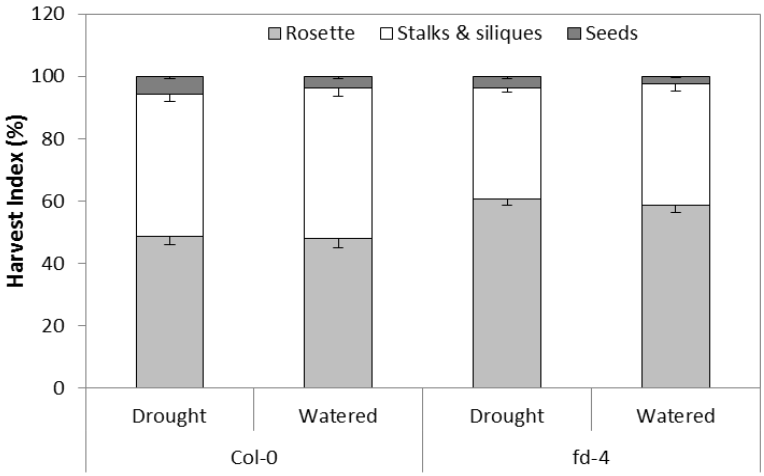


Figure 4.19. Harvest index estimated in the wild-type and the *fd* knockout under drought and watered conditions. All the values are mean - standard error (n=7).

4.4. Phenotyping and analysis of knockouts of *RAP2.12*

4.4.1. Measurement of water use under drought

The potential role of *RAP2.12* in the plant response to drought was studied using two knockouts of the gene, *rap2.12-1* and *rap2.12-3*. Figure 4.20 shows the drying rates for the wild-type and the two knockouts calculated from a progressive drought of 6 week-old plants until 20% rSWC was reached. No difference in drying rate was observed between the two knockouts and the wild-type line. An analysis of the rosette area of the three lines was made on day 0 of the drought, which also showed no difference between the knockout lines and the wild-type (Figure 4.21).

The rLWC content of plants of the wild-type, *rap2.12-1* and *rap2.12-3* lines subjected to drought and watered conditions were also measured (Figure 4.22). The rLWC of the plants subjected to drought was lower than that of the control plants in all three lines, and this difference was significant in Col-0. There was no difference in rLWC between the three lines maintained under droughted and watered conditions.

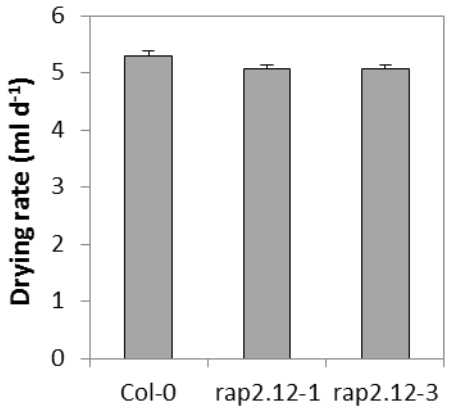


Figure 4.20. Drying rates for *rap2.12* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=10). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

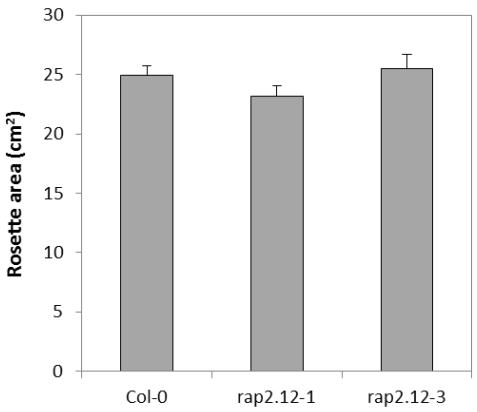


Figure 4.21. Rosette areas for *rap2.12* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=6). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

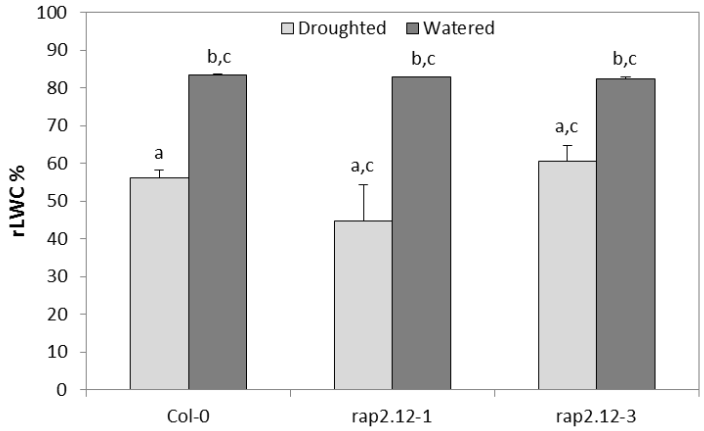


Figure 4.22. Measurement of relative water content (rLWC) in Col-0 wild-type and the *rap2.12* knockouts under drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference (P < 0.05, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.4.2. Measurement of photosynthetic performance

g_s and A of the *rap2.12* knockouts were measured under drought and watered conditions and compared to Col-0 (Figure 4.23). A decrease in g_s was observed in all three lines over the course of the drought (Figure 4.23A), measured at five different rSWC until 20% rSWC. A decrease in *A* was also seen in all three lines as the drought progressed (Figure 4.23C). However, there appeared to be no difference in g_s or *A* between the mutant lines and the wild-type.

An increase in the average temperature of the rosette was observed in all three lines over the duration of the drought stress, despite an anomaly in the data for both the droughted and the corresponding watered plants, when measured at rSWC of 30% (Figure 4.24A). However, here also no difference was seen in the average temperature between the lines at any point of the drought.

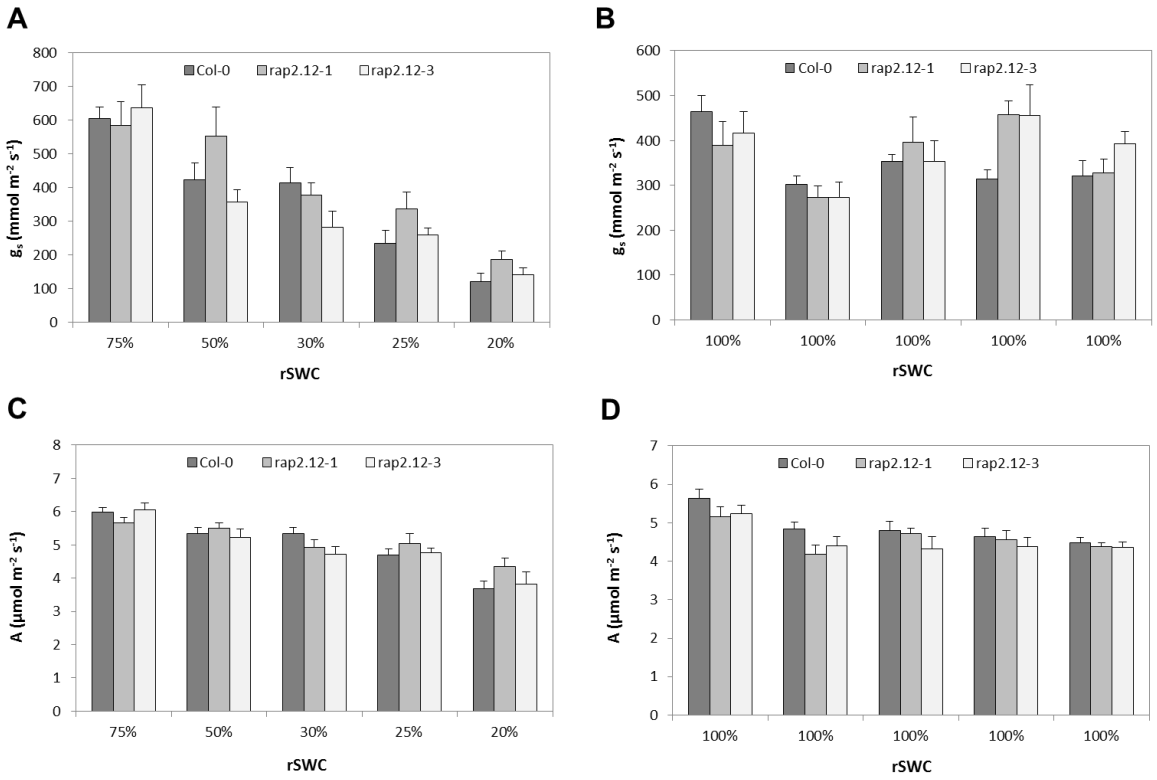


Figure 4.23. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in the *rap2.12* knockouts and Col-0 wild-type at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

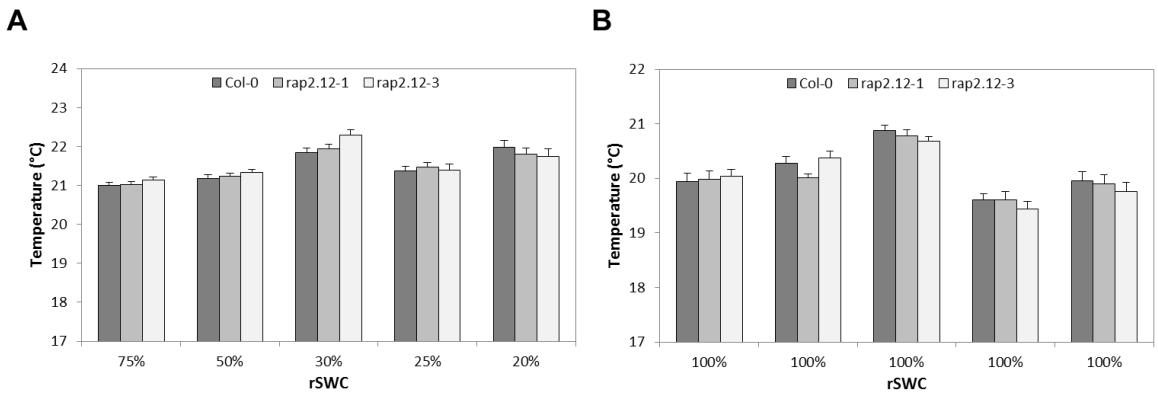


Figure 4.24. Temperature of the *rap2.12* knockouts and Col-0 wild-type measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

4.4.3. Measurement of the stress status of drought-stressed plants

4.4.3.1. Hydrogen peroxide

The levels of H_2O_2 were measured in the wild-type and the two knockouts, *rap2.12-1* and *rap2.12-3* (Figure 4.25). It appeared that more H_2O_2 was produced in the plants subjected to drought than in the watered plants, and that more H_2O_2 was produced in the Col-0 than in the two knockouts under both treatments, however, these results were not significant.

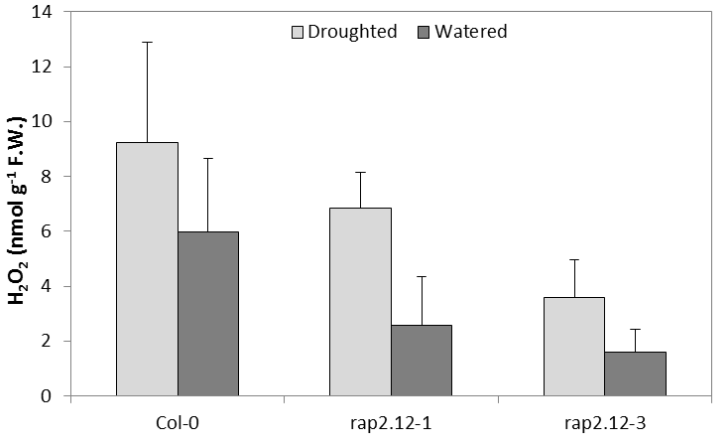


Figure 4.25. Measurement of H_2O_2 levels in the *rap2.12* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

4.4.3.2. Electrolyte leakage

Electrolyte leakage was measured in the wild-type and the two *rap2.12* knockouts to study cellular membrane stability in plants subjected to drought, and in the corresponding watered control plants (Figure 4.26). Surprisingly, the watered plants appeared to have a higher percentage of electrolyte leakage than the droughted plants in all three lines, and this was significant in *rap2.12-3*. However, there appeared to be no difference in electrolyte leakage between the Col-0, *rap2.12-1* and *rap2.12-3* lines under droughted and watered conditions.

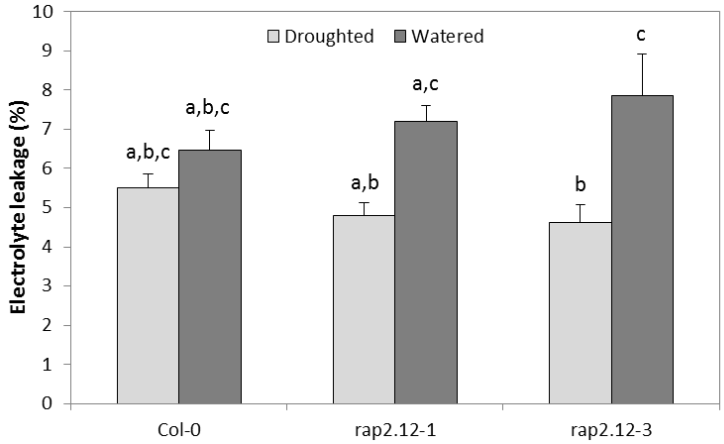


Figure 4.26. Measurement of electrolyte leakage from the *rap2.12* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey HSD test). The experiment was performed once.

4.4.3.3. Chlorophyll and carotenoid content

Drought stress in plants leads to chlorophyll breakdown and subsequently, reduced photosynthetic capacity in the stressed plants (Munné-Bosch and Alegre, 2000; Mafakheri *et al.*, 2010). Hence chlorophyll (a, b and total) and also carotenoid content were measured in plants subjected to drought and watered conditions, in the wild-type and the two knockout lines. Surprisingly, higher levels of chlorophyll and carotenoid content were measured in the droughted plants compared to the watered plants across all three lines (Figure 4.27). There are appeared to be no difference in chlorophyll and carotenoid levels between Col-0 and the knockout lines under drought and watered conditions.

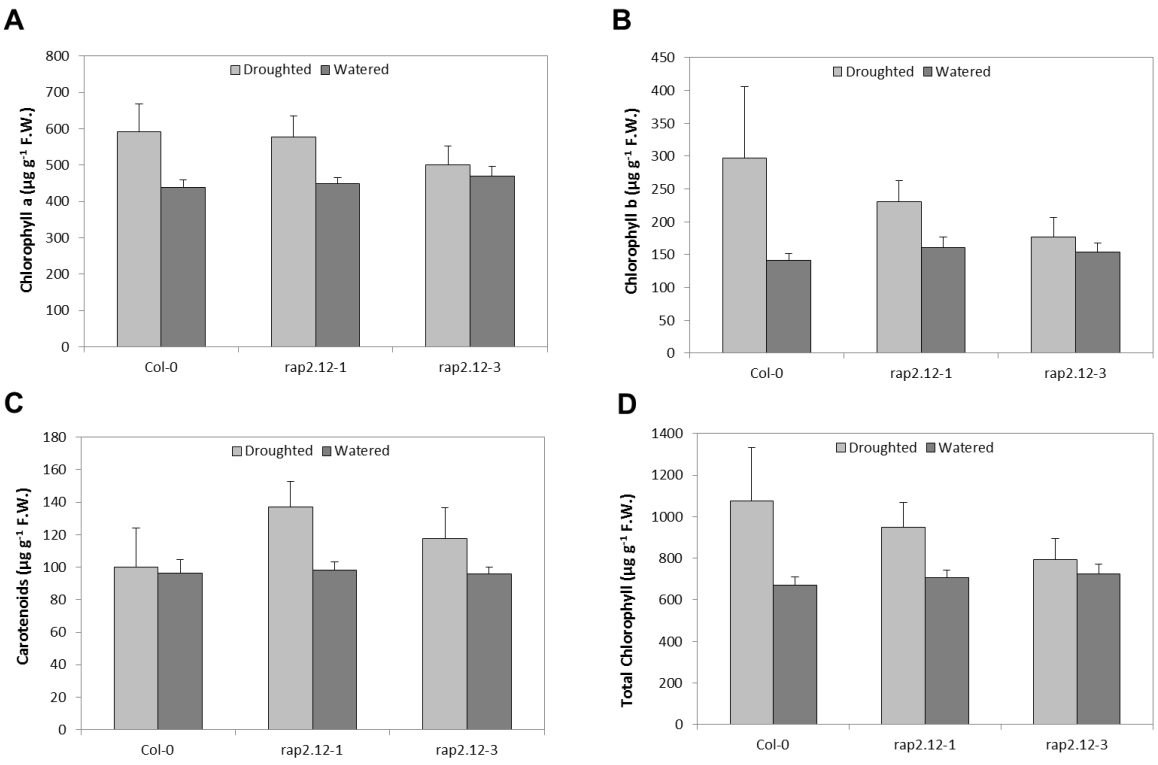


Figure 4.27. Estimation of chlorophyll and carotenoid content of the *rap2.12* knockouts and the Col-0 wild-type. (A) Chlorophyll a (B) Chlorophyll b, (C) Carotenoids and (D) Total Chlorophyll content per fresh weight (F.W.) were measured under both drought and watered conditions. All the values are mean + standard error (n=3). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

4.4.3.4. Anthocyanin content

Anthocyanin levels were measured in the droughted Col-0, *rap2.12-1* and *rap2.12-3* and also in the corresponding watered control plants (Figure 4.28); anthocyanins have an antioxidant role in the cell, during times of stress and levels of anthocyanins are known to increase in response to stress (Neill and Gould, 2003). Higher anthocyanin content was measured in the Col-0 plants subjected to drought, than in the droughted *rap2.12-1* and *rap2.12-3* plants. The amount of anthocyanins detected in the watered plants of all three lines was almost negligible.

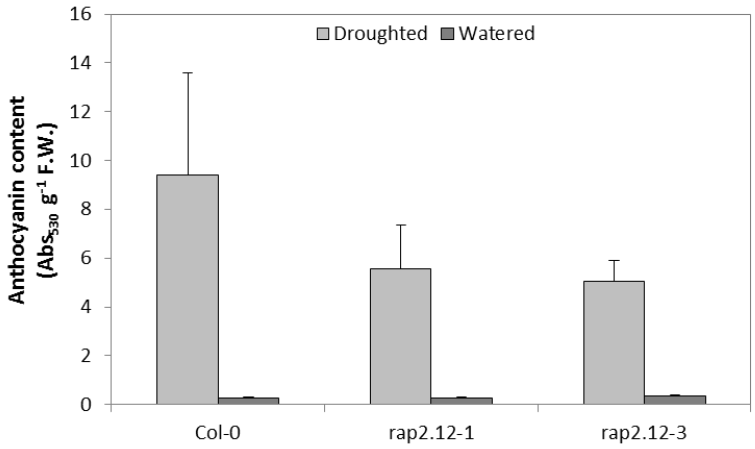


Figure 4.28. Estimation of anthocyanin content of the *rap2.12* knockouts and the Col-0 wild-type under both drought and watered conditions. All the values are mean + standard error (n=3). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.4.4. Measurement of growth and development

4.4.4.1. Flowering time

The flowering time of the wild-type and the two *rap2.12* lines were measured under drought and watered conditions (Figure 4.29). The droughted plants flowered earlier than the watered plants in the case of Col-0 and *rap2.12-1* plants which produced fewer leaves at flowering than the corresponding watered plants, though this difference was not significant (Figure 4.29A). However, *rap2.12-3* plants subjected to drought appeared to flower later than the watered plants. A difference was also seen between the watered *rap2.12-3* and the watered wild-type, though this was also not significant.

In terms of number of days to flowering, the droughted plants flowered earlier in all three lines (Figure 4.29B). The two knockouts flowered earlier than the Col-0 under both drought and watered conditions, and this difference was significant for droughted and watered *rap2.12-3* and also droughted *rap2.12-1*.

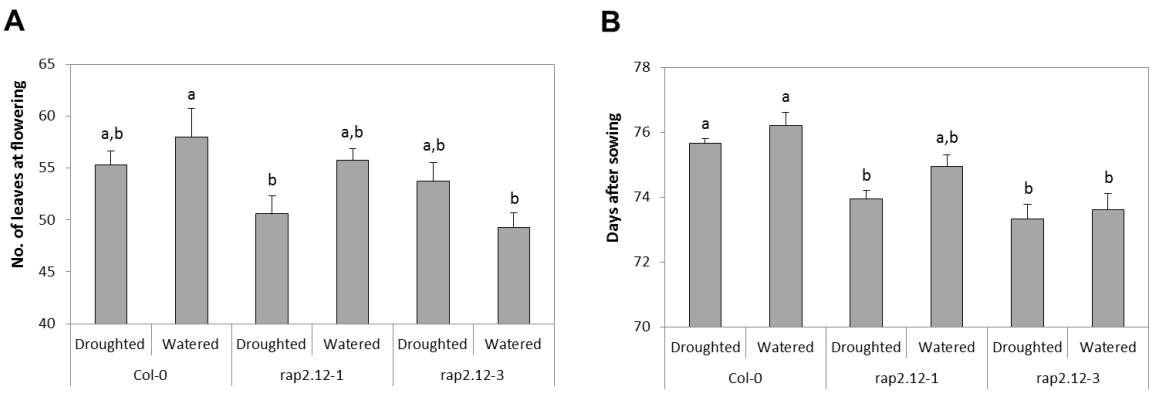


Figure 4.29. Flowering time measured in the Col-0 wild-type and the *rap2.12* knockouts under drought and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=15). Means with different letters indicate significant difference – in Figure A ($P < 0.05$, One-way ANOVA, post-hoc Tukey HSD test); in Figure B ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.4.4.2. Biomass and harvest index

The above-ground biomass of the *rap2.12* knockouts under drought and watered conditions was measured and compared with that of Col-0 plants under the same conditions (Figure 4.30). A small difference in dried rosette weight was observed between the Col-0 plants subjected to drought and those that remained watered (Figure 4.30A). On the other hand, almost no difference in rosette mass was seen between droughted and watered *rap2.12-1* plants, while droughted *rap2.12-3* plants appeared to produce more rosette mass than the watered plants. Overall, it can be seen the two knockouts produced less rosette mass than the wild-type under both conditions, though these differences were not significant.

Droughted Col-0 produced fewer reproductive structures than the watered Col-0 (Figure 4.30B). *rap2.12-1* also produced fewer stalks and siliques under droughted conditions than under watered conditions, while there appeared to be no difference due to the two treatments in the *rap2.12-3* line.

Less seed was produced in droughted Col-0 plants than in the watered plants (Figure 4.30C). Significantly more seed was produced in *rap2.12-1* subjected to drought than in droughted Col-0. However, there was no difference in seed yield between droughted and

watered *rap2.12-1* plants. The seed yield of droughted *rap2.12-3* plants was significantly less than that produced by watered *rap2.12-3* plants. Droughted *rap2.12-3* plants produced similar levels of seed as droughted Col-0, and less than that produced by the droughted *rap2.12-1* plants, while watered *rap2.12-3* plants produced more seed than watered Col-0 and *rap2.12-1* plants.

The total biomass obtained from Col-0 plants subjected to drought was less than that obtained from watered Col-0 (Figure 4.30D). The biomass produced by *rap2.12-1* was similar to that of Col-0 under both treatments, while there was very little difference in total biomass between droughted and watered *rap2.12-3* plants.

There appeared to be no difference in seed harvest index of Col-0 plants subjected to drought was less than that of the watered wild-type plants (Figure 4.31). However, the harvest index was found to be higher in the droughted plants that in the watered plants of *rap2.12-1* plants, while for the *rap2.12-3* line this was found to be the opposite. Under droughted conditions, the harvest indexes of Col-0 and *rap2.12-3* plants were comparable but less than that of *rap2.12-1*, while under watered conditions, the harvest indexes of Col-0 and *rap2.12-1* plants were comparable but less than that of *rap2.12-3*.

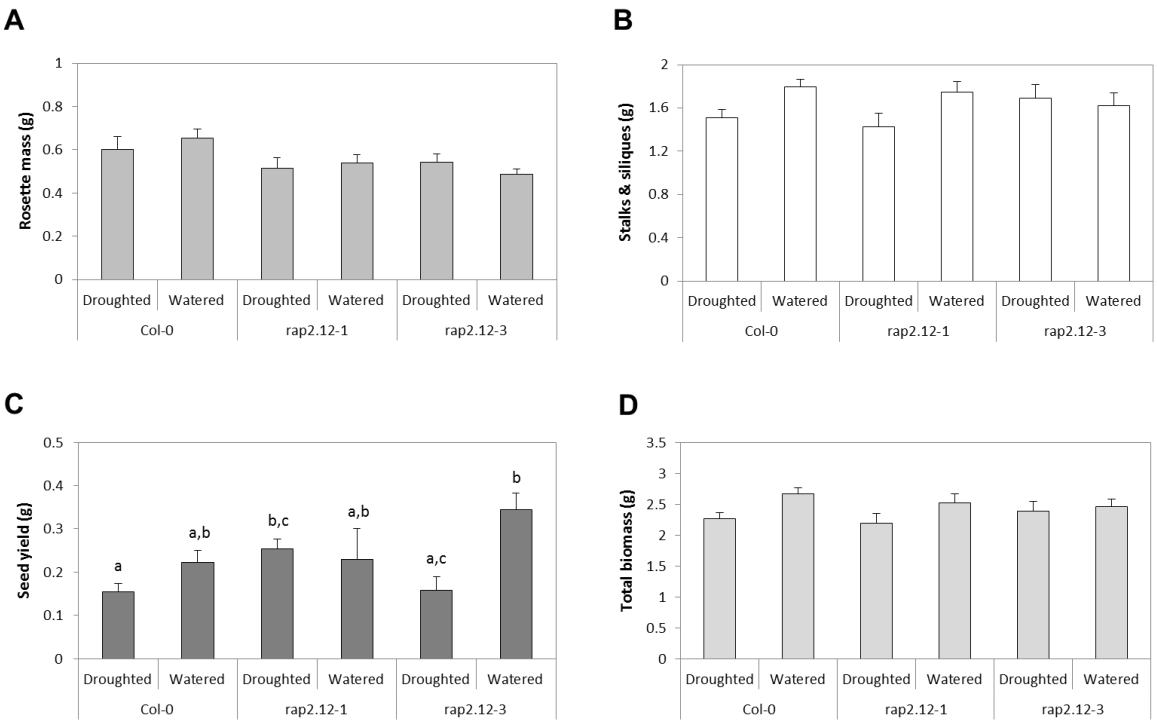


Figure 4.30. Measurement of the vegetative and reproductive biomass for the wild-type and the *rap2.12* knockouts under drought and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). No significant difference was observed in Figures A, B and D ($P > 0.05$, One-way ANOVA). In Figure C, means with different letters indicate significant difference ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

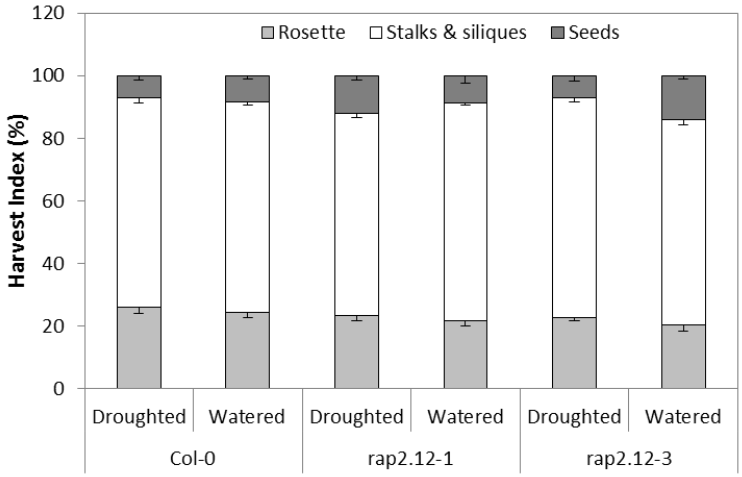


Figure 4.31. Harvest index estimated in the wild-type and the *rap2.12* knockouts under drought and watered conditions. All the values are mean - standard error (n=7).

4.5. Phenotyping and analysis of knockouts of *BHLH038*

4.5.1. Measurement of water use under drought

The drying rates of the two knockouts of *BHLH038* were calculated after a progressive drought till 20% rSWC was reached, and these were compared with the drying rate of the wild-type, Col-0 (Figure 4.32). No difference in drying rate was seen between the wild-type and *bhlh038-2* or *bhlh038-4*. An analysis of the rosette area of the three lines showed that there was no difference in their rosette areas when the drought commenced (Figure 4.33).

The rLWC of Col-0, *bhlh038-2* and *bhlh038-4* subjected to drought and control treatments were measured (Figure 4.34). The droughted plants of all three lines showed a significantly lower rLWC than their respective controls, but between them, the rLWC of *bhlh038-2* was lower than that of the Col-0, and the rLWC of *bhlh038-4* was higher than the wild-type. The rLWC of watered Col-0 was higher than that of watered *bhlh038-2* and *bhlh038-4*, and there was a significant difference in rLWC between the two knockouts maintained under watered conditions.

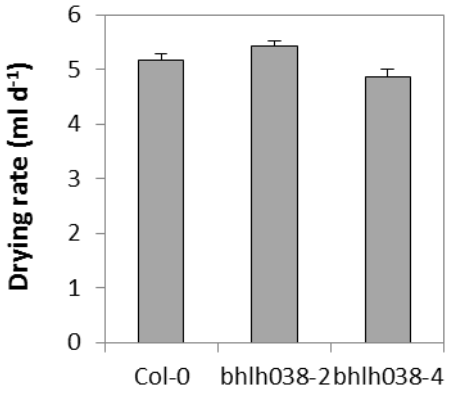


Figure 4.32. Drying rates for *bhlh038* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=10). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

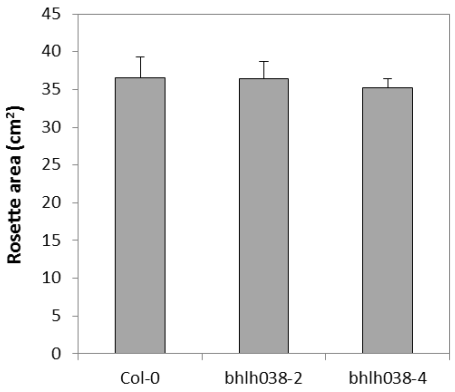


Figure 4.33. Rosette areas for *bhlh038* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=6). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

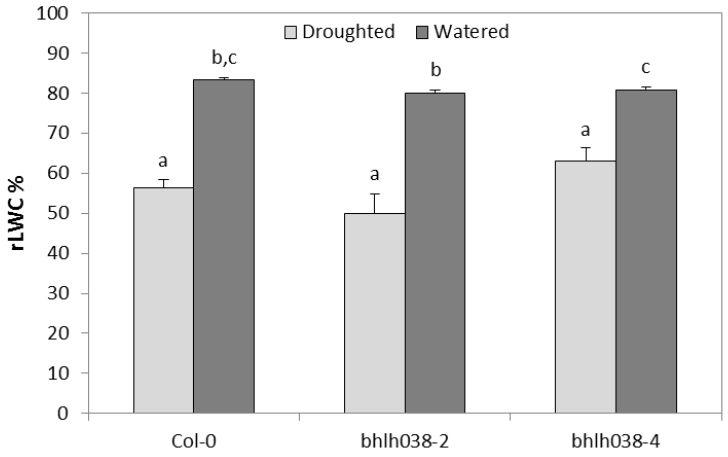


Figure 4.34. Measurement of relative water content (rLWC) in Col-0 wild-type and the *bhlh038* knockouts under drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference (P < 0.05, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

4.5.2. Measurement of photosynthetic performance

g_s and *A* were measured in Col-0, *bhlh038-2* and *bhlh038-4* plants at several points during a progressive drought, and in the corresponding watered control plants (Figure 4.35). g_s of Col-0 plants subjected to drought decreased only towards the end of the drought (Figure 4.35A). g_s of *bhlh038-2* decreased throughout the course of the drought and was consistently lower than that of Col-0 plants. On the other hand, g_s of *bhlh038-4* appeared to increase till 30% rSWC and then remain constant. g_s of droughted plants was less than that of the control watered plants (Figure 4.35B). Over the course of the experiment, g_s of the watered plants continued to increase, and the wild-type had higher g_s than the two knockouts. Overall, it would seem that there was no real difference in g_s between the wild-type and the mutants. There was also no difference in the photosynthetic assimilation rate (*A*) over the course of the drought, or between the three lines (Figure 4.35C).

The average temperatures of the wild-type and the knockouts that were subjected to drought and those maintained under control conditions were also measured (Figure 4.36). There appeared to be no change in the average temperature of the droughted plants over the course of the drought (Figure 4.36A), however the average temperature of the watered plants for all three lines measured at the same time was almost 1 °C lower than that of the droughted plants (Figure 4.36B).

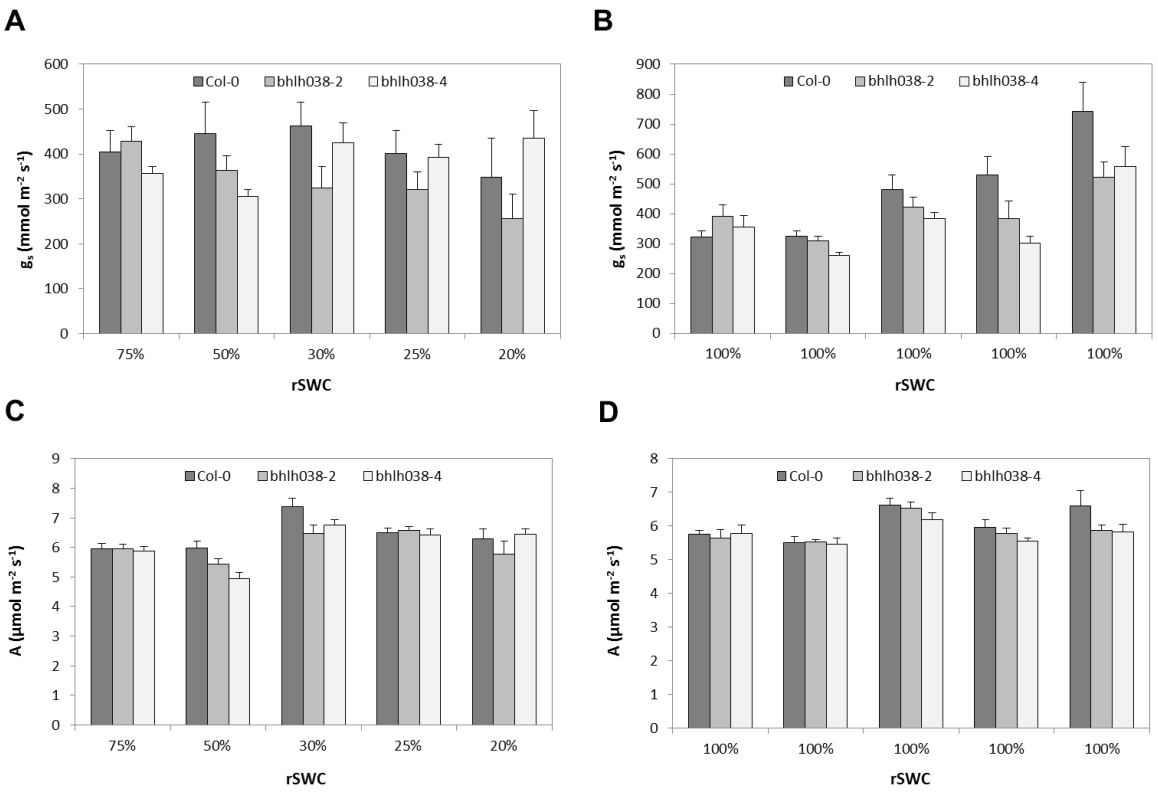


Figure 4.35. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in the *bhlh038* knockouts and Col-0 wild-type at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

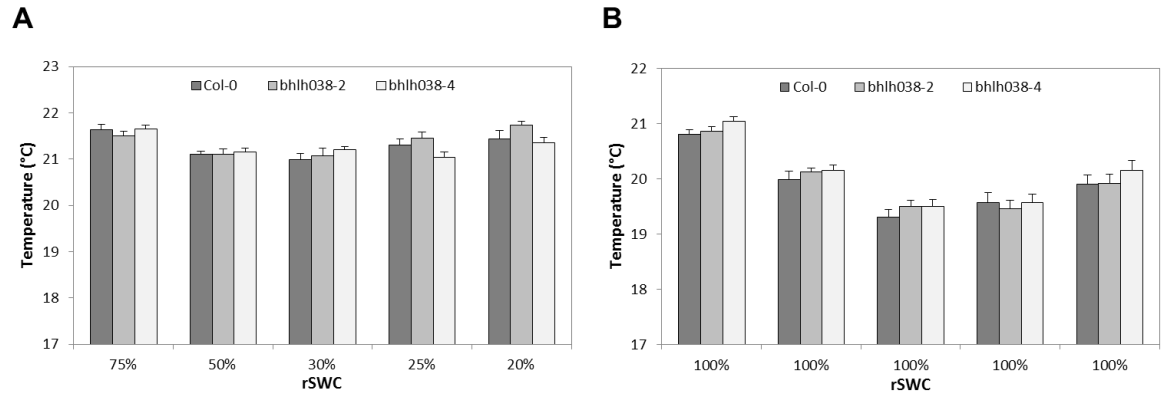


Figure 4.36. Temperature of the *bhlh038* knockouts and Col-0 wild-type measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

4.5.3. Measurement of the stress status of drought-stressed plants

4.5.3.1. Hydrogen peroxide

Levels of the ROS H_2O_2 were measured in the wild-type and the two knockouts of *BHLH038* under drought and watered conditions (Figure 4.37). More H_2O_2 was produced in Col-0 plants subjected to drought than in the watered Col-0, though this was not significant. On the other hand, the two knockouts produced similar levels of H_2O_2 under both droughted and watered conditions. Both *bhlh038-2* and *bhlh038-4* produced less H_2O_2 under droughted conditions than the drought-stressed wild-type plants, although this difference was not significant, while there was no difference in H_2O_2 in the watered plants of all three lines.

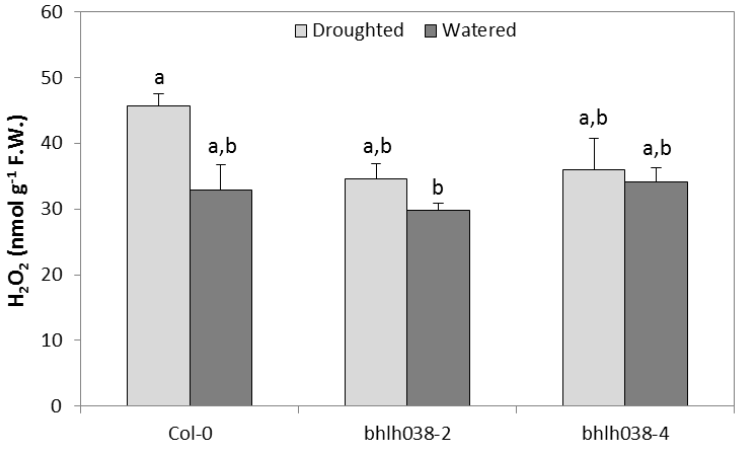


Figure 4.37. Measurement of H_2O_2 levels in the *bhlh038* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

4.5.3.2. Electrolyte leakage

Electrolyte leakage was measured in the three lines under both drought and watered conditions, and these were compared with each other as shown in Figure 4.38. There appeared to be no difference in electrolyte leakage in Col-0 between the drought and watered plants. Lower electrolyte leakage was measured in droughted *bhlh038-2* plants, while droughted *bhlh038-4* plants had a higher percentage of electrolyte leakage than the

wild-type. The watered *bhlh038* knockouts showed a higher electrolyte leakage than the respective droughted plants, and also the watered wild-type.

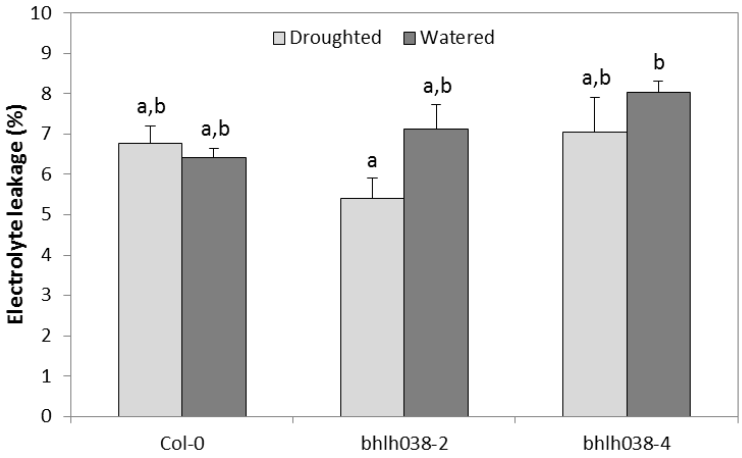


Figure 4.38. Measurement of electrolyte leakage from the *bhlh038* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

4.5.3.3. Chlorophyll and carotenoid content

On examining the chlorophyll and carotenoid content in Col-0, *bhlh038-2* and *bhlh038-4* plants subjected to drought or watered conditions, it was seen that the droughted Col-0 plants had significantly more chlorophyll content (a, b, and total) than their watered counterparts and also higher levels of carotenoids than the watered plants (Figure 4.39). There was no difference in the levels of chlorophyll and carotenoids in the two knockout lines subjected to droughted and watered conditions. The two knockouts subjected to drought also showed significantly less chlorophyll b and total chlorophyll content than droughted Col-0 plants. There was also a difference in chlorophyll a and carotenoid content between the droughted Col-0 plants and the knockouts, though this was not significant. There was no difference in chlorophyll and carotenoid between the wild-type and the mutants under watered conditions.

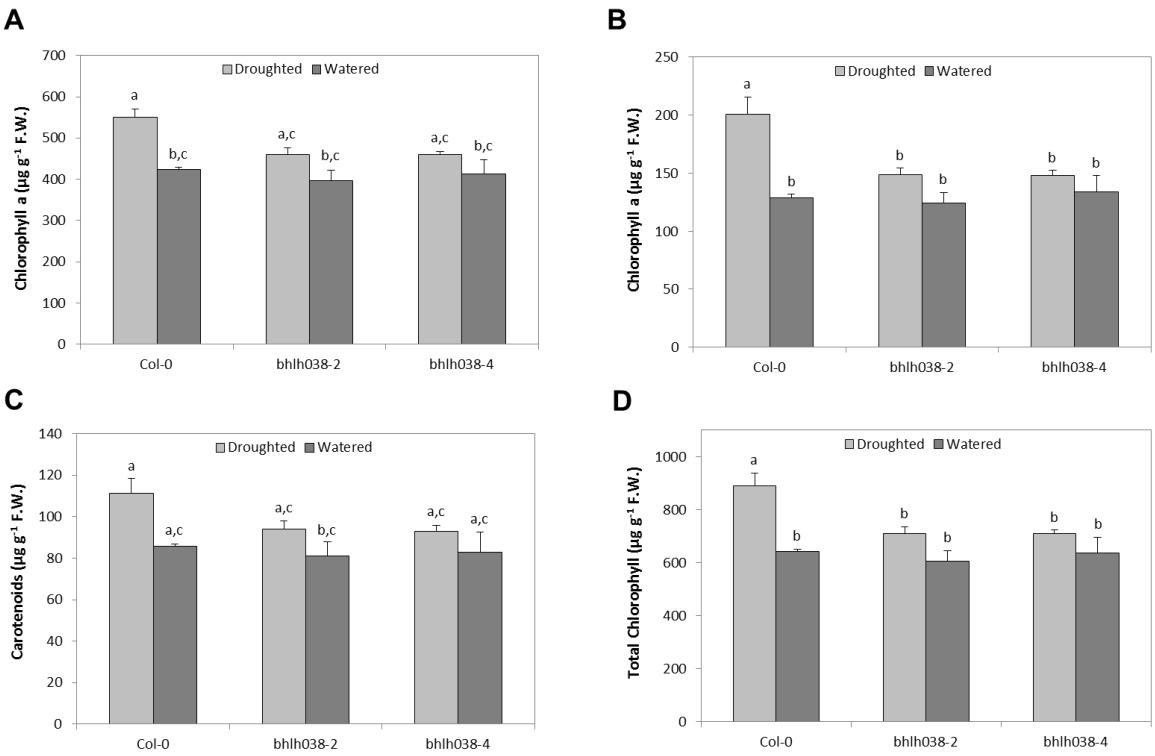


Figure 4.39. Estimation of chlorophyll content of the *bhlh038* knockouts and the Col-0 wild-type. (A) Chlorophyll a (B) Chlorophyll b, (C) Carotenoids and (D) Total Chlorophyll content were measured under both drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

4.5.3.4. Anthocyanin content

The levels of anthocyanin were also measured in the Col-0 and the two *bhlh038* knockouts under drought and watered conditions (Figure 4.40). The droughted plants produced more anthocyanins than the watered plants in the three lines, though this difference was not significant. However, there was no difference in anthocyanin content between the three lines under both droughted and watered conditions.

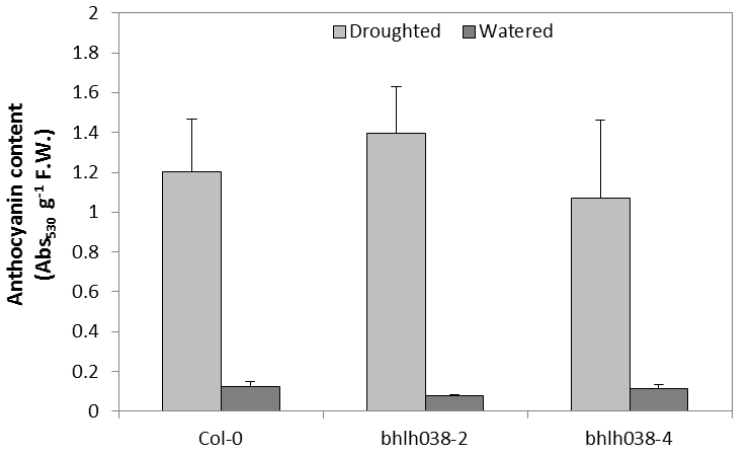


Figure 4.40. Estimation of anthocyanin content of the *bhlh038* knockouts and the Col-0 wild-type under both drought and watered conditions. All the values are mean + standard error (n=3). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.5.4. Measurement of growth and development

4.5.4.1. Flowering time

The flowering times of Col-0, *bhlh038-2* and *bhlh038-4* were determined under drought and watered conditions in terms of number of leaves at flowering and the number of days to flowering (Figure 4.41). In Figure 4.43A, it can be seen that Col-0 plants subjected to drought flowered earlier than the watered Col-0 plants, though this was not significant. However, there was no difference in flowering time between drought-stressed and well-watered *bhlh038-2* plants, while droughted *bhlh038-4* plants flowered much earlier than the watered plants. Also under both treatments, the two *bhlh038* knockouts appeared to flower earlier than the Col-0.

The opposite was seen when the number of days to flowering were analysed (Figure 4.41B). It appeared that the flowering time (in terms of number of days) of the droughted Col-0 plants was later than that of the watered plants, as was the case for droughted *bhlh038-4* plants. On the other hand, *bhlh038-2* plants subjected to drought appeared to flower earlier than the watered plants. However, the two *bhlh038* knockouts appeared to flower earlier than the Col-0 plants of the same treatment, and this was significant between Col-0 and *bhlh038-2* plants.

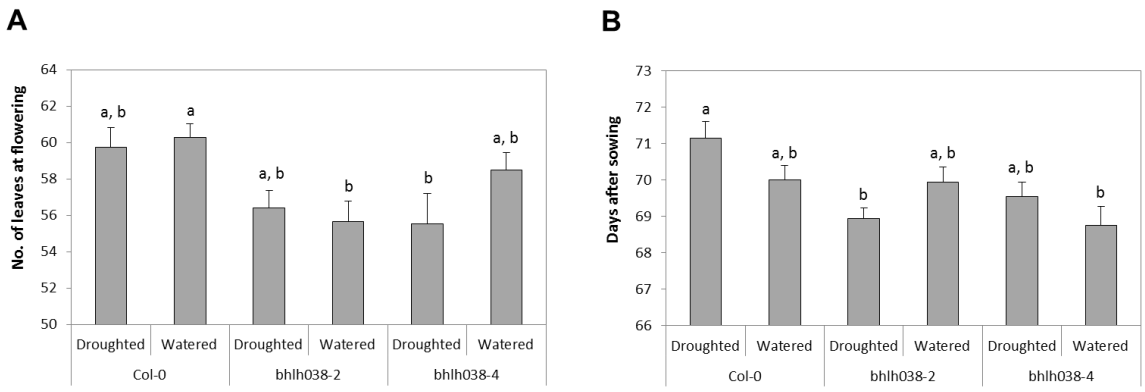


Figure 4.41. Flowering time measured in the Col-0 wild-type and the *bhlh038* knockouts under drought and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=10). Means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Gabriel test). The experiment was performed once.

4.5.4.2. Biomass and harvest index

The vegetative and reproductive biomass were measured in drought-stressed and well-watered Col-0, *bhlh038-2* and *bhlh038-4* plants (Figure 4.42). Figure 4.42A shows that there was no difference in rosette mass between the plants subjected to drought and those that were maintained under watered conditions for all three lines. There was also no difference in rosette mass between the wild-type and the knockout lines.

Figure 4.42B shows the dried mass of reproductive structures produced by the three lines under droughted and watered conditions. No difference was observed in drought-stressed and well-watered Col-0 plants, while droughted *bhlh038-2* plants produced significantly more reproductive biomass than the watered plants. Droughted *bhlh038-4* plants also produced more stalks and siliques than watered *bhlh038-4* plants, though this was not significant. There was no difference in the amount of biomass produced by each of the three lines under droughted conditions, or between Col-0 and *bhlh038-4* watered plants, however, the watered *bhlh038-2* line produced significantly less biomass than the wild-type.

The seed yield of droughted and watered Col-0, *bhlh038-2* and *bhlh038-4* was also measured (Figure 4.42C). There was no difference in seed yield between droughted and watered plants for Col-0 and *bhlh038-2*, while there appeared to be a difference in seed

yield between the droughted and watered *bhlh038-4* plants, which was not significant. There did not appear to be any difference in the yield of *bhlh038-2* plants compared to Col-0 plants, under both treatments. However, *bhlh038-4* plants appeared to produce less seed than both Col-0 and *bhlh038-2* plants under drought and watered conditions, though this was not significant.

Overall, there was no difference in the total amount of biomass produced by Col-0 or *bhlh038-4* plants between the droughted and watered conditions, while watered *bhlh038-2* plants produced significantly less biomass than the droughted plants and also significantly less biomass than watered Col-0 plants.

An analysis of the harvest index between the wild-type and the knockouts of *BHLH038* under watered and droughted conditions (Figure 4.43) showed no difference between the droughted and watered Col-0 and *bhlh038-2* plants, and also no difference between the two lines. The harvest index of *bhlh038-4* plants was less than that of the Col-0 and *bhlh038-2* plants under both conditions of drought and watered, and the harvest index of droughted *bhlh038-4* plants was more than that of the watered plants.

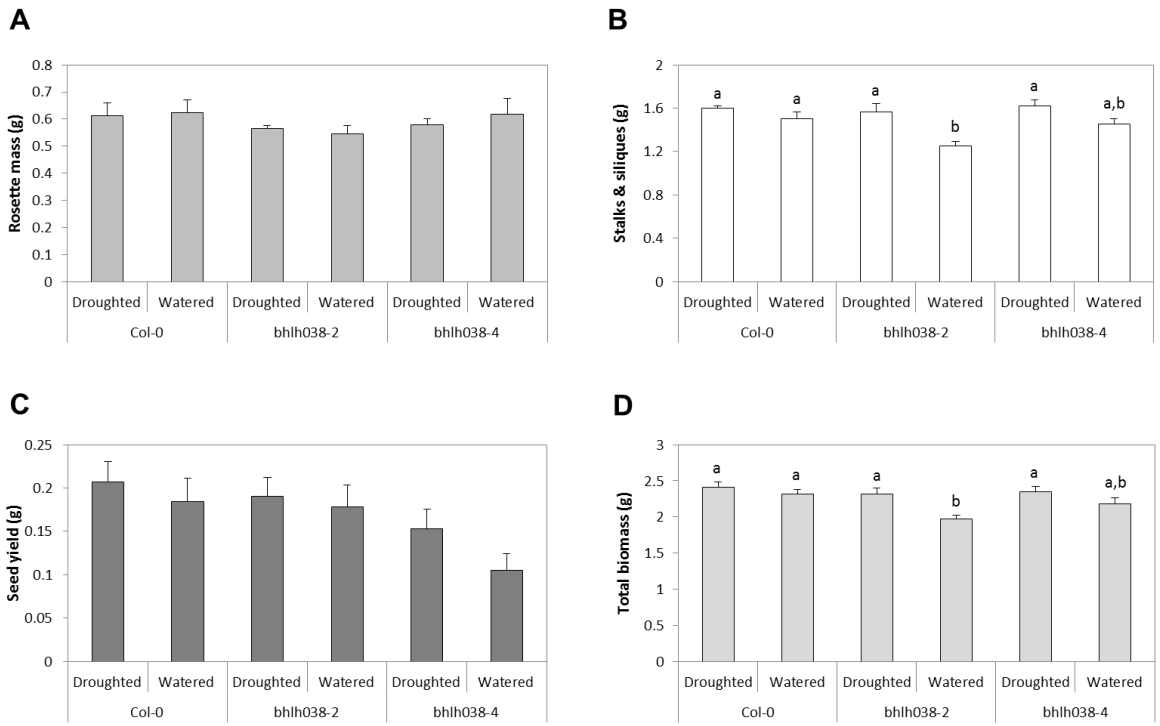


Figure 4.42. Measurement of the vegetative and reproductive biomass for the wild-type and the *bhlh038* knockouts under drought and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). No significant difference was observed in Figures A and C ($P > 0.05$, One-way ANOVA). In Figures B and D, means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

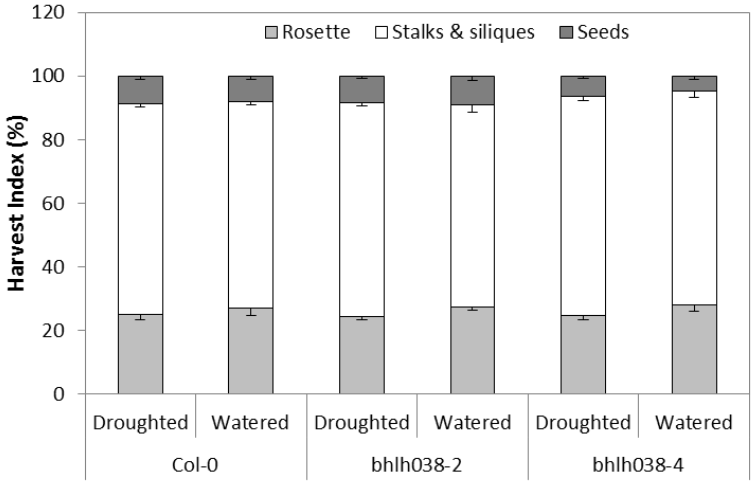


Figure 4.43. Harvest index estimated in the wild-type and the *bhlh038* knockouts under drought and watered conditions. All the values are mean - standard error (n=7).

4.6. Phenotyping and analysis of knockouts of *ANL2*

4.6.1. Measurement of water use under drought

The wild-type Col-0 and the two knockouts of *ANL2*, *anl2-2* and *anl2-4*, were subjected to a progressive drought until 20% rSWC was reached (Figure 4.44). There was no significant difference in the drying rates. An analysis of the rosette area of the three lines (Figure 4.45) showed no difference in rosette area between the wild-type and the two knockouts.

The relative leaf water content (rLWC) was measured in Col-0, *anl2-2* and *anl2-4* plants, subjected to drought and watered conditions (Figure 4.46). The droughted plants had a lower rLWC compared to the watered plants, which was significant for Col-0 and *anl2-4*, but there appeared to be no difference in rLWC between the three lines under watered or droughted conditions.

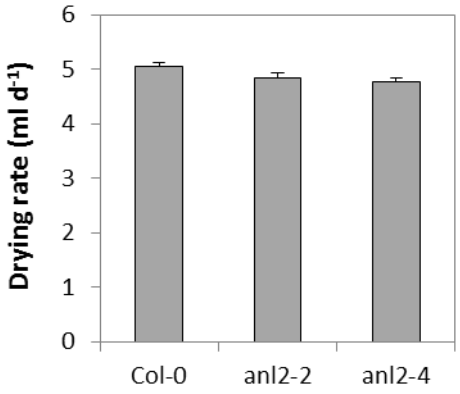


Figure 4.44. Drying rates for *anl2* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=10). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

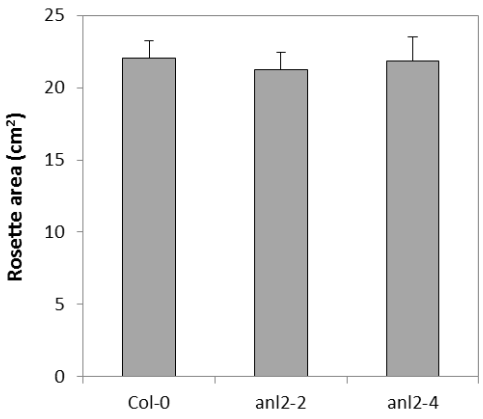


Figure 4.45. Rosette areas for *anl2* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=6). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

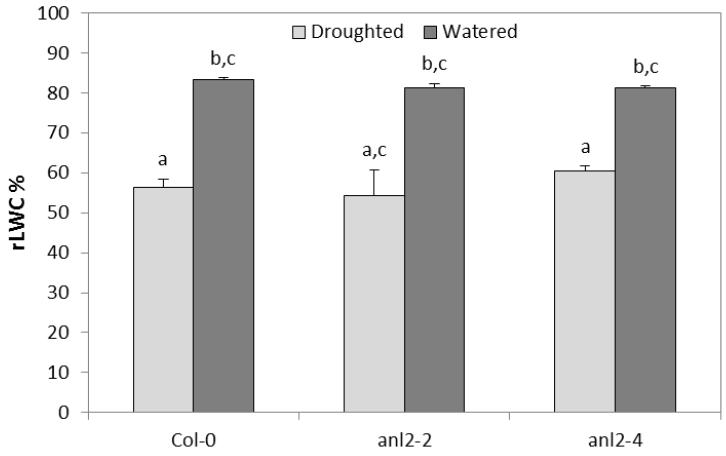


Figure 4.46. Measurement of relative water content (rLWC) in Col-0 wild-type and the *anl2* knockouts under drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference (P < 0.05, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.6.2. Measurement of photosynthetic performance

The stomatal conductance (g_s) and photosynthetic assimilation rate (*A*) of Col-0, *anl2-2* and *anl2-4* were measured under drought and watered conditions (Figure 4.47). g_s was found to decrease in droughted plants from a rSWC of 30% until the end of the drought experiment (Figure 4.47A). A large increase in g_s was also seen from the beginning of the drought to 30% rSWC – however this anomaly was also seen in the data from the watered plants (Figure 4.47B). A comparison of g_s between the droughted plants of all three lines and the watered plants, showed that there was a reduction in g_s towards the end of the drought, compared with no change in g_s in the watered plants. However, there appeared to be no difference in g_s between the three lines.

An increase in *A* was seen over the course of the drought (Figure 4.47C), which was also seen in the watered plants (Figure 4.47D; not taking into consideration the anomalous data measured in the middle of the experiment). Relative to the wild-type, *A* in the knockouts appeared to increase towards the end of the drought, while it did not appear to change under watered conditions.

The average temperature of Col-0, *anl2-2* and *anl2-4* was measured in plants subjected to drought or maintained under watered conditions (Figure 4.48). An increase in the average temperature of the rosette was seen in all three lines only at the end of the drought and this was > 1 °C than the watered control. However, no difference in temperature was seen between the wild-type and the two knockouts under drought or watered conditions.

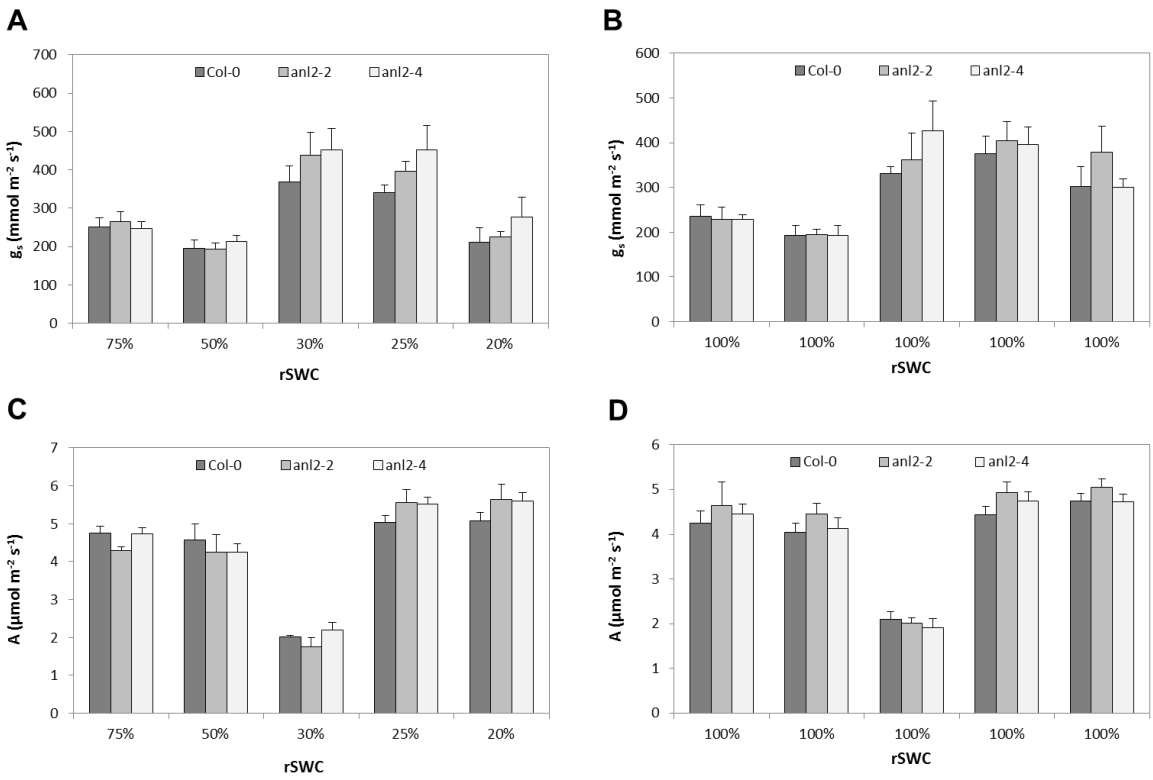


Figure 4.47. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in the *anl2* knockouts and Col-0 wild-type at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

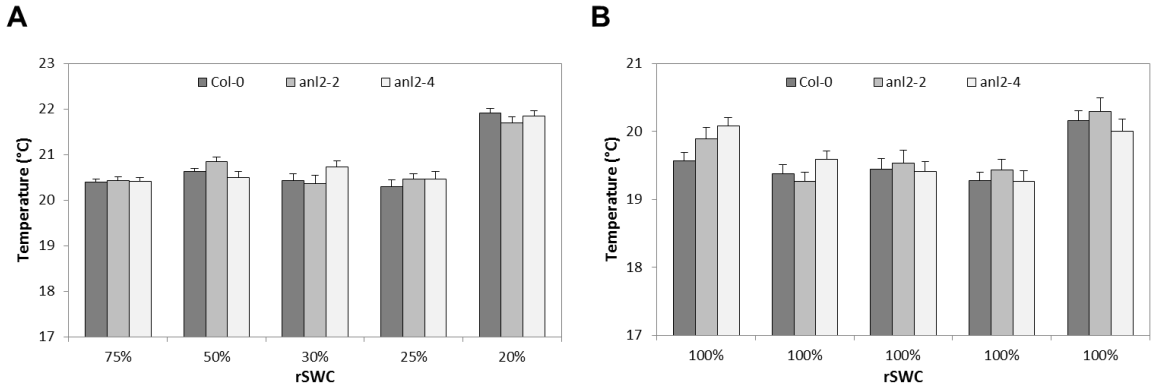


Figure 4.48. Temperature of the *anl2* knockouts and Col-0 wild-type measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once. The experiment was performed once.

4.6.3. Measurement of the stress status of drought-stressed plants

4.6.3.1. Hydrogen peroxide

Hydrogen peroxide (H_2O_2) levels were measured in the Col-0, *anl2-2* and *anl2-4* lines under drought and watered conditions (Figure 4.49). Similar levels of H_2O_2 were detected in drought-stressed and well-watered plants for all three lines, indicating that the plants were not stressed.

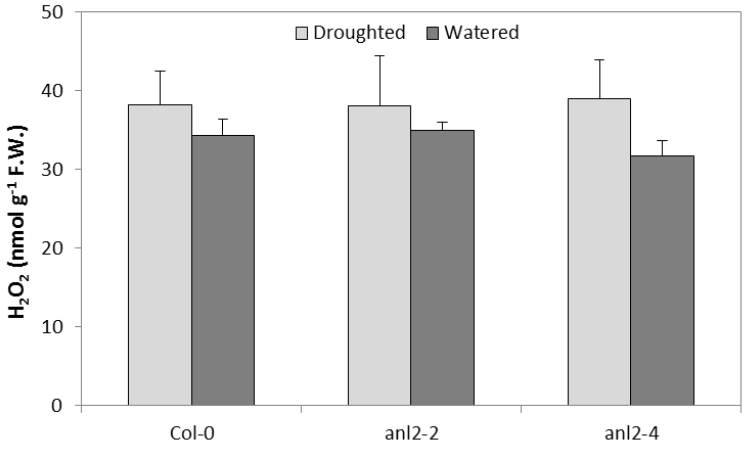


Figure 4.49. Measurement of H_2O_2 levels in the *anl2* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.6.3.2. Electrolyte leakage

Electrolyte leakage was also measured in all three lines under drought and watered conditions (Figure 4.50). No difference in electrolyte leakage was seen between the two treatments in Col-0 and *anl2-2* plants, while higher electrolyte leakage was seen in the watered *anl2-4* plants compared to the droughted plants, though this was not significant.

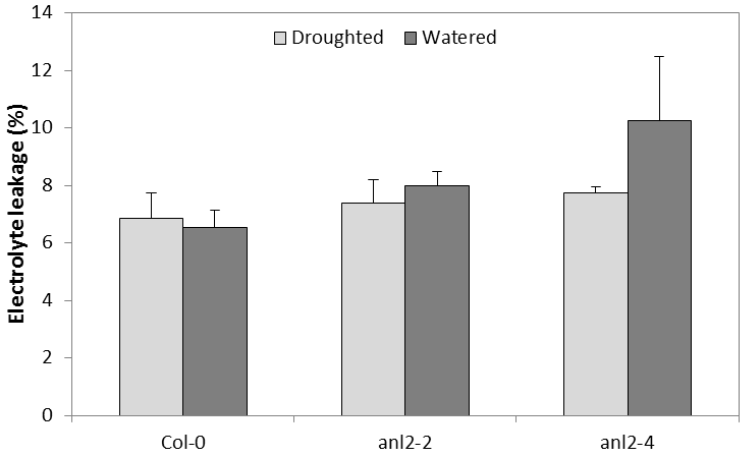


Figure 4.50. Measurement of electrolyte leakage from the *anl2* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.6.3.3. Chlorophyll and carotenoid content

The chlorophyll and carotenoid content of Col-0, *anl2-2* and *anl2-4* were measured in plants subjected to drought and watered conditions (Figure 4.51). No difference in chlorophyll and carotenoid content was seen between droughted and watered Col-0 plants, while a difference was seen in the two knockouts, though this was not significant. Under watered conditions, there appeared to be less chlorophyll (a, b and total) and carotenoids in the knockouts compared to the wild-type, while under droughted conditions, there appeared to be more chlorophyll and carotenoid content in the knockouts compared to the wild-type; however, this difference was not significant.

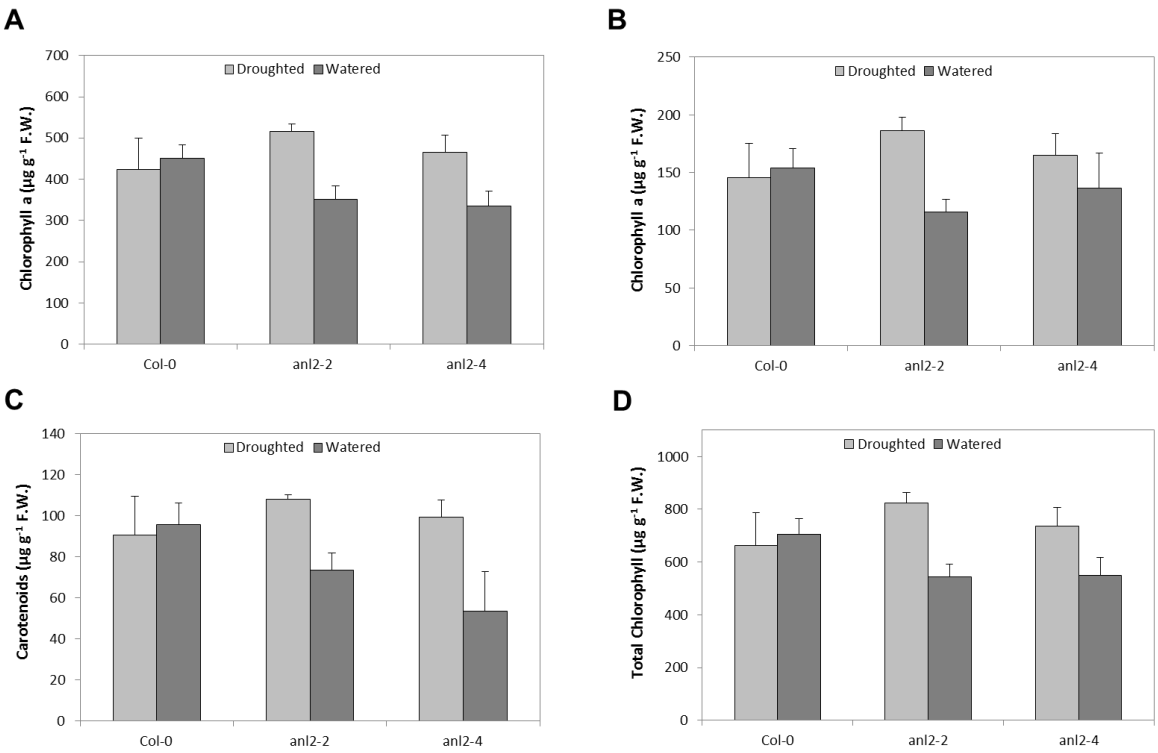


Figure 4.51. Estimation of chlorophyll content of the *anl2* knockouts and the Col-0 wild-type. (A) Chlorophyll a (B) Chlorophyll b, (C) Carotenoids and (D) Total Chlorophyll content were measured under both drought and watered conditions. All the values are mean + standard error (n=3). * indicates significant difference relative to the watered control. No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.6.3.4. Anthocyanin content

Anthocyanin levels were measured in the wild-type and the two knockouts under drought and watered conditions (Figure 4.52). Despite the name, *ANL2* only affects accumulation of anthocyanin in the epidermal cells, but does not affect the biosynthesis of anthocyanins, so it is possible to measure anthocyanin content in these mutants.

Surprisingly higher, though not significant, levels of anthocyanins were measured in the watered plants of Col-0 relative to droughted plants. There was no difference in the levels of anthocyanins between the differently treated *anl2-2* plants, while higher levels of anthocyanins were measured in droughted *anl2-4* plants than in the watered plants, though this was not significant.

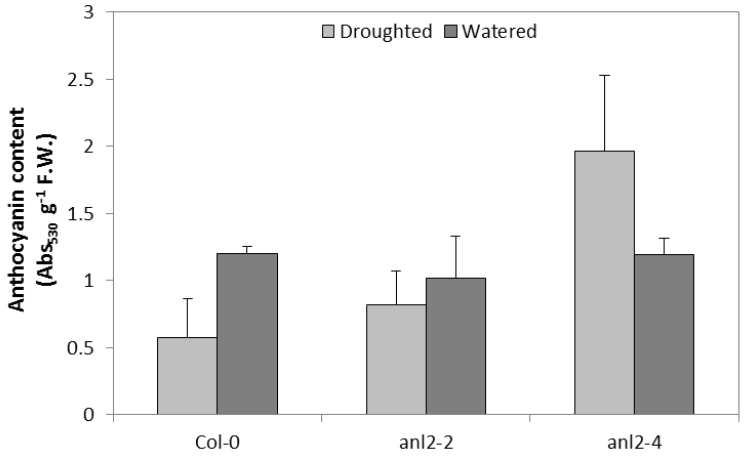


Figure 4.52. Estimation of anthocyanin content of the *anl2* knockouts and the Col-0 wild-type under both drought and watered conditions. All the values are mean + standard error (n=3). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.6.4. Measurement of growth and development

4.6.4.1. Flowering time

The flowering time of the three lines under drought and watered conditions was measured as number of leaves at flowering and number of days from sowing to floral transition (Figure 4.53). No difference was seen in flowering time between Col-0 plants subjected to drought and those that were maintained under watered conditions (Figure 4.53A). The droughted plants of *anl2-2* and *anl2-4* flowered earlier than their respective controls, though this was not significant. Under droughted conditions, there appeared to be no difference in flowering time between the knockouts and the wild-type. Under watered conditions, there was no difference in flowering time between Col-0 and *anl2-2*, while *anl2-4* plants flowered later than these lines, though not significantly.

Col-0 plants subjected to drought took longer to flower, in terms of number of days from sowing, than the watered plants, though this was not significant (Figure 4.53B). There appeared to be no difference in flowering time between the droughted and watered plants of *anl2-2* and *anl2-4*. On the other hand, *anl2-4* plants flowered significantly later than both Col-0 and *anl2-2* under droughted and watered conditions.

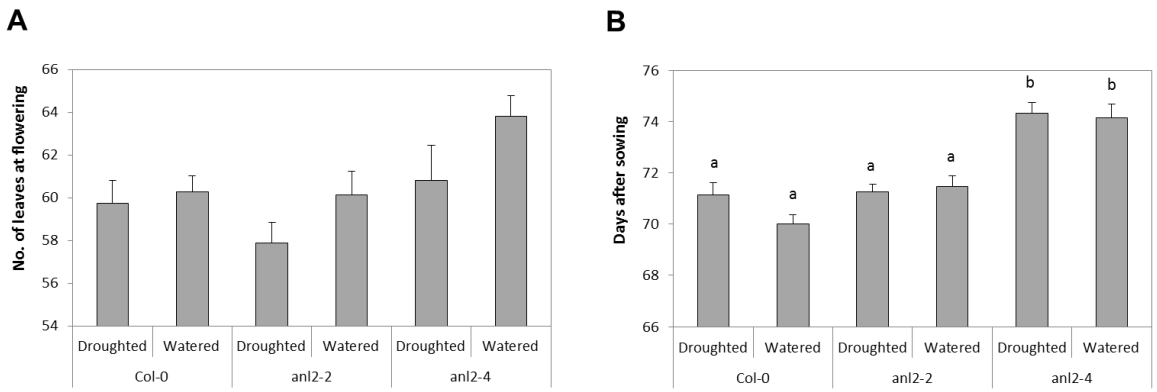


Figure 4.53. Flowering time measured in the Col-0 wild-type and the *anl2* knockouts under drought and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=10). No significant difference was observed in Figure A ($P > 0.05$, One-way ANOVA). In Figure B, means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

4.6.4.2. Biomass and harvest index

An analysis was made of the yield of the wild-type and the two *anl2* knockouts under drought and watered conditions (Figure 4.54). No difference in vegetative biomass was observed between the droughted and watered plants of Col-0, *anl2-2* and *anl2-4*, or between Col-0 and *anl2-2* under droughted and watered conditions (Figure 4.54A). On the other hand, *anl2-4* produced less dried rosette mass than Col-0 and *anl2-2* under both conditions, though this was not significant. In terms of the dried mass of stalks and siliques, it was seen that there was no difference between treatments or lines for the mass of reproductive structures produced (Figure 4.54B).

There was no difference in the amount of seed produced by Col-0 and *anl2-2* plants subjected to drought or watered conditions, while there was no difference in seed yield between the droughted Col-0 and *anl2-2* plants (Figure 4.54C). Under drought stress, *anl2-4* plants appeared to produce more seed than droughted Col-0 and *anl2-2* plants, but this difference was not significant. On the other hand, there was no difference in the seed yield between three lines under watered conditions. Thus, large variations in seed yield were seen between the two knockouts of *ANL2*.

An analysis of the total above-ground biomass produced by the three lines under drought and watered conditions showed there was no difference between the two treatments within a line or between the lines themselves (Figure 4.54D).

There was no difference in the harvest index of Col-0 plants subjected to drought and those that were maintained under watered conditions (Figure 4.55). Under drought conditions, both *anl2* knockouts appeared to have a better seed harvest index than the corresponding watered plants, and similar and better harvest indexes for *anl2-2* and *anl2-4* plants, respectively, compared to the wild-type. Watered *anl2-2* plants had a lower harvest index than watered Col-0 plants, which in turn had a lower harvest index than watered *anl2-4* plants. Also, under both droughted and watered conditions, *anl2-4* plants had a better seed yield relative to above-ground biomass compared to *anl2-2* plants.

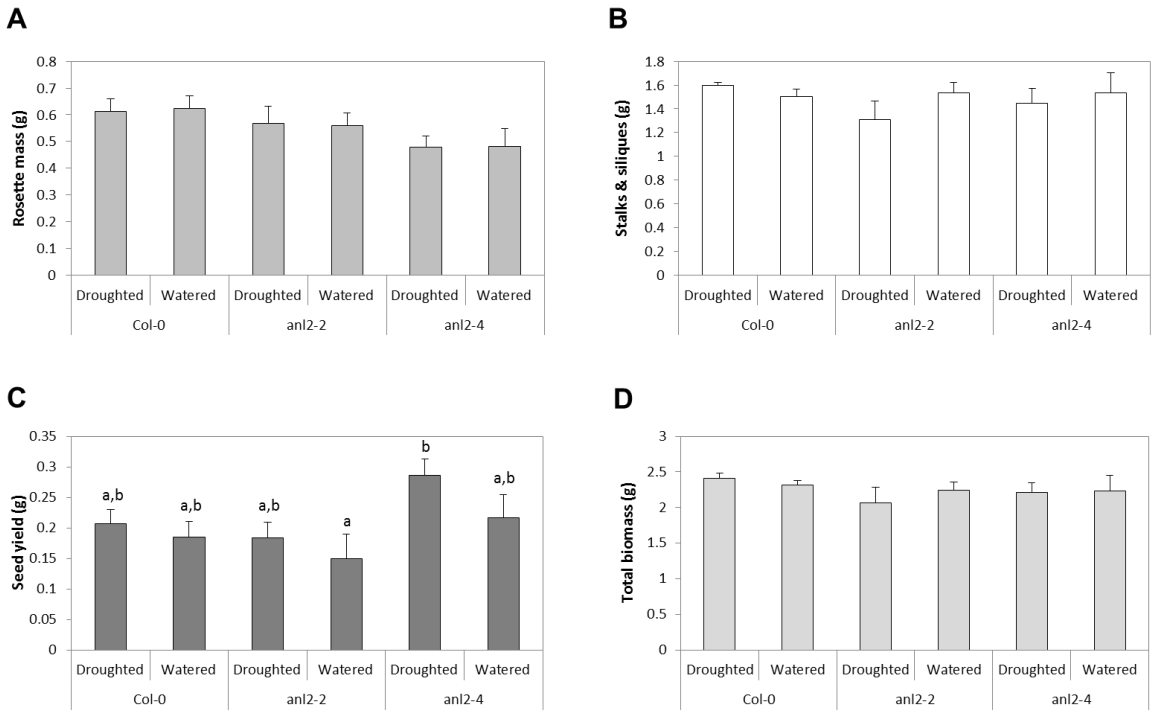


Figure 4.54. Measurement of the vegetative and reproductive biomass for the wild-type and the *anl2* knockouts under drought and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). No significant difference was observed in Figures A, B and D ($P > 0.05$, One-way ANOVA). In Figure C, means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

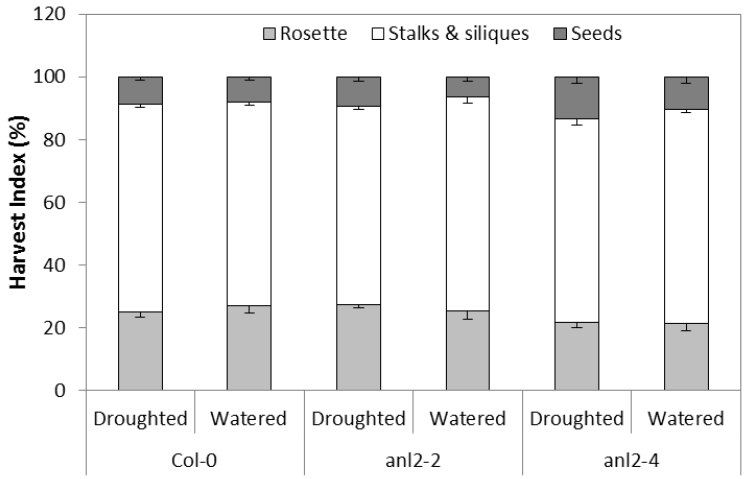


Figure 4.55. Harvest index estimated in the wild-type and the *anl2* knockouts under drought and watered conditions. All the values are mean - standard error (n=7).

4.7. Phenotyping and analysis of overexpressors of *UKTF*

4.7.1. Measurement of water use under drought

The screening of T-DNA insertional mutants of *UKTF* led to the identification of two overexpressors of the gene, named UKTF-3 and UKTF-4; however, no reduced expression mutants could be identified. Six-week old Col-0 plants and both *UKTF* overexpressors were subjected to a progressive drought down to 20%, and the drying rates of both mutants were calculated and compared to the wild-type, Col-0 (Figure 4.56). No difference in drying rate or rosette area (Figure 4.57) was observed between the Col-0 and the two *UKTF* overexpressing lines.

The rLWC of Col-0, UKTF-3 and UKTF-4 plants subjected to drought and watered conditions was also measured (Figure 4.58). The rLWC of the watered plants of all three lines was significantly higher than that of the droughted plants. No difference was seen between the three lines under watered conditions, but there appeared to be a significant difference between the two mutant lines under drought conditions and no difference between the knockouts and the wild-type.

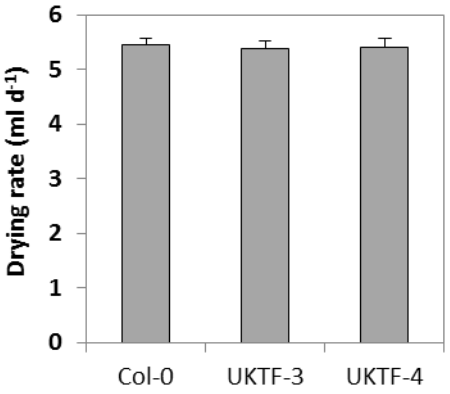


Figure 4.56. Drying rates for *UKTF* overexpressors compared to the wild-type, Col-0. The values are mean + standard error (n=10). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed twice.

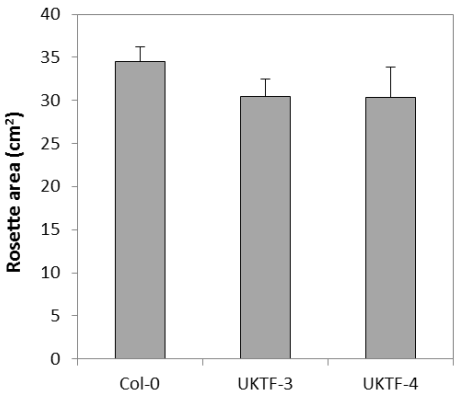


Figure 4.57. Rosette areas for *UKTF* overexpressors compared to the wild-type, Col-0. The values are mean + standard error (n=6). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed twice.

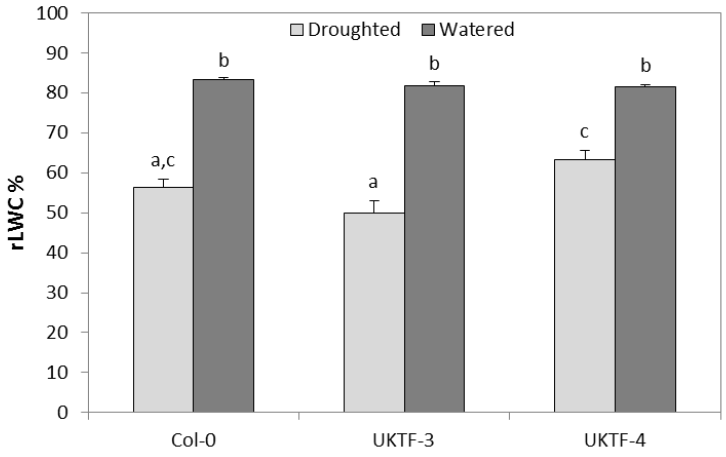


Figure 4.58. Measurement of relative water content (rLWC) in Col-0 wild-type and the *UKTF* overexpressors under drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference (P < 0.05, One-way ANOVA, post-hoc Tukey test). The experiment was performed once.

4.7.2. Measurement of photosynthetic performance

g_s and *A* of Col-0 and the two overexpressing lines of *UKTF* were measured under drought and watered conditions (Figure 4.59). The g_s of droughted Col-0, UKTF-3 and UKTF-4 plants decreased over the course of the drought, and at 20% rSWC was almost 50% of the g_s at the start of the drought period (Figure 4.59A). However, there no was difference in g_s between the wild-type and the mutant lines. *A* also decreased throughout the course of the drought, but there appeared to be no difference in *A* between the three lines (Figure 4.59C).

An increase in the average temperature of the plants subjected to drought was seen in the wild-type and the *UKTF* overexpressors (Figure 4.50). However, no difference in average temperature was seen between the three lines over the course of the drought or in the watered plants.

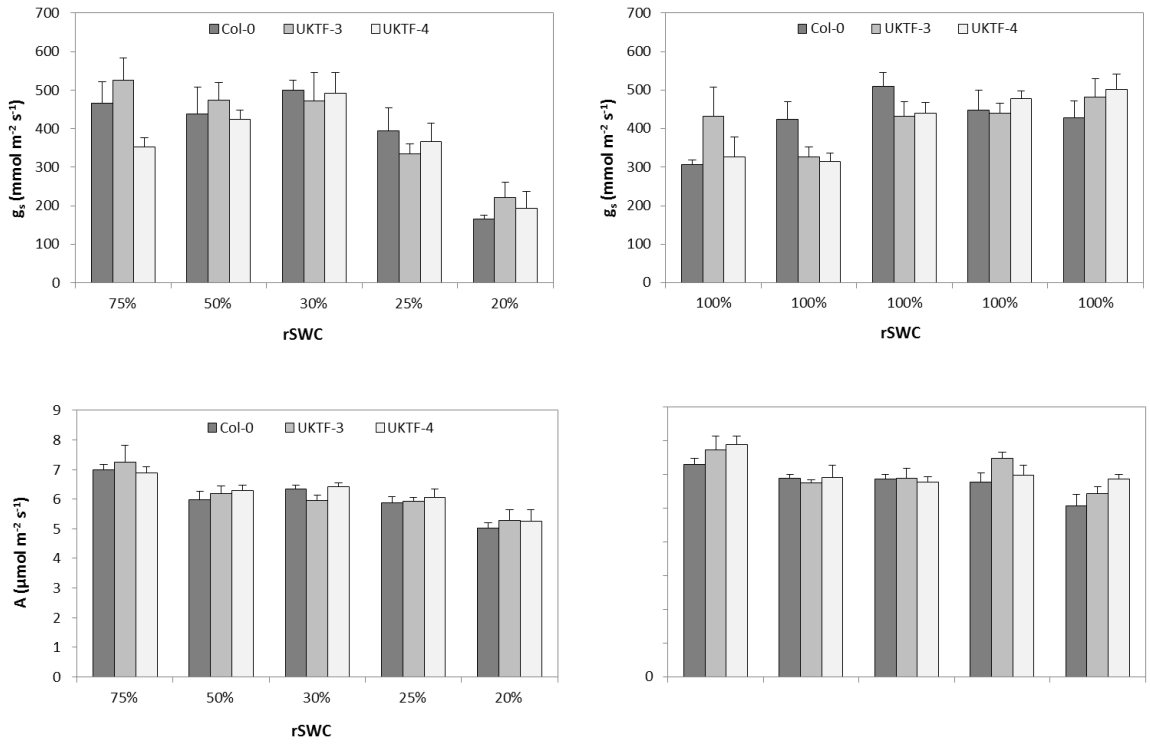


Figure 4.59. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in the *UKTF* overexpressors and Col-0 wild-type at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

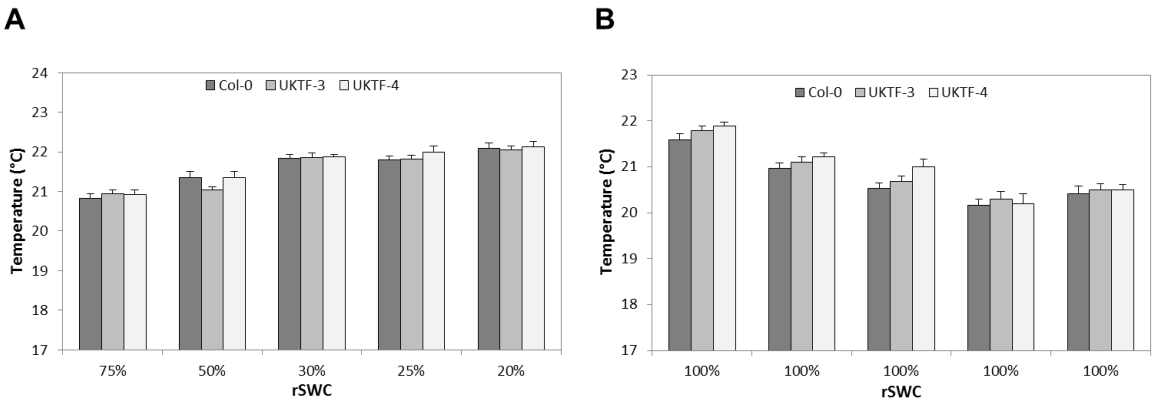


Figure 4.60. Temperature of the *UKTF* overexpressors and Col-0 wild-type measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

4.7.3. Measurement of the stress status of drought-stressed plants

4.7.3.1. Hydrogen peroxide

The levels of H_2O_2 were measured in Col-0, UKTF-3 and UKTF-4 plants under drought and watered conditions (Figure 4.61). Surprisingly, there was no difference in the levels of H_2O_2 between watered and droughted Col-0 plants. Thus, it appears that in this experiment the Col-0 plants did not experience drought stress, unlike the H_2O_2 measurements performed previously with the other mutants. On the other hand, higher levels of H_2O_2 were measured in both mutant lines subjected to drought, compared to watered plants, though this was not significant.

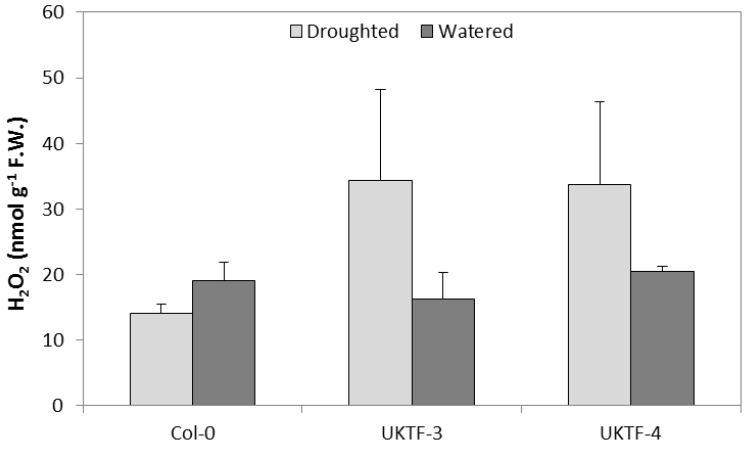


Figure 4.61. Measurement of H_2O_2 levels in the *UKTF* overexpressors and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.7.3.2. Electrolyte leakage

Electrolyte leakage was measured in the wild-type and overexpressors of *UKTF* subjected to drought or maintained under watered conditions to test cellular membrane stability (Figure 4.62). Electrolyte leakage was lower in the droughted plants than in the watered plants in all three lines, though this was not significant. There was no difference in levels of leakage between the three lines under droughted and watered conditions.

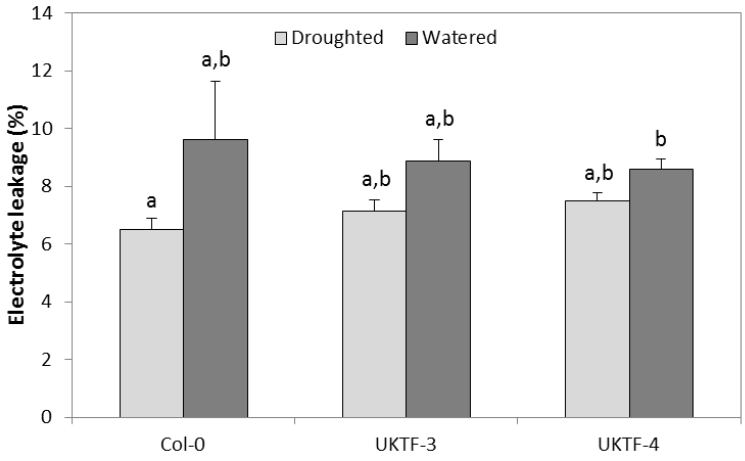


Figure 4.62. Measurement of electrolyte leakage from the *UKTF* overexpressors and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). Means with different letters indicate significant difference (P < 0.05, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.7.3.3. Chlorophyll and carotenoid content

The levels of chlorophyll (a, b and total) and carotenoids were measured in Col-0, UKTF-3 and UKTF-4 plants subjected to drought and watered conditions (Figure 4.63). The levels of chlorophyll and carotenoids were found to be higher in the droughted Col-0 and UKTF-3 plants compared to the watered plants, while no difference was seen in UKTF-4 plants. However, these differences were not significant. No difference in the levels of chlorophyll or carotenoids was seen between the three lines under watered conditions, while under drought less chlorophyll and carotenoids were measured in the overexpressors, though this was not significant.

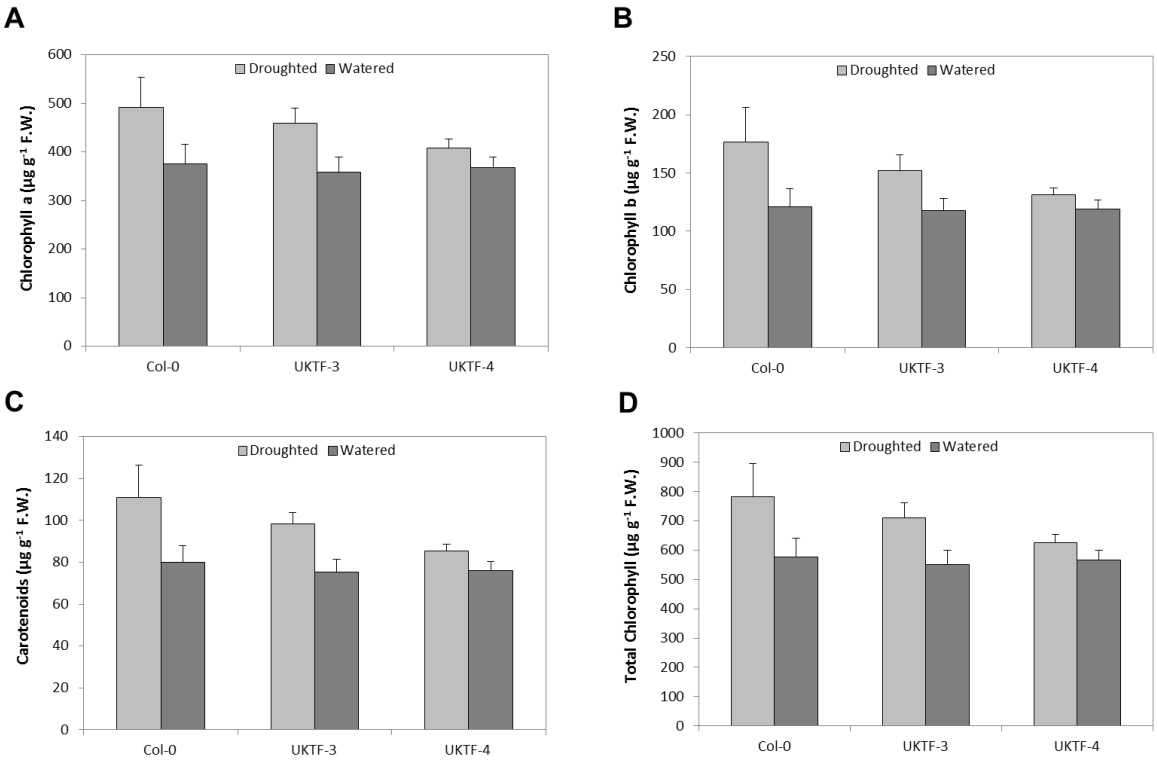


Figure 4.63. Estimation of chlorophyll content of the *UKTF* overexpressors and the Col-0 wild-type. (A) Chlorophyll a (B) Chlorophyll b, (C) Carotenoids and (D) Total Chlorophyll content were measured under both drought and watered conditions. All the values are mean + standard error (n=3). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

4.7.3.4. Anthocyanin content

Anthocyanin levels were also measured in the three lines, in plants that were droughted or watered (Figure 4.64). Higher levels of anthocyanins were observed in the droughted plants in all three lines; however, this difference was very small in the Col-0 and UKTF-4 lines. Higher anthocyanin levels were measured in droughted UKTF-3 plants compared to UKTF-3 watered plants, and also compared to droughted Col-0 and UKTF-4 plants, though these differences were not significant.

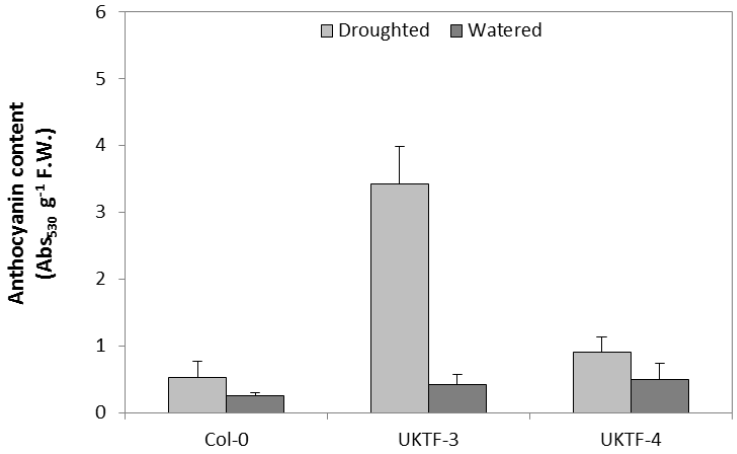


Figure 4.64. Estimation of anthocyanin content of the *UKTF* overexpressors and the Col-0 wild-type under both drought and watered conditions. All the values are mean + standard error (n=3). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

4.7.4. Measurement of growth and development

4.7.4.1. Flowering time

The flowering time of Col-0, UKTF-3 and UKTF-4 plants were measured under drought and watered conditions by measuring the number of leaves at flowering (Figure 4.65A) and the number of days to flowering from sowing (Figure 4.65B). Figure 4.65A shows that droughted plants of all three lines appeared to flower earlier than the watered plants, though this difference was not significant. Under droughted conditions, UKTF-3 plants flowered later than the wild-type plants, while UKTF-4 plants flowered earlier than wild-type plants and earlier than UKTF-3 plants. Under watered conditions, no difference in

flowering time was seen between Col-0 and UKTF-4 plants, while UKTF-3 plants flowered later than both of these.

In terms of number of days to flowering, no difference in flowering time was seen between Col-0 plants under both conditions of drought and watered (Figure 4.65B), while the droughted mutants flowered earlier than the respective watered plants, which was significant for UKTF-3. Under droughted conditions, no difference in flowering time was seen between Col-0 and UKTF-3 plants, while UKTF-4 plants flowered significantly earlier than the wild-type. On the other hand, under watered conditions UKTF-3 plants flowered later than the wild-type, and significantly later than watered UKTF-4 plants.

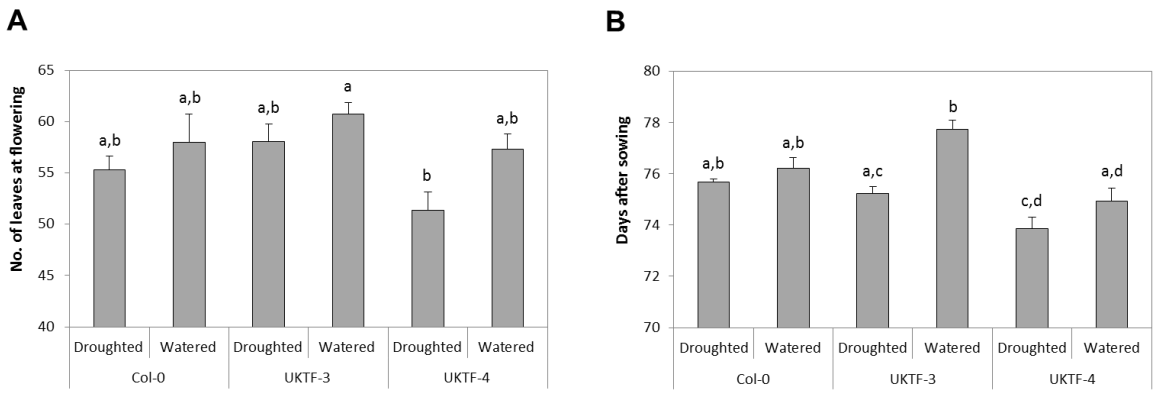


Figure 4.65. Flowering time measured in the Col-0 wild-type and the *UKTF* overexpressors under drought and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=10). Means with different letters indicate significant difference – in Figure A ($P < 0.05$, One-way ANOVA, post-hoc Gabriel test); in Figure B ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.7.4.2. Biomass and harvest index

The vegetative and reproductive biomass produced by the overexpressors of *UKTF* under droughted and watered conditions were measured and compared against the wild-type, Col-0 (Figure 4.66). The rosette biomass of the three lines is shown in Figure 4.66A. No difference was seen in the rosette mass produced by droughted and watered Col-0 and UKTF-3 plants, while droughted UKTF-4 plants produced less rosette biomass than the watered plants, though this was not significant. Droughted UKTF-4 plants also

produced less rosette mass than the Col-0 and UKTF-3 lines, but this was also not significant.

The reproductive structures (stalks and siliques) were also measured in Col-0, UKTF-3 and UKTF-4 subjected to drought and watered conditions (Figure 4.66B). Fewer reproductive structures were produced in droughted Col-0 plants than in the watered plants, though this was not significant. However, there was no difference in amount of reproductive structures produced by the overexpressing lines under both droughted and watered conditions.

The seed yield of the three lines subjected to drought and watered conditions was analysed (Figure 4.66C). Droughted Col-0 plants produced less seed than the watered plants, though not significantly, while there was no difference in seed yield between the droughted and watered UKTF-3 plants. However, more seed was produced in UKTF-4 plants subjected to drought than the watered plants, though this was also not significant. Under droughted conditions the Col-0 plants appeared to produce less seed than the two mutants, while under watered conditions UKTF-4 produced less seed than Col-0 and UKTF-3 plants.

There was a difference in the total biomass produced by Col-0 plants subjected to drought compared to that produced by watered plants (Figure 4.66D), though this was not significant, while there was no difference in total biomass between droughted and watered UKTF-3 and UKTF-4 plants. There was no difference in total biomass between the three lines under droughted conditions, while under watered conditions UKTF-4 plants produced significantly less total biomass than the wild-type.

There was no difference in seed harvest index between droughted and watered Col-0 and UKTF-3 plants (Figure 4.67), while watered UKTF-4 had a lower harvest index than the droughted plants.

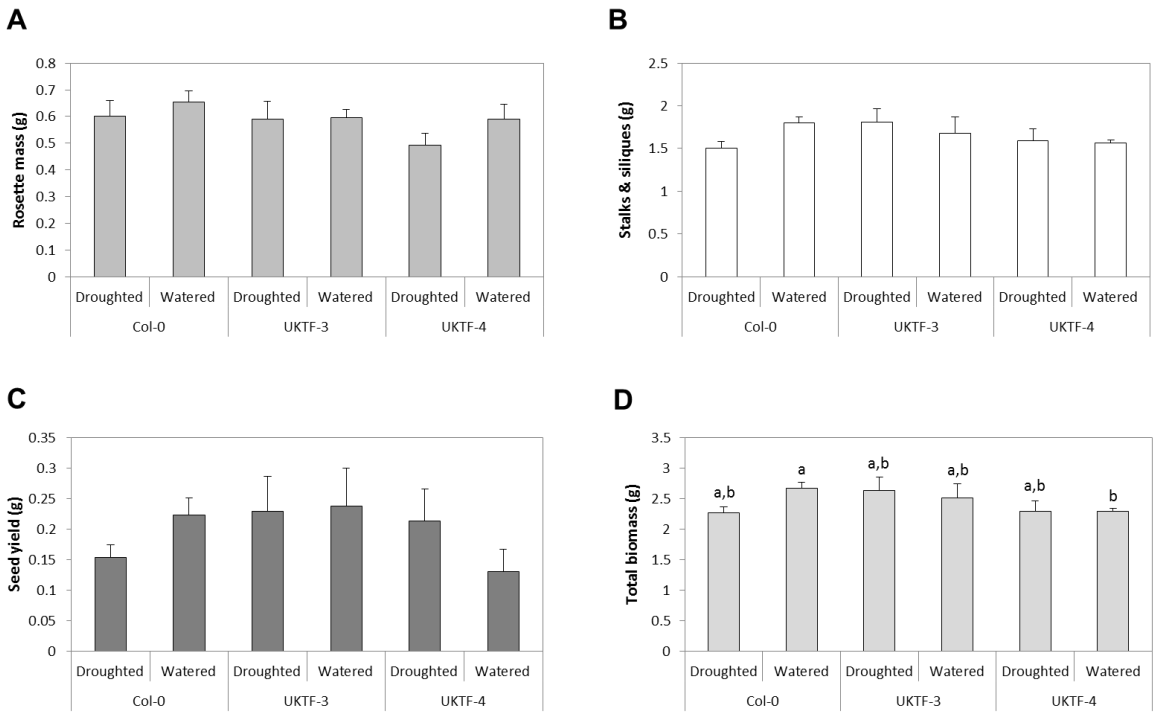


Figure 4.66. Measurement of the vegetative and reproductive biomass for the wild-type and the *UKTF* overexpressors under drought and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). No significant difference was observed in Figures A, B and C ($P > 0.05$, One-way ANOVA). In Figure D, means with different letters indicate significant difference ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

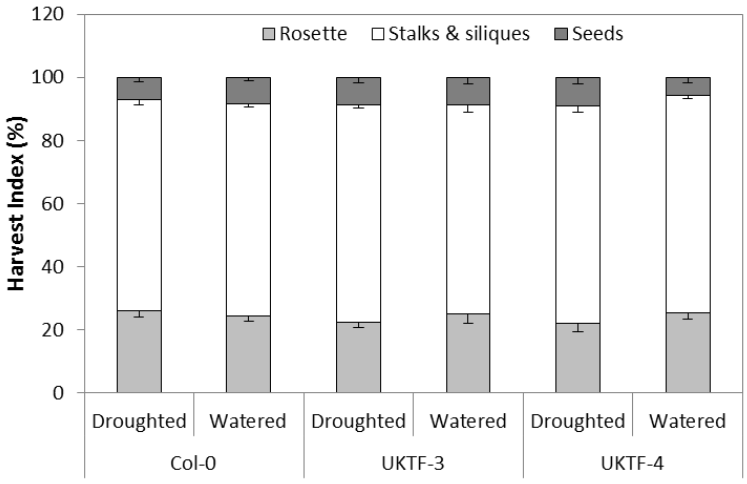


Figure 4.67. Harvest index estimated in the wild-type and the *UKTF* overexpressors under drought and watered conditions. All the values are mean - standard error (n=7).

4.8. Phenotyping and analysis of knockouts of *POZ*

4.8.1. Measurement of water use under drought

Two knockouts of *POZ* were used to analyse a potential role of the gene in the drought response. The wild-type, Col-0, and the two knockouts, *poz-2* and *poz-3*, were subjected to a progressive drought until 20% rSWC was reached, and their drying rates were subsequently calculated (Figure 4.68). No difference in the drying rates was found between the three lines; also, no difference in the rosette areas was found between the wild-type and the mutant lines (Figure 4.69).

The rLWC of the wild-type and the two knockouts of *POZ* were measured in plants subjected to drought and watered conditions (Figure 4.70). A significantly lower rLWC was seen in the droughted plants of Col-0, while the difference in rLWC between the droughted and watered knockouts was not significant. No difference was seen between the three lines under droughted conditions, while under watered conditions the rLWC of the wild-type was significantly higher than that of *poz-2*.

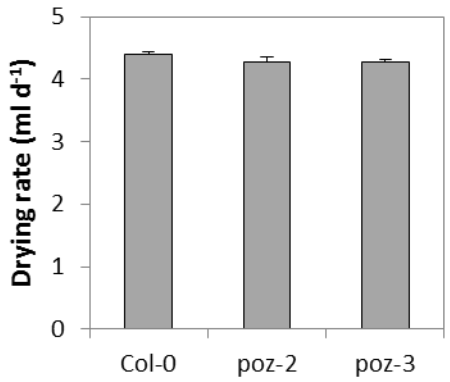


Figure 4.68. Drying rates for *poz* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=10). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

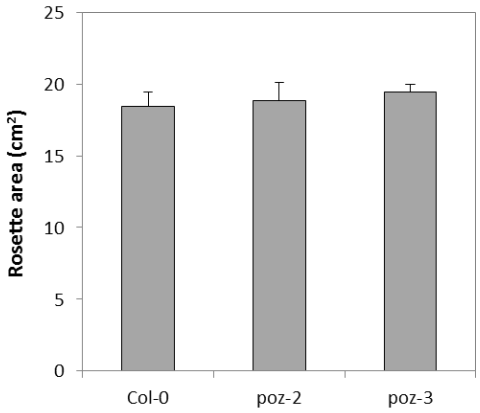


Figure 4.69. Rosette areas for *poz* knockouts compared to the wild-type, Col-0. The values are mean + standard error (n=6). No significant difference was observed (P > 0.05, Student's t-test). The experiment was performed twice.

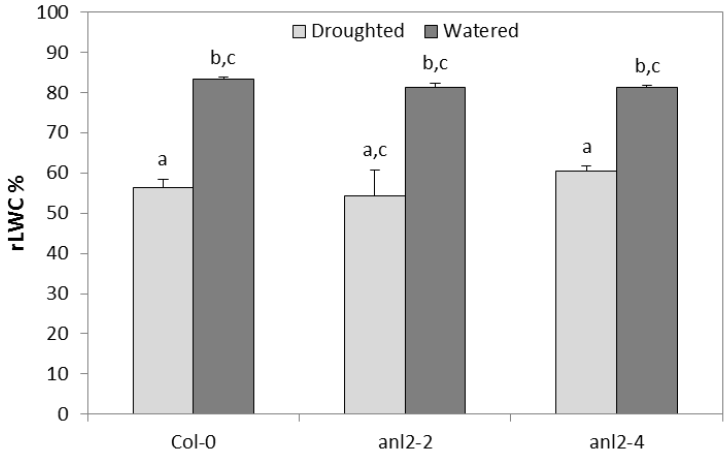


Figure 4.70. Measurement of relative water content (rLWC) in Col-0 wild-type and the *poz* knockouts under drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference (P < 0.05, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.8.2. Measurement of photosynthetic performance

The stomatal conductance (g_s) and carbon assimilation rate (A) of the wild-type and the two knockouts of *POZ* were measured under drought and watered conditions (Figure 4.71). g_s of all three lines decreased as the drought progressed, but there appeared to be no difference in g_s between the three lines under drought or watered conditions (Figures 4.71A and B). *A* did not appear to decrease in all three lines over the course of the drought (Figure 4.71C).

The average temperatures of the Col-0, *poz-2* and *poz-3* lines subjected to drought were measured and compared to the plants maintained under watered conditions (Figure 4.72). An increase in temperature of about 1 °C was observed in the all three lines between the start and the end of the drought (Figure 4.72A), which was not seen in the watered plants (Figure 4.72B). However, no difference in the average temperature was observed between the three lines under both drought and watered conditions.

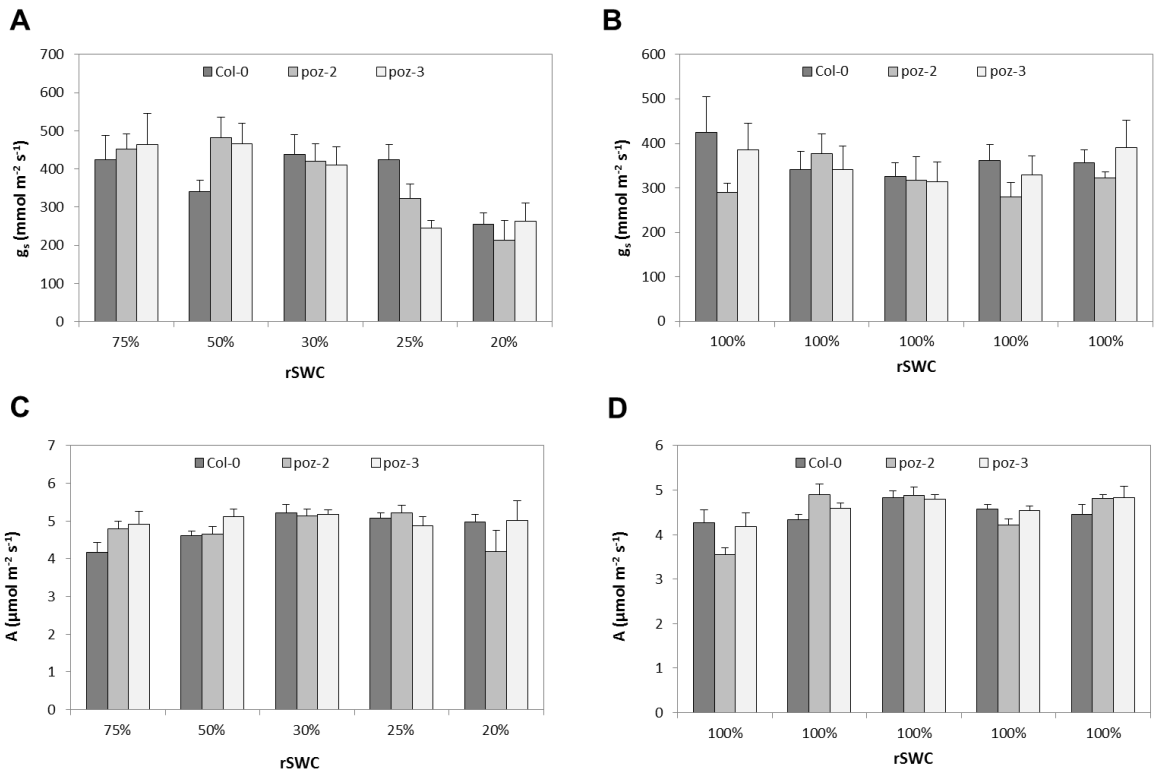


Figure 4.71. Measurement of stomatal conductance (g_s) and photosynthetic assimilation rate (A) in the *poz* knockouts and Col-0 wild-type at regular intervals during the drought stress. (A) The stomatal conductance of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The stomatal conductance of the corresponding watered control plants (n=6). (C) The assimilation rate of the plants subjected to drought at different relative soil water content (rSWC; n=6). (D) The assimilation rate of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

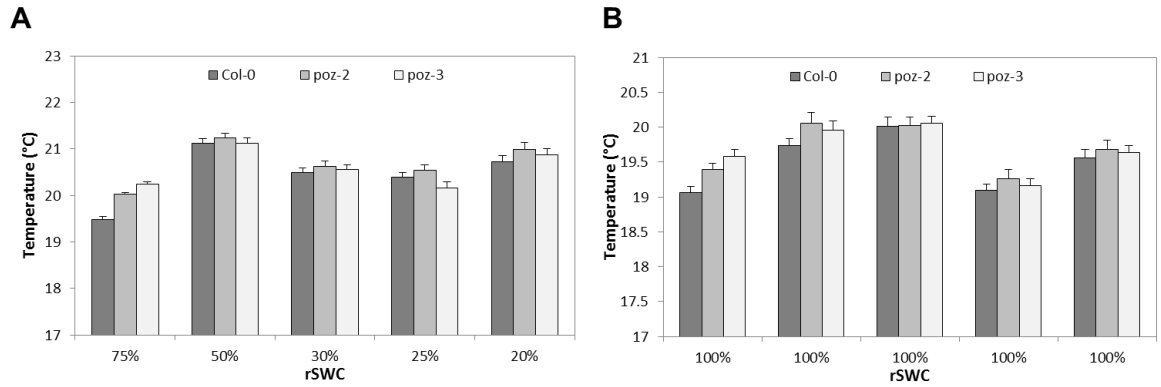


Figure 4.72. Temperature of the *poz* knockouts and Col-0 wild-type measured at regular intervals throughout the duration of the drought. (A) The temperature of the plants subjected to drought at different relative soil water content (rSWC; n=6). (B) The temperature of the corresponding watered control plants (n=6). All the values are mean + standard error. The experiment was performed once.

4.8.3. Measurement of the stress status of drought-stressed plants

4.8.3.1. Hydrogen peroxide

Levels of H_2O_2 were measured in the wild-type and the knockouts of *POZ* subjected to drought and watered conditions (Figure 4.73). No difference in the levels of H_2O_2 was seen between all three lines or between droughted and watered conditions.

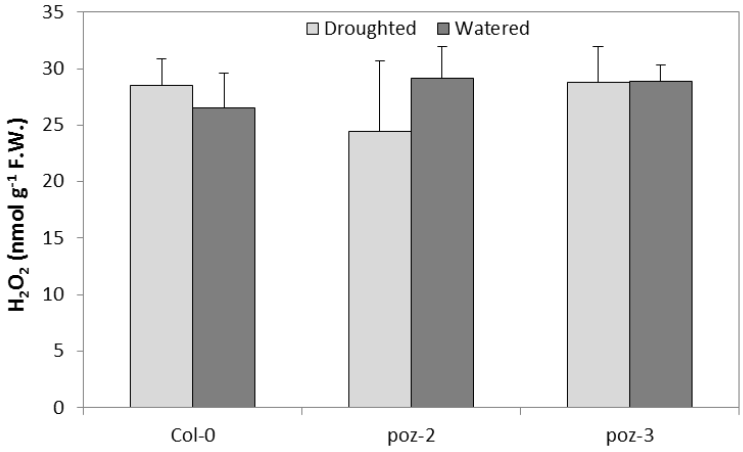


Figure 4.73. Measurement of H_2O_2 levels in the *poz* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

4.8.3.2. Electrolyte leakage

Electrolyte leakage was measured in Col-0, *poz-2* and *poz-3* under droughted and watered conditions (Figure 4.74). No difference in electrolyte leakage was observed in the wild-type plants between the two treatments, while in the two knockouts higher electrolyte leakage was seen in the watered plants than in the droughted plants, though this difference was not significant.

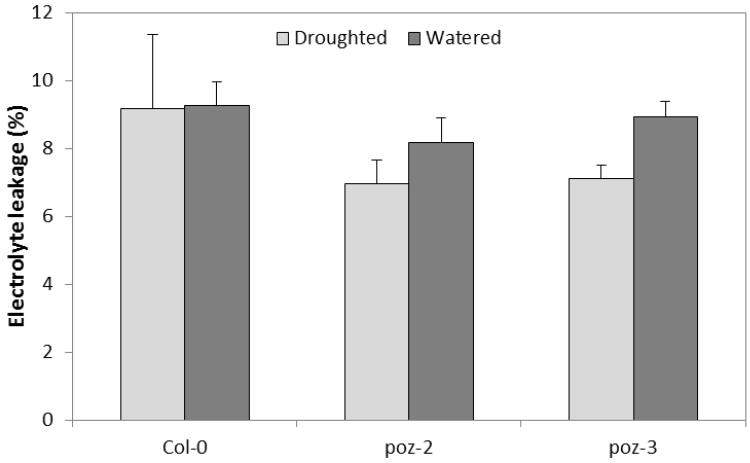


Figure 4.74. Measurement of electrolyte leakage from the *poz* knockouts and wild-type Col-0 under drought and watered conditions. All the values are mean + standard error (n=5). No significant difference was observed (P > 0.05, One-way ANOVA). The experiment was performed once.

4.8.3.3. Chlorophyll and carotenoid content

Chlorophyll (a, b and total) and carotenoid content were measured in the three lines under drought and watered conditions (Figure 4.75). Higher levels of chlorophyll and carotenoid content were measured in the droughted plants than in the watered plants for all three lines; this difference was significant in Col-0 for carotenoid content and total chlorophyll levels and in *poz-3* for all four measurements. Under droughted conditions, *poz-2* appeared to contain lower levels of chlorophyll and carotenoid than the wild-type and *poz-3*, though this was not significant.

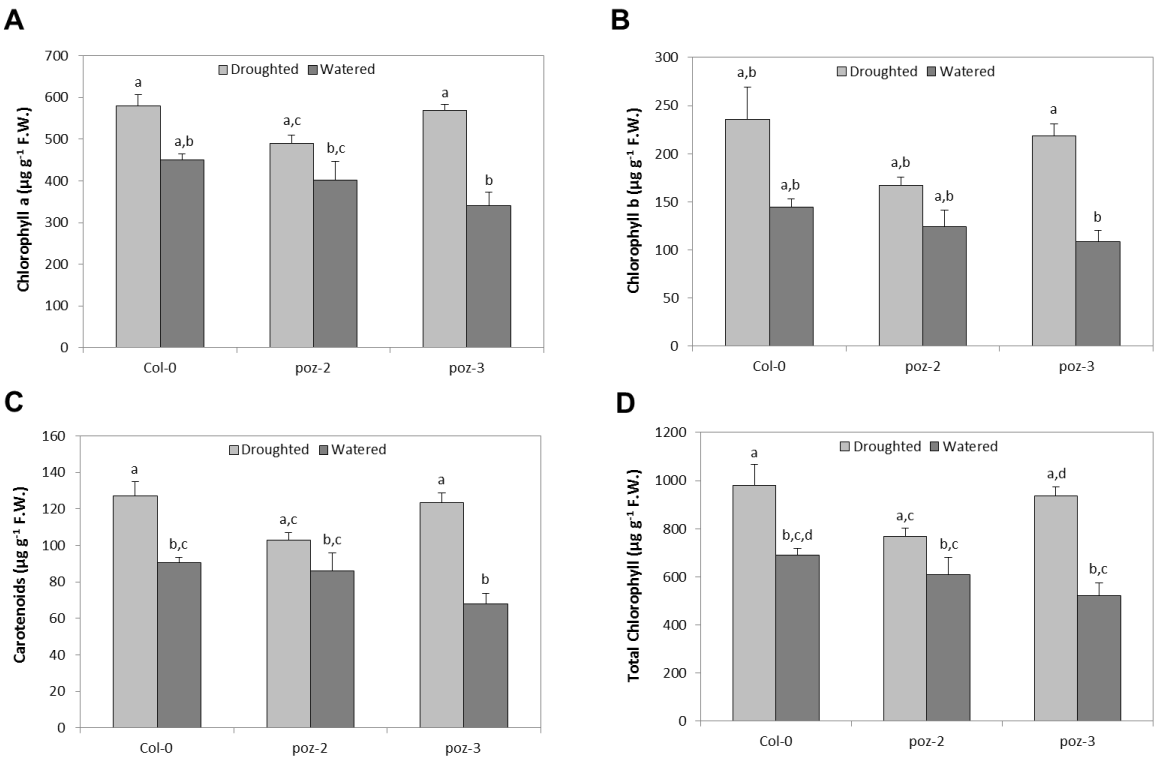


Figure 4.75. Estimation of chlorophyll content of the *poz* knockouts and the Col-0 wild-type. (A) Chlorophyll a (B) Chlorophyll b, (C) Carotenoids and (D) Total Chlorophyll content were measured under both drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.8.3.4. Anthocyanin content

Anthocyanin content was increased in the droughted plants in the wild-type and the two knockouts of *POZ*, though this was significant only for *poz-3* (Figure 4.76). Both mutants had higher anthocyanin content during drought compared to the wild-type, though this was not significant.

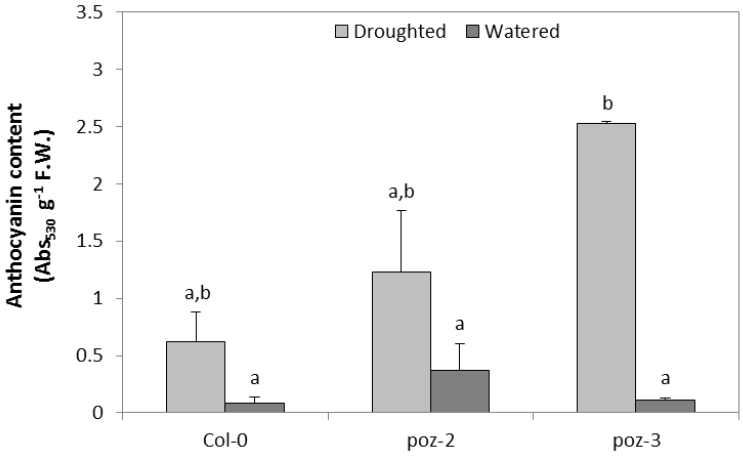


Figure 4.76. Estimation of anthocyanin content of the *poz* knockouts and the Col-0 wild-type under both drought and watered conditions. All the values are mean + standard error (n=3). Means with different letters indicate significant difference ($P < 0.05$, Welch ANOVA, post-hoc Games-Howell test). The experiment was performed once.

4.8.4. Measurement of growth and development

4.8.4.1. Flowering time

The flowering times of Col-0 and the knockouts of *POZ* were measured in terms of the number of leaves at flowering (Figure 4.77A). No difference in flowering time was seen between the droughted and watered Col-0 plants. Droughted *poz-2* plants appeared to flower later than the watered plants, while droughted *poz-3* plants flowered earlier than the respective watered plants, though both were not significant. Under both conditions of drought and watered, the wild-type flowered later than the two knockouts, though this difference was also not significant.

The flowering times of the wild-type and the knockouts of *POZ* were also measured in terms of the number of days to flowering from sowing (Figure 4.77B). In this case, droughted Col-0 plants appeared to flower later than watered Col-0 plants, though this was not significant. No difference in flowering time was seen in the knockout lines between the drought and watered plants. Under droughted conditions, it appeared that *poz-2* flowered later than the wild-type, while *poz-3* plants flowered earlier than Col-0. However, these differences were not significant. Under watered conditions, *poz-2* flowered significantly later than the wild-type, while no difference in flowering time was seen between the wild-type and *poz-3* plants.

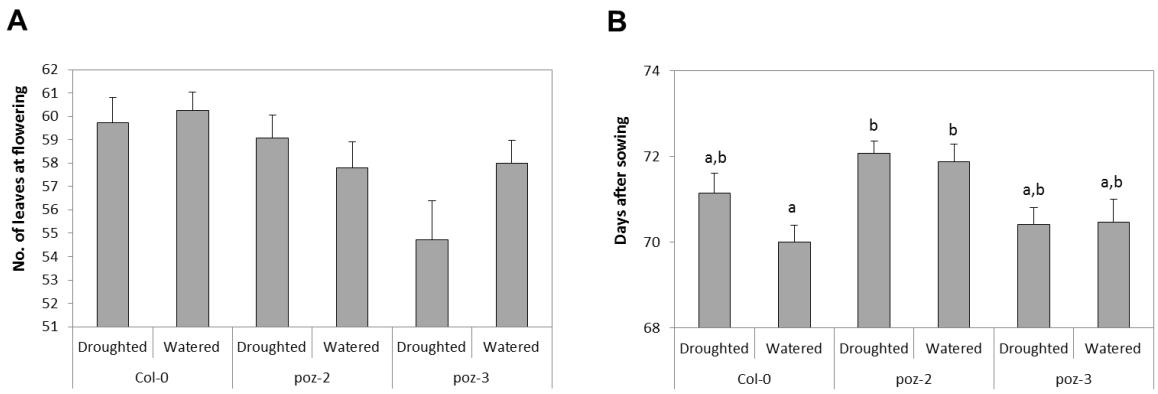


Figure 4.77. Flowering time measured in the Col-0 wild-type and the *poz* knockouts under drought and watered conditions. Flowering time was estimated in terms of (A) the number of leaves at flowering and (B) days to flowering after sowing. All the values are mean + standard error (n=10). No significant difference was observed in Figure A ($P > 0.05$, One-way ANOVA). In Figure B, means with different letters indicate significant difference ($P < 0.05$, One-way ANOVA, post-hoc Gabriel test). The experiment was performed once.

4.8.4.2. Biomass and harvest index

The vegetative and reproductive yield of the two knockouts of *POZ* were measured and compared to that of Col-0 plants (Figure 4.78). There was no difference in rosette biomass between the droughted and watered plants of the wild-type and the two mutant lines (Figure 4.78A). Overall, under both conditions of drought and watered, Col-0 appeared to produce more rosette mass than *poz-2* and *poz-3*, though this was not significant.

The mass of stalks and siliques (reproductive structures) were measured in the three lines in droughted and watered plants (Figure 4.78B). There appeared to be very little difference in the amount of stalks and siliques produced in droughted and watered plants of all three lines.

There appeared to be no difference in seed yield between the droughted and watered plants of Col-0 and *poz-2*, whereas droughted *poz-3* plants produced more seed than the watered plants, though this was not significant (Figure 4.78C). Under watered conditions, there appeared to be no difference in yield between the three lines. However, under droughted conditions *poz-2* produced less seed than *poz-3* plants, though this was not significant.

The total biomass produced by Col-0, *poz-2* and *poz-3* plants under drought and watered conditions shows that there was no difference in biomass between the lines or between drought-stressed and well-watered plants (Figure 4.78D).

The harvest index of the Col-0, *poz-2* and *poz-3* plants subjected to drought or under well-watered conditions was also calculated (Figure 4.79). There was no difference in harvest index between droughted and watered wild-type and *poz-2* plants. However, the seed harvest index of droughted *poz-3* was more than that of the watered plants.

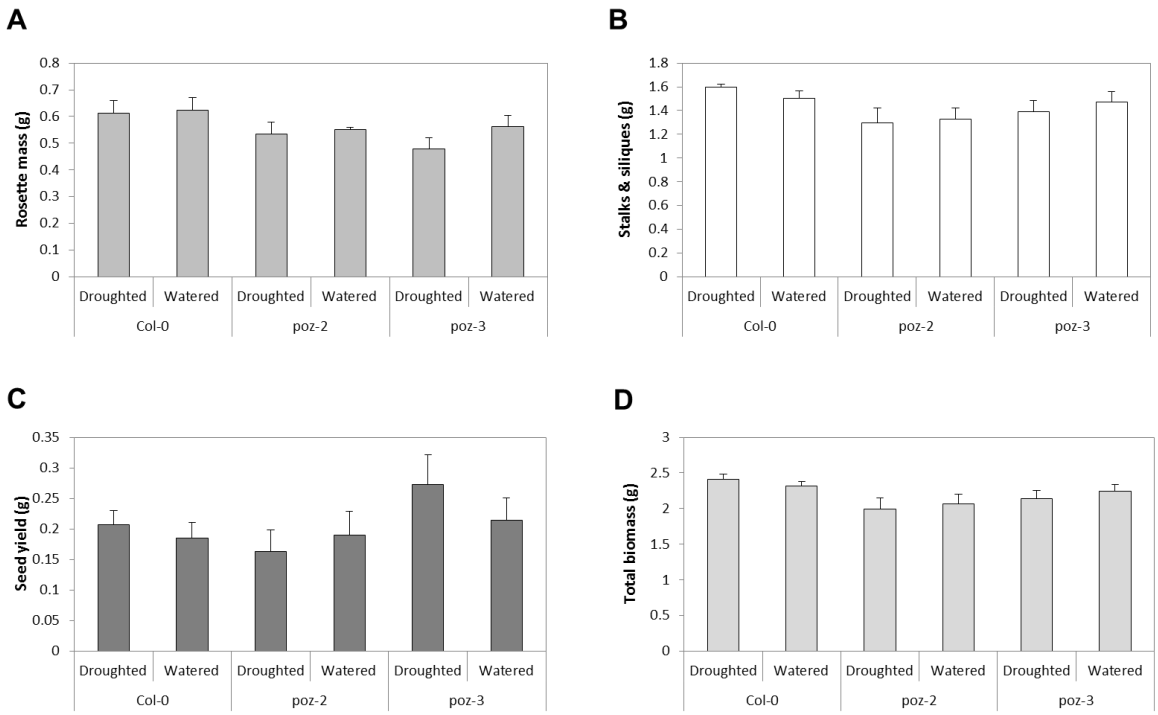


Figure 4.78. Measurement of the vegetative and reproductive biomass for the wild-type and the *poz* knockouts under drought and watered conditions. (A) Dry rosette weight, (B) the stalks and siliques, (C) seed yield and (D) total above-ground biomass were measured. All the values are mean + standard error (n=7). No significant difference was observed ($P > 0.05$, One-way ANOVA). The experiment was performed once.

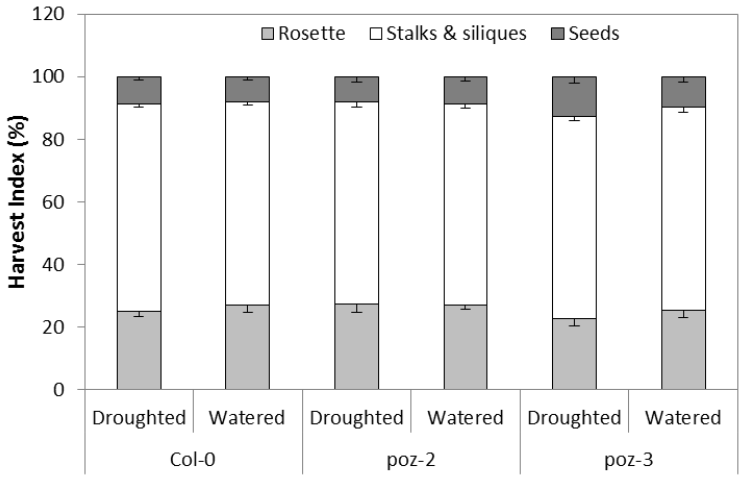


Figure 4.79. Harvest index estimated in the wild-type and the *poz* knockouts under drought and watered conditions. All the values are mean - standard error (n=7).

4.9. Discussion

In this chapter, mutant lines of *AGL22*, *FD*, *RAP2.12*, *BHLH038*, *ANL2*, *UKTF* and *POZ* were phenotyped under drought conditions. *AGL22* was identified as a hub gene in Bechtold *et al.* (2016), while the other genes were identified in chapter 3 (Section 3.3, including Figures 3.3 to 3.8). Only mutants for *AGL22* and *FD* showed a drought phenotype, while the other five genes did not.

However, *RAP2.12* is functionally redundant with *RAP2.2* (Nakano *et al.*, 2006) which may explain why the single knockout does not have a drought phenotype. *BHLH038* has high sequence similarity and possible functional redundancy with *BHLH039* (Wang *et al.*, 2007), while as shown in chapter 3, *UKTF* and *POZ* (Gingerich *et al.*, 2005) both belong to gene families. Thus, it is possible that modelling drought-responsive gene networks using VBSSM has predicted these gene families that are important in drought, and this would require multi-gene knockouts to verify this. Also, in the case of *UKTF* in particular, T-DNA insertional knockouts are not available and so it will not be entirely possible to say that *UKTF* is not involved in drought stress response.

4.9.1. *AGL22* could play a regulatory role during drought

When knockouts of *AGL22* were subjected to a progressive drought experiment, these mutants showed a faster drying rate than wild-type plants, and one of the reasons for this was seen to be due to a larger rosette area compared to the wild-types. It was also seen that these knockouts had a higher average temperature than Col-0, indicating a lower stomatal conductance, particularly towards the end of the drought. Since this was not observed throughout the course of the drought, it can be assumed that the lowered stomatal conductance was not a phenotypic effect of knocking out *AGL22*, but rather due to the decreasing water content in the tissue. The drying rates of the two overexpressing lines also appear to correlate with their measured rosette areas, and were higher than the Col-0 in AGL22ox-2, but there was no difference between AGL22ox-10 and Col-0. The observed differences between the two overexpressors of *AGL22* could be due to the difference in the levels of expression of *AGL22* in these two lines (Figure 3.24).

The reactive oxygen species, hydrogen peroxide (H_2O_2), is produced in plants in response to abiotic and biotic stresses (Apostol *et al.*, 1989; Moran *et al.*, 1994; Prasad *et al.*, 1994; Sgherri and Navari-Izzo, 1995; Tsugane *et al.*, 1999; Apel and Hirt, 2004).

Levels of H_2O_2 were measured in Col-0, the segregating wild-type, AGL22 WT, the two knockouts and the two *AGL22* overexpressing lines. In the wild-type lines, a clear difference can be seen in the amount of H_2O_2 measured between the two treatments; however, there appears to be no such difference in the levels of H_2O_2 measured in the knockouts. The levels of H_2O_2 in the knockouts under watered conditions are similar to those of the wild-type lines, but the amount measured under droughted conditions is much less in the knockouts than in the wild-type plants. This could be due to decreased production of H_2O_2, but possibly due to increased scavenging of ROS. Increased protection from oxidative stress could be a consequence of the early flowering phenotype of these mutants – the plants had begun flowering before the completion of the drought period and so would want to ensure production and survival of the next generation. The overexpressors of *AGL22*, on the other hand, had less H_2O_2 under watered conditions than the wild-types and knockouts, but produced more H_2O_2 under drought conditions than any of the other lines. There was a much larger difference in the levels of H_2O_2 measured in the overexpressors between the two treatments. This indicates that *AGL22* might play a role in the ability of plants to scavenge H_2O_2 produced under periods of drought. However, the expression levels of *AGL22* increase early in the drought period. This could mean that *AGL22* could play a role in H_2O_2-mediated stress signalling during drought. It would be interesting to study the expression levels of signalling genes, other ROS and calcium levels in the knockouts and overexpressors under drought to verify this.

The flowering time of the knockouts was significantly earlier than that of the wild-type plants. The number of leaves at flowering was less in the droughted plants than in the watered plants, but there was no difference in the number of days to flowering due to the two treatments. This indicates that the droughted knockouts flowered at an earlier developmental stage than the watered plants, but also took longer to grow and develop than the watered plants. Thus, it appears that in *AGL22* knockouts drought stress has a negative impact on the vegetative development of the plants, but does not hinder the timing of the reproductive development in these mutants. The mildly overexpressing AGL22ox-2 appeared to flower earlier than the wild-type lines, while the higher overexpressor, AGL22ox-10, flowered later.

There appears to be some discrepancy between the two overexpressing lines for which differences in expression levels of *AGL22* cannot account. In addition to the larger rosette area, AGL22ox-2 plants flowered significantly earlier than both AGL22ox-10 and Col-0. Since *AGL22* is a negative regulator of flowering (Hartmann *et al.*, 2000), increased levels of the gene should delay flowering relative to the wild-type, but the opposite is seen

with AGL22ox-2. It could be that though transcript levels are increased, there is reduction below wild-type levels of the amount of protein produced, due to post-transcriptional or translational changes. However, this does not explain why these plants produced sepaloid petals similar to the second overexpressor, AGL22ox-10. Thus, it is important to check the level of AGL22 protein in AGL22ox-2, and also to phenotype the other generated *AGL22* overexpressing lines shown in Figure 3.24. This will help to verify if mildly overexpressing lines of *AGL22* do produce a similar phenotype to knockouts, or if this is an artifact of the transformation process.

The vegetative and reproductive yield of the wild-type, knockout and overexpressing lines were determined under watered and droughted conditions. The yield of the knockouts was far less than any of the other lines, and in particular the seed yield was quite poor even under normal watered conditions, even though the biomass of stalks and siliques was comparable to the wild-type levels. On the other hand, the vegetative and reproductive yield of the overexpressors was similar to that of the wild-type, except the seed yield of AGL22ox-2. The low seed yield of the knockouts, *agl22-3* and *agl22-4*, could be explained by the fact that the plants had started flowering almost a week before the end of the drought experiment, which was conducted under short days. The suppressive effect of short day photoperiod on flowering might explain the poor yield of these lines, rather than the effect of the mutation. Unstressed knockouts of *AGL22* that were moved to long days soon after bolting were observed to produce almost twice the amount of seed of Col-0 plants (data not shown).

Thus, it would appear that *AGL22* has a possible role in the plant's response to drought, particularly with respect to regulating H_2O_2 levels, and could possibly produce a similar response to other abiotic stresses.

4.9.2. *FD* is drought-responsive

Similar to *AGL22*, the knockout of *FD* had a higher drying rate due to a larger rosette area, despite both genes having opposite functions in the regulation of flowering time (Hartmann *et al.*, 2000; Abe *et al.*, 2005). Thus, it would appear that *FD* has a similar effect on vegetative development as *AGL22*, but a difference in reproductive development. However, it is very difficult to make any conclusion on the role of *FD* in drought as only one knockout and overexpressor each were available. It is important to

have at least two, ideally three, mutant lines to study the effect of a gene, and thus the function of *FD* during drought remains to be seen.

4.9.3. Nature of drought stress

In this work, drought stress was initiated by withdrawing water until 20% rSWC was reached, which typically took two weeks. However, it was evident from the analysis in this chapter that the drought stress imposed was not severe – droughted plants showed no difference to the watered plants, or in some cases performed better than the watered plants. In particular, there appeared to be no reduction in photosynthetic assimilation rate as the drought progressed, and the droughted plants produced better yield than the watered plants. Similarly, results from electrolyte leakage and chlorophyll and carotenoid content indicate that the droughted plants were not stressed, and in fact were better off than the watered plants. On the other hand, results from the H_2O_2 and anthocyanin experiments indicate that the plants were experiencing drought stress. Thus, it appears that the drought stress imposed in this thesis is not sudden and severe, and hence it might be that the plants are able to adapt to the drought.

One reason for this could be the size of the pots used for the drought. This prolonged the drought for a total period two weeks for most mutants, while for the *AGL22* experiments took up to four weeks, and thus could have induced drought acclimation in the plants. This appeared to be the case based on the electrolyte leakage data in which the droughted plants had lower electrolyte leakage than the watered plants, and this was similar to that observed by Bouchabke *et al.* (2008). In their work, watered plants had higher electrolyte leakage than the droughted plants, and this was attributed to increased membrane hardening and thus drought adaptation. Thus it could be that to simulate drought stress in laboratory conditions, the size of the pot used relative to the size of the plant / root structure may have to be taken into consideration, to accurately portray field-like drought stress.

4.10. Conclusion

- Mutant lines of *AGL22*, *FD*, *RAP2.12*, *BHLH038*, *ANL2*, *UKTF* and *POZ* were phenotyped under drought conditions. Water use, photosynthetic performance, stress status, and growth and yield parameters were measured.

- Mutants of *AGL22* and *FD* had a drought phenotype compared to the wild-type.
- *AGL22* appears to regulate the levels of H_2O_2 in the plant during drought stress.
- *RAP2.12*, *BHLH038*, *ANL2*, *UKTF* and *POZ* showed no drought phenotype, but there may be functional redundancy for *RAP2.12*, *BHLH038*, *UKTF* and *POZ*.
- The drought stress applied was not severe and so the plants are likely to have experienced drought adaption, as indicated by the electrolyte leakage data.

Chapter 5

Testing network connections of gene regulatory networks

5.1. Introduction

In chapter 3, differentially expressed drought-responsive genes were used to model gene regulatory networks (GRNs) using VBSSM (Beal *et al.*, 2005). A number of GRNs were modelled which led to the identification of hub genes in these networks: *POZ*, *FD*, *UKTF*, *RAP2.12*, *BHLH038* and *ANL2*. Loss-of-function and gain-of-function mutants of the hub genes of these networks were screened and identified. *AGL22* was identified as a hub gene by Bechtold *et al.* (2016) and mutants of this gene were also identified. In chapter 4, these mutants were phenotyped under drought to ascertain if these genes are involved in drought stress response. Mutants of the flowering time gene *AGL22* showed a drought phenotype compared to wild-type plants. The knockout of *FD* also showed a higher drying rate than Col-0, however only one line was available for this. Mutants of the other genes did not show a different drought response compared to the wild-type.

This chapter describes the testing and analysis of some of the networks modelled in chapter 3. Only the network connections for *AGL22*, *RAP2.12*, *BHLH038* and *UKTF* were tested. Mutants of selected hub genes were subjected to a progressive drought stress and network connections from the gene regulatory networks in chapter 3 were analysed by qPCR. This was done to verify the connections predicted by the networks and to evaluate the accuracy of the modelling.

As shown in chapter 4, knockouts of *AGL22* showed a quicker drying rate phenotype compared to the wild-type when subjected to progressive drought. The drying rates of the overexpressors were ambiguous. Both knockouts and overexpressors produced different levels of H_2O_2 under drought, compared to the wild-type lines. *AGL22* has been shown to respond to changes in environmental conditions and has been implicated in thermal regulation of flowering (Lee *et al.*, 2007). Riboni *et al.* (2013) showed that *AGL22* is responsive to drought, and also that during drought under short days, inhibition of flowering was due to the action of *AGL22* and another floral repressor, *Flowering Locus C* (*FLC*).

Though there appeared to be no difference in drought phenotype due to knocking out either *RAP2.12* or *BHLH038*, some of the genes connected to these hub genes were also tested. The single knockout of *RAP2.12* did not show a difference to the wild-type in its drought response, however the double knockout of *rap2.12 rap2.3* was drought shown to be drought sensitive (Papdi *et al.*, 2015), indicating that *RAP2.12* may be functionally redundant with *RAP2.3*.

BHLH038 is involved in maintaining iron homeostasis and uptake, along with *BHLH039*, with which it is very similar in sequence. Thus, there may be functional redundancy between these two genes which may account for the lack of phenotype in the *bhlh038* knockout. Maintaining iron homeostasis is important because excess iron is toxic to plants, resulting in increased levels of ROS through the Fenton reaction and causing oxidative stress (Kampfenkel *et al.*, 1995; Gellego *et al.*, 1996). As drought stress also induces oxidative stress (Moran *et al.*, 1994; Munné-Bosch and Peñuelas, 2004; Sharma and Dubey, 2005), the plant's ability to maintain iron homeostasis is important to control oxidative stress during drought.

Some of the network connections for the gene network identifying the unknown gene *UKTF* were also tested. No knockouts were available for this gene – only T-DNA insertion-induced overexpressors. Though these mutants also did not show a drought phenotype, some connections in the gene network were investigated as this gene may be involved in regulating flower development in Arabidopsis and could be an important regulatory gene, since flowering is affected during drought.

5.2. Analysis of gene regulatory network potentially controlled by *AGL22*

As mentioned in Chapter 3, gene regulatory networks (GRN) were modelled using the algorithm VBSSM (Beal *et al.*, 2005) and transcriptomics data obtained from a progressive drought of Col-0 plants. Bechtold *et al.* (2016) describes one such GRN that was modelled and led to the identification of *AGL22* as a 'hub' gene of that GRN, and a potentially important gene in the drought response of plants. Figure 5.1 shows some of the connections that were modelled, as shown in Chapter 3. The green nodes in Figure 5.1 show some of the genes that were modelled by Bechtold *et al.* and which were tested in this chapter. The yellow nodes in Figure 5.1 are some of the connections that were predicted by modelling flowering-related genes with *AGL22*.

The model predicts that under drought conditions, *AGL22* induces the expression of *AUXIN RESPONSE FACTOR 2* (*ARF2*), *VERNALIZATION 1* (*VRN1*), *WRKY20*, *DICER-LIKE 4* (*DCL4*), *RS-CONTAINING ZINC FINGER PROTEIN 21* (*RSZ21*), *PLEIOTROPIC REGULATORY LOCUS 1* (*PRL1*), a nuclease-phosphatase, and *FLOWERING BHLH 3* (*FBH3*). It also predicts that *AGL22* negatively regulates the expression of *DREB1A*, *ERF023*, *ERF034* and a peroxidase, under drought conditions.

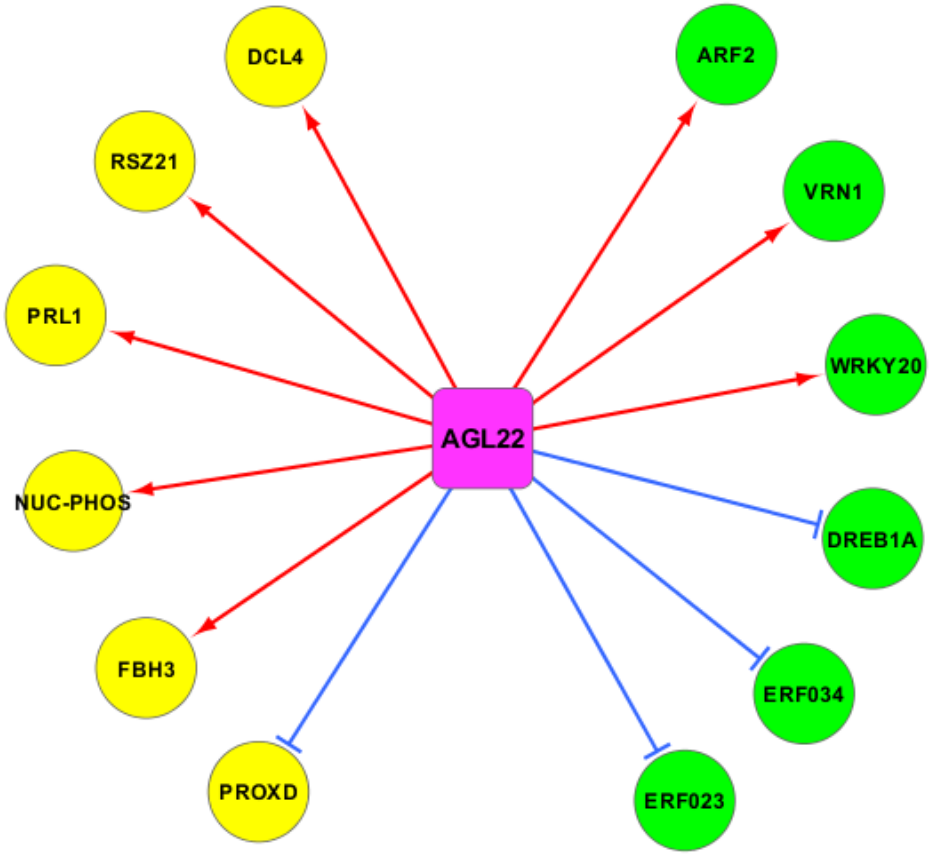


Figure 5.1. Gene regulatory network potentially regulated by *AGL22*, which was created by VBSSM. The nodes in green are some of the connections in the network from Bechtold *et al.* (2016). The yellow nodes are connections from modelling *AGL22* with flowering-related genes. The red connections indicate that the connections modelled are inductive, and the blue connections indicate inhibition.

The transcriptional regulation of the above genes by *AGL22* was verified by qPCR using the knockouts, *agl22-3* and *agl22-4*, and the wild-type, Col-0, which were subjected to drought conditions. Figure 5.2 shows the results of the qPCR analysis in the two knockouts of *AGL22*. The data is expressed as $\log_2$ (fold expression of gene in mutant line relative to that in the wild-type).

The genes *WRKY20*, *VRN1*, *ARF2* and *DCL4* which were modelled to be induced by *AGL22* under drought conditions (Figure 5.1), were found to be down-regulated in both knockout lines of *AGL22* under drought conditions, thus verifying these predicted connections. The other genes that were predicted to be induced by *AGL22* during drought were *RSZ21*, *PRL1*, the nuclease-phosphatase and *FBH3*. *RSZ21* appeared to be mildly down-regulated in *agl22-4*, but there was no change in the level of its

expression in *agl22-3*. There appeared to be no change in the expression levels of *PRL1* and the nuclease-phosphatase in both knockout lines of *AGL22* under drought conditions, while *FBH3* showed the opposite expression pattern to that predicted by the network. It was found to be induced in *agl22-3* and *agl22-4* under drought, indicating that *AGL22* negatively regulates *FBH3* under drought conditions.

Of the genes that were predicted to be negatively regulated by *AGL22* under drought conditions, only *ERF023* and *ERF034* were found to agree with this prediction, and were up-regulated in the two knockouts under drought conditions. *DREB1A* and the peroxidase were found to be down-regulated in both *agl22-3* and *agl22-4*. This implies that under drought conditions, these genes appear to be induced by *AGL22*.

These genes were also tested in *agl22-3* and *agl22-4* plants that were maintained under watered conditions (Figure 5.3) to verify if the observed changes in expression levels of the above genes were due to the effect of the drought or solely due to the mutation in the *AGL22* gene. The levels of *ERF023* in normally watered knockouts of *AGL22* were found to be similar to those in the droughted mutant lines. Similar levels of expression of *DREB1A*, *VRN1* and *RSZ21* were also seen in normally watered *agl22-3* and *agl22-4*, as in the droughted plants. These indicate that the reduced expression of *AGL22* contributed to the differential expression of these genes, rather than the stress treatment. Similar to the droughted plants, no difference in the expression of *PRL1* and the nuclease-phosphatase were seen in the watered knockouts of *AGL22*. Also, no differential expression was seen in the genes *WRKY20*, *ARF2*, *FBH3* and the peroxidase in the watered plants, while all of these were down-regulated under drought conditions. *DCL4* was found to be only mildly down-regulated in the watered knockout lines of *AGL22*, compared to the droughted mutants. On the other hand, under watered conditions, *ERF034* was found to be further induced than under drought conditions, indicating that *AGL22* down-regulates *ERF034* under normal conditions, and that this control was mildly relaxed under drought conditions.

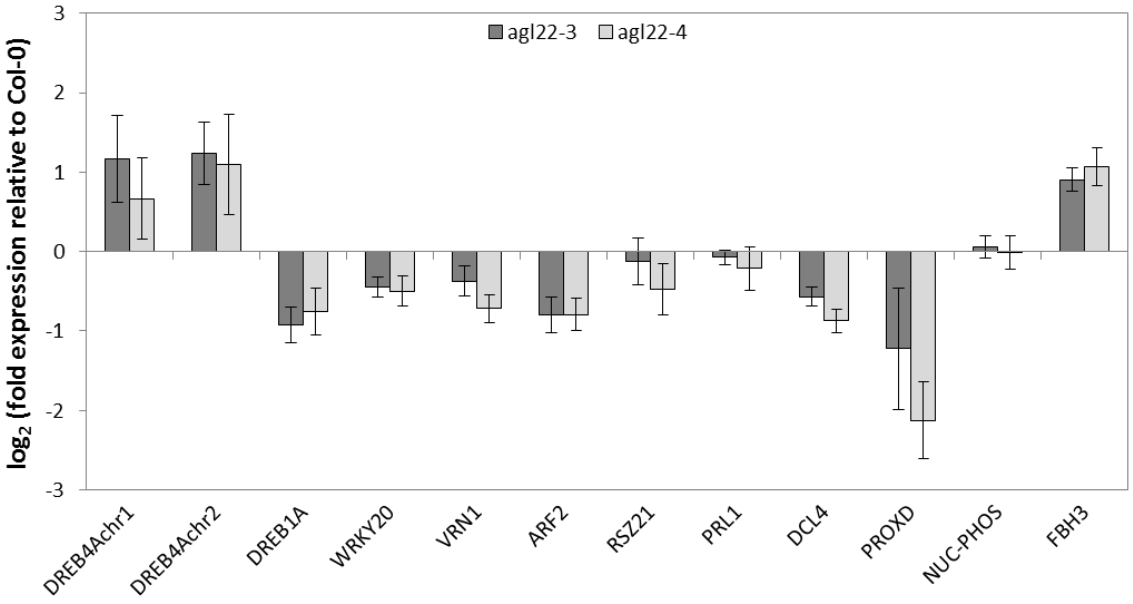


Figure 5.2. qPCR analysis of the genes identified in Figure 5.1 in *agl22-3* and *agl22-4* subjected to drought. Values are mean ± standard error (no. of biological replicates, n=5).

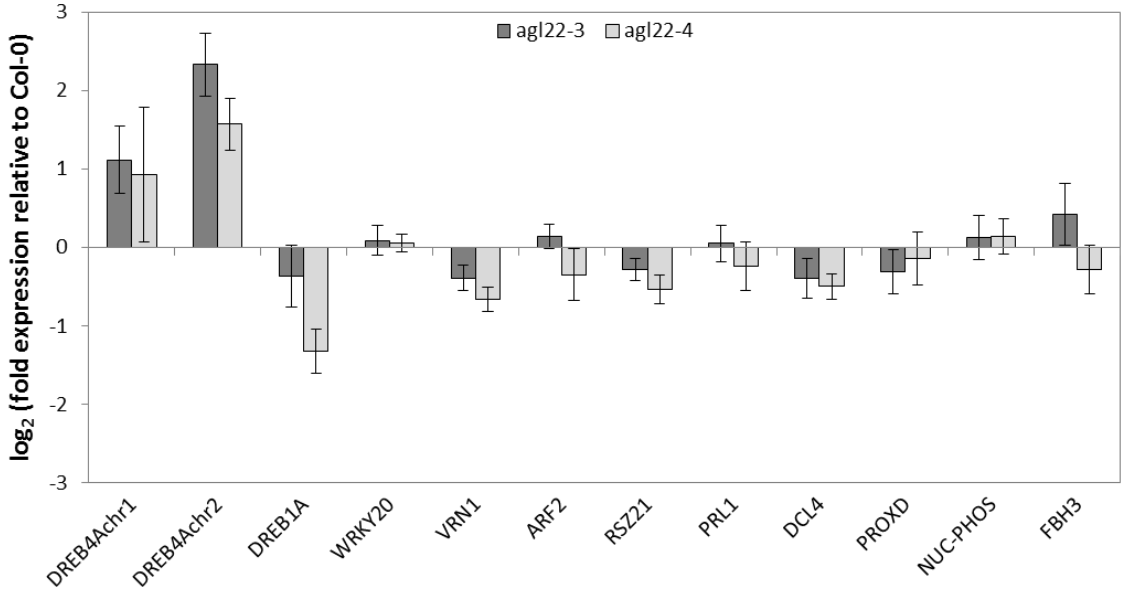


Figure 5.3. qPCR analysis of the genes identified in Figure 5.1 in *agl22-3* and *agl22-4* maintained under watered conditions. Values are mean ± standard error (no. of biological replicates, n=5).

5.3. Analysis of gene regulatory network potentially controlled by *RAP2.12*

Another GRN was modelled using VBSSM in which RAP2.12 was the hub gene. Due to lack of time, only a few of the genes were tested in the following models. Some of the genes that were predicted to be up-regulated by RAP2.12 are shown in Figure 5.4. These genes are *PHYTOCHROME INTERACTING FACTOR 3-LIKE 2* (*PIL2*), *BASIC PENTACYSTEINE 7* (*BPC7*), a bZIP-encoding gene *BZO2H1* and *DP-E2F-LIKE 2* (*DEL2*).

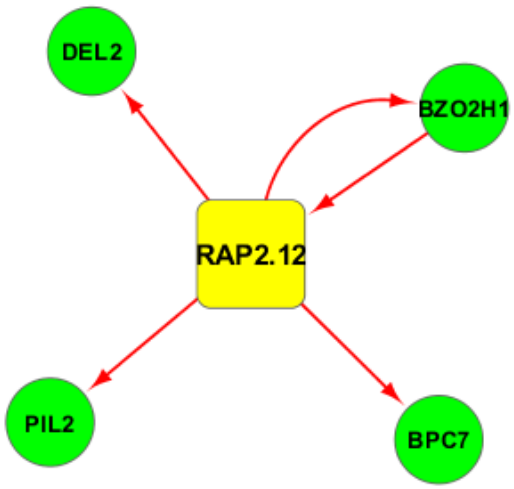


Figure 5.4. Gene regulatory network potentially regulated by *RAP2.12*, which was created by VBSSM. The red connections indicate that the connections modelled are inductive.

The wild-type, Col-0, and knockouts of *RAP2.12*, *rap2.12-1* and *rap2.12-3*, were subjected to drought conditions and leaf tissue was taken from each line to analyse the expression levels of the above genes using qPCR. The result of the expression analysis is shown in Figure 5.5. There appeared to be no difference in the expression levels of all the genes analysed between the mutant and the wild-type lines, under drought conditions. Under watered conditions, there also appeared to be no difference in the expression levels of *BZO2H1* and *PIL2* (Figure 5.6). *BPC7* appeared to be mildly down-regulated in both mutant lines, under watered conditions, while *DEL2* was down-regulated in only *rap2.12-1*.

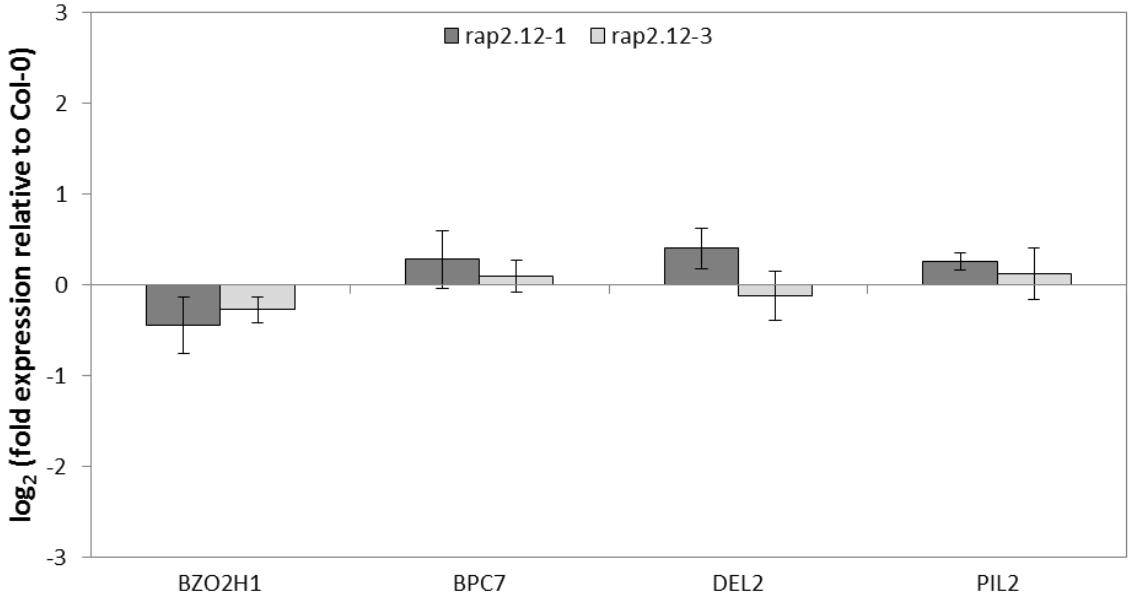


Figure 5.5. qPCR analysis of the genes identified in Figure 5.4 in *rap2.12-1* and *rap2.12-3* subjected to drought. Values are mean ± standard error (no. of biological replicates, n=5).

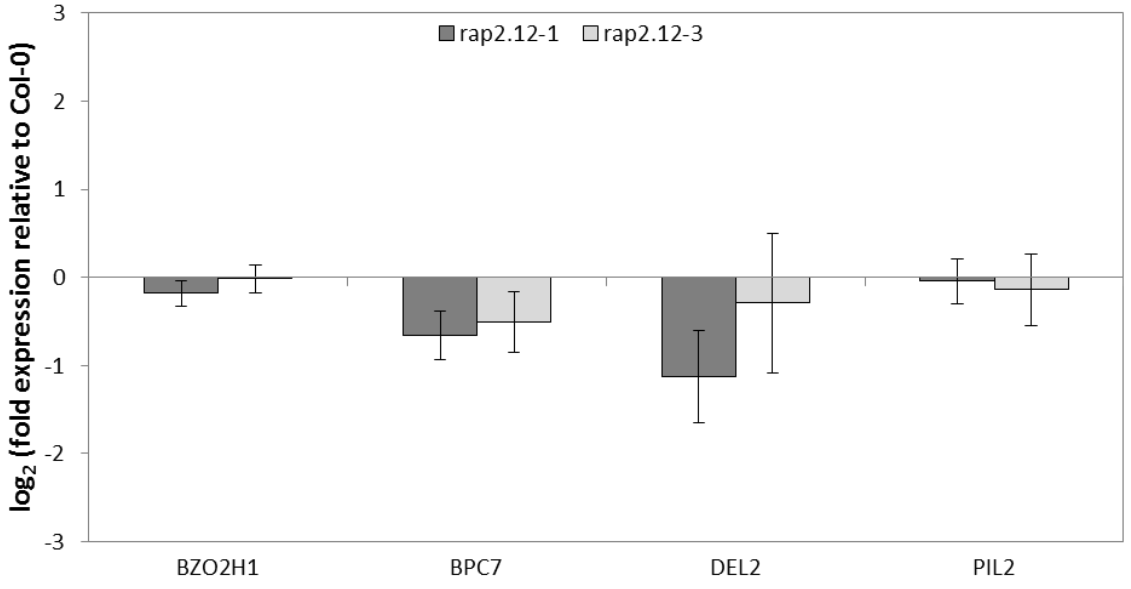


Figure 5.6. qPCR analysis of the genes identified in Figure 5.4 in *rap2.12-1* and *rap2.12-3* maintained under watered conditions. Values are mean ± standard error (no. of biological replicates, n=5).

5.4. Analysis of gene regulatory network potentially controlled by *BHLH038*

BHLH038 was modelled as a hub gene in a GRN in which it is predicted to inhibit the expression levels of *PIL2*, *BZO2H1*, *BPC7* and *DEL2* (Figure 5.7). However, as seen in Figures 5.8 and 5.9, no difference in the expression of the above genes was seen in the knockouts of *BHLH038* under both drought and watered conditions.

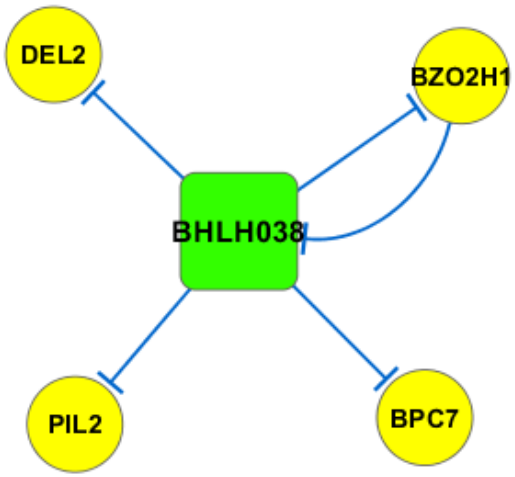


Figure 5.7. Gene regulatory network potentially regulated by *BHLH038*, which was created by VBSSM. The blue connections indicate that the connections modelled are inhibitory.

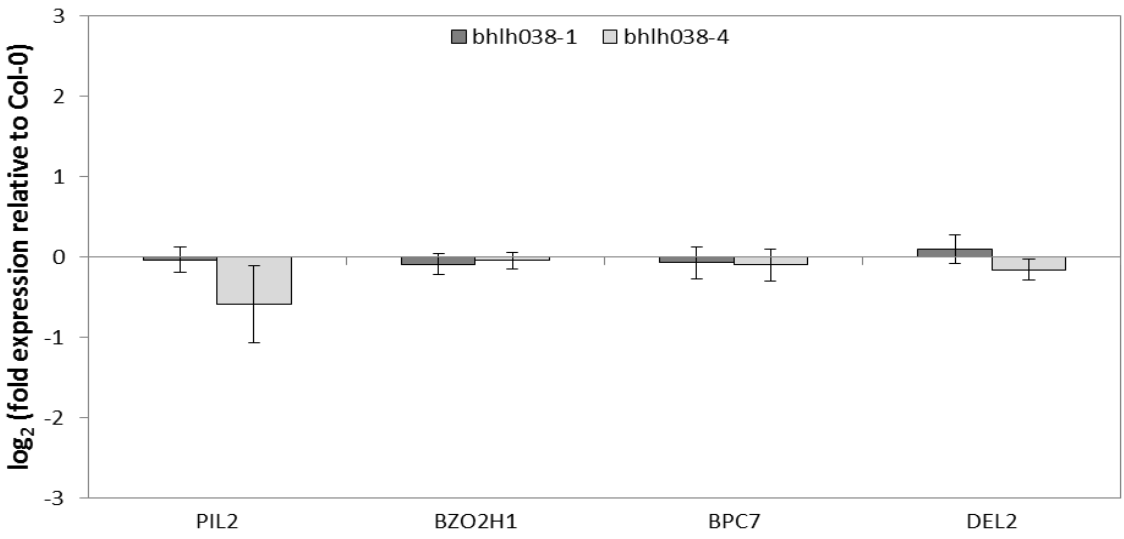


Figure 5.8. qPCR analysis of the genes identified in Figure 5.7 in *bhlh038-1* and *bhlh038-4* subjected to drought. Values are mean ± standard error (no. of biological replicates, n=5).

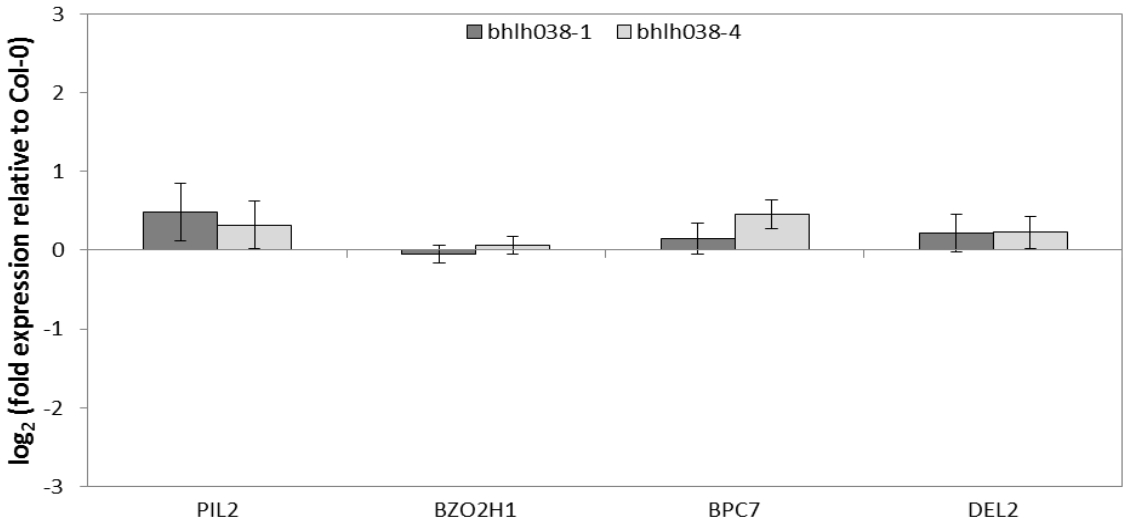


Figure 5.9. qPCR analysis of the genes identified in Figure 5.7 in *bhlh038-1* and *bhlh038-4* maintained under watered conditions. Values are mean ± standard error (no. of biological replicates, n=5).

5.5. Analysis of gene regulatory network potentially controlled by *UKTF*

As described in Chapter 3, *UKTF* was identified as a potential hub gene in a predicted GRN involved in a plant's drought response. Figure 5.10 shows some of the connections that were modelled using VBSSM, namely *NUCLEAR FACTOR Y, SUBUNIT A10* (*NF-YA10*) and *HOMOLOG OF XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP B 1* (*XPB1*) (green nodes). Having identified *UKTF* as having a MIP1 domain, *UKTF* was also modelled with flowering- and flowering time-related genes to identify potential *UKTF*-regulated genes that are involved in flowering. *GA INSENSITIVE DWARF1C* (*GID1C*), *WITH NO LYSINE (K) KINASE 5* (*WNK5*) and *SUGAR INSENSITIVE 8* (*SIS8*) are some of the genes identified in this network and are also shown in Figure 5.10 (purple nodes).

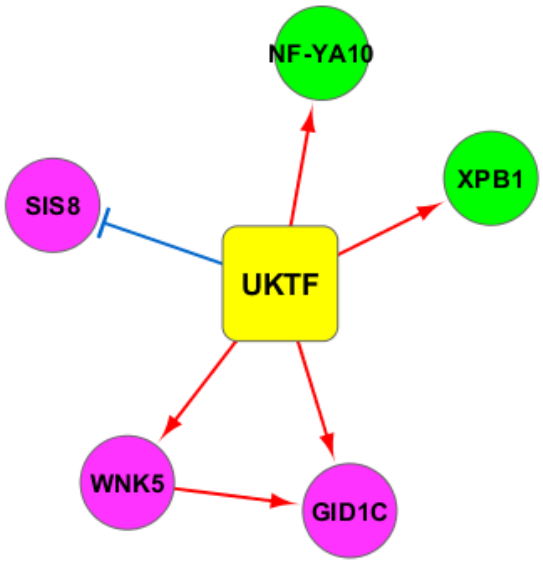


Figure 5.10. Gene regulatory network potentially regulated by *UKTF*, which was created by VBSSM. The nodes in green are some of the connections in the network from Chapter 3. The purple nodes are connections from modelling *UKTF* with flowering-related genes. The red connections indicate that the connections modelled are inductive, and the blue connection indicates inhibition.

As seen in Figures 5.11 and 5.12, there appeared to be no difference in the expression of *NF-YA10* during drought, but it was up-regulated in UKTF-3 under watered conditions. There appeared to be no difference in the expression of *NF-YA10* in UKTF-4 under both drought and watered conditions. Very little difference in the expression of *XPB1*, *GID1C* and *WNK5* was seen between the mutants and the wild-type plants under both drought and watered conditions. SIS8 was down-regulated in both UKTF-3 and UKTF-4 under drought conditions, as per the model. Under watered conditions, SIS8 was induced in UKTF-3, while there appeared to be no difference in its expression in UKTF-4.

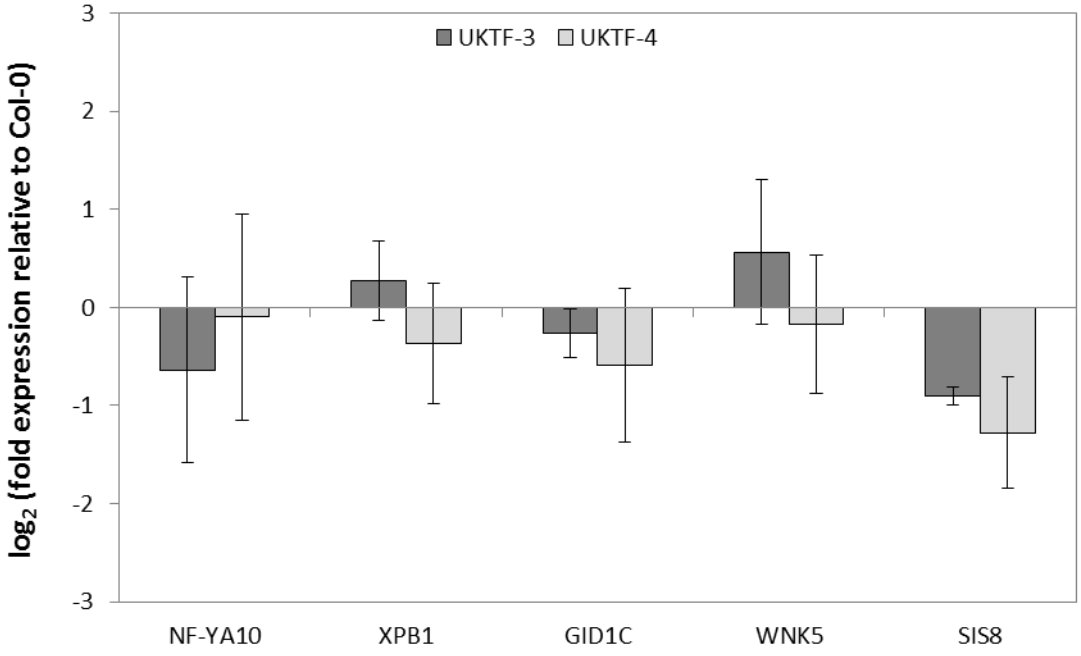


Figure 5.11. qPCR analysis of the genes identified in Figure 5.10 in UKTF-3 and UKTF-4 subjected to drought. Values are mean ± standard error (no. of biological replicates, n=5).

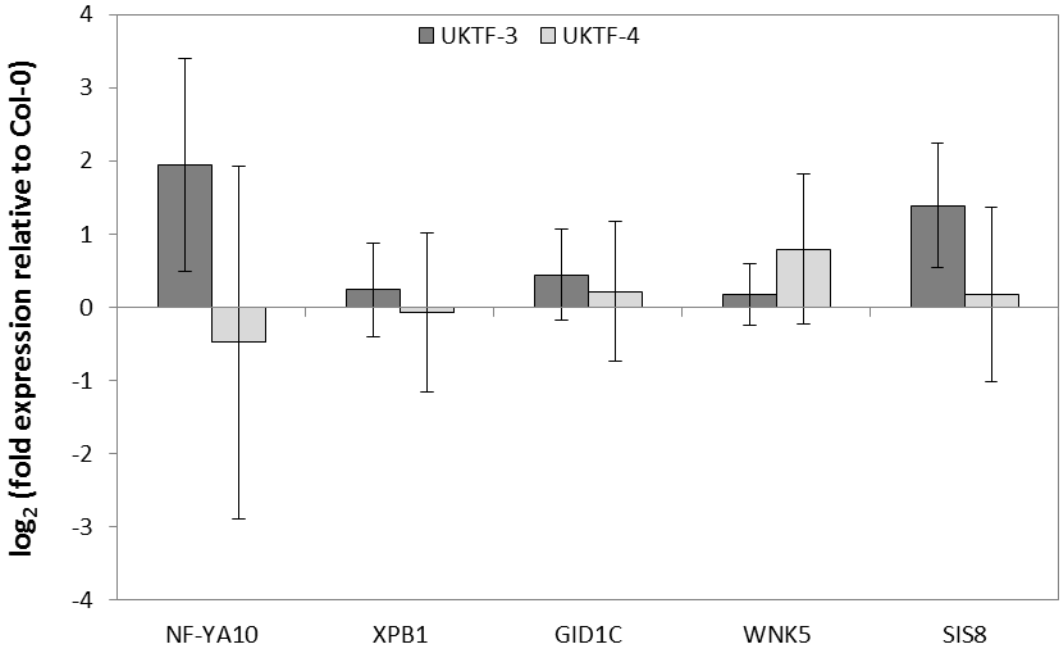


Figure 5.12. qPCR analysis of the genes identified in Figure 5.10 in UKTF-3 and UKTF-4 maintained under drought conditions. Values are mean ± standard error (no. of biological replicates, n=5).

5.6. Discussion

Gene regulatory networks for *AGL22*, *RAP2.12*, *BHLH038* and *UKTF* were tested in this chapter. This was done to ascertain the accuracy of the VBSSM modelling and also to test the regulation of the expression of genes that are not known to be regulated by the above four, and thus establish new targets of these genes and understand the role they may play during drought.

5.6.1. Gene regulatory network for *AGL22*

In section 5.2, it was confirmed that *AGL22* induces the expression of *WRKY20*, *VRN1*, *ARF2* and *DCL4*, as predicted by the model; however, only *WRKY20*, *ARF2* and *DCL4* were induced during drought, while *VRN1* appeared to be induced as a result of the reduced expression of *AGL22*. On the other hand, *ERF023* and *ERF034* were predicted by VBSSM to be down-regulated by *AGL22* during drought. These genes were found to be induced in knockouts of *AGL22*, implying that they are down-regulated by *AGL22*, but this was not specific to droughted conditions. There appeared to be no difference in the expression levels of *ERF023* under both drought and watered conditions, while *ERF034* appeared to be further down-regulated during drought. This implies that *ERF023* is regulated by *AGL22* regardless of the conditions, while under drought another gene(s) mediates the regulation of the expression of *ERF034*.

FBH3 and *PROXD* were predicted to be positively and negatively regulated by *AGL22* during drought, respectively, while in reality the opposite was observed – *FBH3* was found to be negatively regulated by *AGL22* under drought stress, while *PROXD* was positively regulated. *DREB1A* was also found to be induced by *AGL22*, even though the model predicted the opposite; however this gene was up-regulated under both droughted and watered conditions, and thus its expression is regulated by *AGL22* regardless of the conditions. It has been shown in literature that that *DREB1A* is only cold-responsive and not drought-inducible (Liu *et al.*, 1998), so it is surprising that this gene has come up in the drought transcriptomics data of Bechtold *et al.* (2016). However, the difference could be down to the different methods used to induce drought. Liu *et al.* allowed rosettes to dehydrate on the bench as a method of drought and lasted only hours, while Bechtold *et al.* used a soil-based slow drying progressive drought over a two-week period, which is likely to be a more realistic situation of drought stress in the field.

WRKY20 has been shown to be induced by jasmonic acid treatment (Schluttenhofer *et al.*, 2014) and a homologue of it in soybean, *GmWRKY1*, was found to be induced by pathogen infection (Kang *et al.*, 2009). These indicate that *WRKY20* is likely to be induced during biotic stresses. *WRKY20* was also shown to bind to the promoter of *Pathogenesis-Related4* (*PR4*) and induce its expression in Arabidopsis (Proietti *et al.*, 2011). PR4 contains a lectin domain (Potter *et al.*, 1993) and has been shown to be induced during pathogen infection (Mukherjee *et al.*, 2010). *PR4* was also found to be induced by drought, and constitutive overexpression of this gene conferred Arabidopsis plants with drought tolerance due to increased cell membrane stability (Cabello and Chan, 2012). Increased cell membrane stability is a measure of plant tolerance to abiotic stresses like drought, freezing and salinity, as it indicates increased osmotic adjustment and protection of the cell membrane, leading to decreased permeability of ions through the membrane (Ebercon and Blum, 1981). Thus, during drought, upregulation of *WRKY20* by *AGL22* could result in induction of *PR4*, and potentially better cell membrane stability and increased drought tolerance.

WRKY20 was also shown to positively regulate the expression of *ApL3*, which codes for one of the large subunits of AGPase (Nagata *et al.*, 2012). *ApL3* is the predominantly-expressed large subunit in inflorescences and roots, and is also expressed in leaves (Crevillén *et al.*, 2005). *WRKY20* was also found to be induced by mannitol and a subsequent increase in *ApL3* was observed (Nagata *et al.*, 2012). Prasch *et al.* (2015) observed increased levels of *ApL3* and other starch synthesis genes, and consequently starch accumulation, in guard cells during drought. A knockout mutant of *β-amylase1* (*BAM1*) confers these plants unable to breakdown starch in stomata, and these mutants are more drought-tolerant than Col-0 plants due to reduced stomatal opening (Prasch *et al.*, 2015). Thus, through *WRKY20*, *AGL22* may contribute to the drought stress response by increased starch accumulation in guard cells, and subsequently minimal stomatal opening.

ARF2 (*Auxin Response Factor2*) is involved in auxin signalling and plays a role in repressing cell division (Schruff *et al.*, 2006). It also affects cell elongation and is involved in various developmental processes – it induces early flowering, floral organ abscission and early senescence (Okushima *et al.*, 2005). *ARF2* also negatively regulates ABA control of seed germination and primary root growth (Wang *et al.*, 2011). Knockouts of *ARF2* in Arabidopsis experience poor yield due to the formation of long gynoecia, resulting in failure in self-fertilisation as pollen from the stamen cannot reach the stigma (Okushima *et al.*, 2005; Schruff *et al.*, 2006). This phenotype is predominant in the early-

developed flowers and becomes less pronounced in the later flowers, enabling self-fertilisation (Okushima *et al.*, 2005). Thus, *ARF2* not only induces early flowering, but also enables successful self-fertilisation in those early flowers.

In the work of Meng *et al.* (2015), *arf2* mutants were shown to have increased levels of abscisic acid and were also found to be more drought-tolerant, due to increased stomatal closure and reduced water loss. *ARF2* was also found to negatively regulate the transcription factor *AINTEGUMENTA* (*ANT*), a positive regulator of the drought-responsive gene *COR15A*, which would further contribute to the drought-tolerant phenotype of *arf2* mutants (Meng *et al.*, 2015). The fact that *AGL22* induces *ARF2* during drought indicates that the plant places a priority of successful reproduction over drought avoidance.

DCL4 (*Dicer-like4*) is involved in the production of transactivating-siRNA (ta-siRNA) from three families of ta-siRNA-encoding genes, *TAS1*, *TAS2* and *TAS3* (Xie *et al.*, 2005). The *TAS3* family consists of two members that target the degradation of *ARF2*, *ARF3* and *ARF4* (Williams *et al.*, 2005). Other genes involved in the generation of these tasiRNA-ARFs are *RNA-dependent RNA polymerase6* (*RDR6*), *Suppressor of Gene Silencing3* (*SGS3*) and *Argonaute7* (*AGO7*). Compared to Col-0, knockouts of these genes were shown to produce less seed after the plants were recovered from a period of drought (Matsui *et al.*, 2014). It was also shown that this phenotype was due to the lack of degradation of *ARF3*, which caused the formation of shortened stamens and subsequently reproductive failure.

Thus, during drought, *AGL22* appears to induce *DCL4* which promotes the degradation of *ARF3*, ensuring that reproduction in the early-formed flowers is successful. However, *ARF2* is also degraded by the tasiRNA-ARF pathway, and the lack of *ARF2* appears to negatively influence reproduction. It may be that *AGL22* also induces the expression of *ARF2* to ensure that when self-fertilisation occurs, either during drought or after recovery, there is sufficient production of seed allowing survival of the next generation. The fact that the plants sampled for the expression analysis in this chapter were 8 weeks old and very close to flowering (they were under an 8/16-hour light/dark period), might explain the reason for the induction of genes and processes related to the reproductive phase and successful reproduction. *AGL22* was found to be up-regulated in the microarray analysis of drought stressed plants (Bechtold *et al.*, 2016). Thus, it would appear that it may be involved in not only repressing flowering during the drought period, but also preparing the plant for a relatively successful reproductive phase, when conditions become favourable.

5.6.2. Gene regulatory network for *RAP2.12* and *BHLH038*

The GRNs for *RAP2.12* and *BHLH038* were also tested but none of the connections appeared to hold true to the modelling. This may be because only a small number of genes in the model were tested. It is important to test a larger number of genes to ascertain the accuracy of the model, because as seen in the GRN for *AGL22*, only 5 of the 12 genes that were tested were found to be regulated by *AGL22* under droughted conditions (section 5.2). So, four genes may not be adequate to test the GRN.

5.6.3. Gene regulatory network for *UKTF*

Genes for the GRN of *UKTF* were analysed and only *SIS8* appeared to be down-regulated by *UKTF* during drought. *SIS8* has been identified as a possible mitogen-activated kinase kinase kinase which regulates resistance to high sugar levels in seedlings. Knockouts of *SIS8* are less sensitive to the inhibitory effects of high levels of sugar in seedlings (Huang *et al.*, 2014). This may imply that these plants are osmotolerant as they can accumulate more sugars (Gibson *et al.*, 2001), and indicating that *UKTF* may be involved in sugar signalling during drought.

SIS8 is also a negative regulator of salt stress and is down-regulated by the stress, while knockouts of *SIS8* are tolerant to salinity stress (Gao and Xiang, 2008), although the mechanism for this is unknown. As shown in Figure 5.11, *UKTF* down-regulates the expression of *SIS8* during drought stress, which in turn may act in a similar way as documented during salinity stress, and confer the plant with drought tolerance.

5.7. Conclusion

- GRNs for *AGL22*, *RAP2.12*, *BHLH038* and *UKTF* were tested
- Five of the twelve genes in the GRN for *AGL22* were differentially expressed by this gene under drought conditions including *WRKY20*, *ARF2*, *DCL4*, *FBH3* and a peroxidase.
- *AGL22* may confer drought through *WRKY20* by increased cell membrane stability.
- *AGL22* may also ensure successful reproduction after drought by acting through *ARF2* and *DCL4*.

- The GRNs for *RAP2.12* and *BHLH038* were not found to hold true, but more genes in these networks need to be tested.
- *UKTF* was found to down-regulate *SIS8* and hence might be involved in sugar signalling during drought.

Chapter 6

General Discussion

6.1. VBSSM was used to model GRNs involved in drought

Drought has a devastating effect on plants and severely affects growth and development. Many studies have focussed on the effect of drought on plant transcriptomes, particularly in Arabidopsis (Kreps *et al.*, 2002; Seki *et al.*, 2002; Killian *et al.*, 2007; Harb *et al.*, 2010). Though these studies were useful in identifying transcription factors (TFs) and other genes as important to the drought response, they were subjected to unnatural water-deficit situations such as dehydration of rosettes and addition of mannitol (Kreps *et al.*, 2002; Seki *et al.*, 2002; Killian *et al.*, 2007). Also, single time-point studies have been carried out which do not show changes in the transcriptome as the drought progresses and the early response events involved (Seki *et al.*, 2001).

The work presented in this thesis is based on the work described in Bechtold *et al.* (2016). In that work, a soil-based progressive drought was applied to Col-0 plants until the relative soil water content (rSWC) reached 20% and a time-series transcriptomics analysis was performed. In Chapter 3 of this thesis, a list of 2330 differentially expressed genes were identified which were used to model gene networks using a dynamic Bayesian algorithm called Variational Bayesian State-Space Modelling (VBSSM; Beal *et al.*, 2005). In addition to analysing these genes, the gene *AGL22* which was identified in Bechtold *et al.* (2016) was also analysed in this work.

VBSSM was used to model Gene Regulatory Networks (GRNs) involved in the drought response. A number of such networks were predicted and particular notice was taken of the highly connected or 'hub' genes in the network, as these could act as potentially important genes during drought that directly or indirectly control the expression of the members of the networks. Six such 'hub' genes were identified and along with *AGL22*, these were: *FD* (*Flowering locus D*), a bZIP protein involved in the positive regulation of flowering (AT4G35900; Abe *et al.*, 2005); *RAP2.12* (*Related to Apetala2*), a member of the ERF subfamily of the ERF/AP2 transcription factor family, and has been found to be involved in response to hypoxia (AT1G53910; Licausi *et al.*, 2011); *BHLH038*, a basic helix-loop-helix protein that regulates iron homeostasis (AT3G56970; Wang *et al.*, 2007); *ANL2* (*Anthocyaninless 2*), a protein involved in the accumulation of anthocyanin and in root development (AT4G00730; Kubo *et al.*, 1999); a protein of unknown function, designated *UKTF* (AT1G16750; The Arabidopsis Information Resource); a protein of unknown function containing the BTB/POZ (Bric-a-Brac, Tramtrack, Broad-complex / Pox virus and Zinc finger) domain, designated in this work as *POZ* (AT1G55760; Gingerich *et al.*, 2005).

UKTF and POZ were characterised *in silico* to identify possible functions of the protein products (Chapter 3). POZ was identified as having a BTB-POZ domain and is phylogenetically very similar to AT1G21780 which has been shown to bind to Cullin3a and 3b (Gingerich *et al.*, 2005). Together, BTB and Cullin proteins form part of the E3 ligase complex that ubiquitinates proteins and targets them for protein degradation. Thus, it is quite likely that POZ is also involved in targeting proteins for degradation.

UKTF was identified as a possible MADS-Box interacting protein1 (MIP1) containing the MIP1 domain, which is a basic leucine zipper region. In Antirrhinum, a MIP1 protein was found to interact with the C class proteins, PLENA and FARINELLI, and the E class proteins, DEFH72 and DEFH200, forming a ternary complex (Causier *et al.*, 2003). It was hypothesised that MIP1 enables interaction between the two groups of proteins to ensure proper development of carpels in flowers. The equivalent C and E class genes in Arabidopsis are *AGAMOUS* and the *SEPALLATA* genes, respectively (Gutierrez-Cortines and Davies, 2000). Thus, UKTF and other MIP1 proteins in Arabidopsis may be similarly involved in ensuring carpel development.

With the exception of *RAP2.12* (Licausi *et al.*, 2011; Papdi *et al.*, 2015), none of the other genes have been shown to have a role in drought or any other stress response. Thus VBSSM has the potential to identify genes and gene networks that are not known to be stress-responsive. However, the theoretical identification of these potentially important genes must be verified experimentally by analysing the role of these proteins in plants during drought.

To this end, as shown in Chapter 3, T-DNA insertional lines were screened to identify loss-of-function mutants of the above seven genes. Two knockouts were obtained for *AGL22*, *RAP2.12*, *BHLH038*, *ANL2* and *POZ*. Only one knockout was obtained for *FD* (Wigge *et al.*, 2005; Riboni *et al.*, 2013) which was kindly provided by Dr. Lucio Conti from the University of Milan, while a knockdown was screened from T-DNA insertional lines obtained from NASC. The two positive T-DNA insertional lines identified for *UKTF* were found to be overexpressing the gene, and no loss-of-function mutants could be identified.

In addition to the loss-of-function mutants, plants overexpressing the genes of interest were also created by transforming wild-type (WT) Col-0 plants with pEarleyGate (Earley *et al.*, 2005) overexpression constructs, and screening the seeds to obtain gain-of-function mutants for each of the genes. Only one non-segregating, overexpressing line

each for *AGL22* and *FD* were used for further analysis. The knockouts and overexpressors identified in Chapter 3 were phenotyped under drought conditions, as described in Chapter 4.

6.2. Flowering time genes appear to play a role during drought

6.2.1. *AGL22* was the most drought-responsive hub gene

Early flowering knockouts of *AGL22* showed a quicker drying rate due to a larger rosette area compared to Col-0 plants. The highly overexpressing AGL22ox-10 (Masiero *et al.*, 2004; Riboni *et al.*, 2013; obtained from Dr. Lucio Conti of the University of Milan) showed no difference in drying rate with the wild-type. On the other hand, the mild overexpressor AGL22ox-2 had a larger rosette area and quicker drying rate compared to WT plants, and thus was more like the knockouts in this respect. However, AGL22ox-2 plants also produced the characteristic sepaloid petals similar to the AGL22ox-10 line (Masiero *et al.*, 2004). Thus the phenotype observed may in fact be due to the mild overexpression of *AGL22* and not gene silencing, although multiple lines expressing different levels of *AGL22* will be required to verify if that is so.

It was also seen that during drought, the overexpressors produced higher levels of H_2O_2 in the droughted plants compared to the watered plants. On the other hand, very little difference in H_2O_2 levels was observed between the droughted and control knockout plants. Since the expression levels of *AGL22* increase during drought, it is possible that *AGL22* positively regulates H_2O_2 levels during drought leading to increased H_2O_2-mediated signalling events. The measurement of signalling genes in *AGL22* overexpressors and knockouts subjected to drought will help to verify if it is involved in signalling during drought.

In Chapter 5, the gene networks for *AGL22* and some of their connections were tested using qPCR. Many of the connections predicted for *AGL22* were found to be true and indicated that *AGL22* might regulate a number of stress-responsive genes, such as *DREB1A* and *WRKY20*. It was also seen that *AGL22* may play a role in preparing the plants for effective reproduction when the drought has passed, by the regulation of *DCL4* and *ARF2*. *AGL22* comes up early in the time-series analysis (day 7) and this indicates that it may play a key role in regulating the plant's response to drought; the fact that it regulates important stress-responsive genes such as *DREB1A* and *WRKY20* reinforces

this, and it will be interesting to test if other key drought-stress regulatory genes are also regulated by *AGL22*.

6.2.2. *FD* and *UKTF* may also be important regulatory genes

The knockout of *FD* showed a quicker drying rate during drought compared to the wild-type, and this mutant was also found to have a larger rosette area. On the other hand, the overexpressor and the knockdown mutants showed no difference in drying rate compared to the wild-type. Since only one knockout line for *FD* could be obtained, it is important to isolate and test other knockouts of *FD* to verify this phenotype and its possible role in drought. It is also necessary to test other stress parameters such as ROS levels and electrolyte leakage, and to determine the water status of mutants of *FD* by measuring rLWC, to ascertain if the gene does indeed play a role in drought stress response. *FD* is a late-responsive gene in the time-series and is induced at a low rSWC. This may indicate that *FD* may have more of a functional role during drought, rather than a regulatory one, although this must be explored further.

Two gain-of-function T-DNA lines of *UKTF* were obtained and phenotyped under droughted conditions, however, there appeared to be no difference between the two lines and the wild-type. Even though the overexpressing lines of *UKTF* did not show a drought phenotype, it is important to isolate a knockout for *UKTF* to properly assess if this gene plays a role in drought. Also, as it a member of a protein family, there may be functional redundancy with other proteins of the family, and multiple knockouts may be required to study the role of *UKTF*.

Some of the connections in the gene regulatory network for *UKTF* were tested as this gene product could potentially be involved in reproduction and carpel development. Only *SIS8* was found to be regulated by it, indicating that *UKTF* may integrate sugar signalling in the drought response. Also, knockouts of *SIS8* are resistant to negative effects of sugar and able to accumulate more sugars without growth inhibition (Huang *et al.*, 2014). This may indicate that *UKTF* mediates osmo-protection during drought. *SIS8* is also a negative regulator of salt stress and knockouts of this gene were salt-tolerant (Gao and Xiang, 2008). It may be similarly involved in drought as it is down-regulated by *UKTF*. The fact that *UKTF* is induced early during drought (day 6), it would indicate that down-regulation of *SIS8* is important for the plant to successfully withstand the drought stress.

From the analysis in this study, it appears that flowering time and flower development are important themes during drought. However, the age of plants sampled has to be taken in to consideration, as 5-week old plants maintained under short days were used for the transcriptomics analysis. It would be interesting to test the gene network in drought-stressed plants of a younger developmental stage. It would be interesting to study other flowering time mutants under drought, especially as there is a known correlation between flowering time and drought (McKay *et al.*, 2003).

6.2.3. The relationship between drought response and flowering and productivity

A number of studies have identified a relationship between drought stress response and flowering time. McKay *et al.* (2003) found a positive correlation between flowering time and water-use efficiency (ratio of carbon assimilation to transpirational water loss), indicating that plants that flower later have better water-use efficiency. They suggested that this indicates a trade-off between drought adaptation and drought escape – plants that escape drought are less likely to be water-use efficient, while plants that flower later are able to recover eventually from the effects of the stress and produce better yield. Corroborating this, Schmalenbach *et al.* (2014) observed that under mild drought stress plants that flowered later produced better seed yield because the longer vegetative phase allowed the plants to recover from the stress. Plants that flowered early produced less seed as flowering occurred during the drought. They also stated that the best strategy for plants to reproduce during drought depended on the type of drought experienced – drought avoidance is better under mild drought stress, while drought escape is the better strategy under severe drought conditions.

Riboni *et al.* (2013; 2016) identified that drought escape is promoted by abscisic acid and the flowering time genes *Flowering Locus T* (*FT*), *Twin Sister of FT* (*TSF*), *Constans* (*CO*) and *Suppressor of Overexpression of Constans 1* (*SOC1*). These genes are positive regulators of floral transition, and this further shows that flowering time and drought response, particularly drought escape, are co-regulated.

In addition, it was also seen that flower development were affected by drought stress which led to reduced seed yield (Su *et al.*, 2013). The development of stamen, in particular, was affected and the length of the filaments in flowers of drought-stressed plants was shortened compared to that of the well-watered plants. This led to poor pollination and, consequently, sterility. However, as the drought continued and the plants

acclimatised to the drought, flowers that developed later were able to mature and produce seed, though the yield was greatly reduced compared to the well-watered plants.

Thus flowering time and development are closely linked with response to drought stress, and this is highlighted by the fact that *AGL22*, *FD* and *UKTF* in the time-series and VBSSM modelling. Thus VBSSM has the potential to identify other genes involved in flowering that have not been analysed.

6.3. The other hub genes did not show a drought phenotype

Knockouts of *RAP2.12*, *BHLH038*, *ANL2* and *POZ* were phenotyped under droughted conditions but showed no difference was observed between the mutants and the WT plants. Gene regulatory networks for *RAP2.12* and *BHLH038* were also tested by qPCR. Only a few of the connections were tested in each network, but it appeared that none of these connections held true. The lack of negative results for *RAP2.12*, *BHLH038* and *POZ* may be attributed to possible functional redundancy for these genes.

POZ belongs to a family of BTB/POZ proteins in Arabidopsis, consisting of 80 members (Gingerich *et al.*, 2005). The lack of phenotype between the knockouts and the WT may due to functional redundancy between POZ and another member of the family, possibly AT1G21780. As these proteins may be involved in targeting proteins for degradation, a double knockout of both of these genes could lead to a build-up of proteins that act antagonistically to vital signalling pathways during drought stress, which could have a detrimental effect on the plant (Lyzenga and Stone, 2012).

BHLH038 has 80% sequence identity with *BHLH039*, another gene involved in iron homeostasis, and both are likely to be functionally redundant (Yuan *et al.*, 2008). A double mutant will be required to verify if *BHLH038* is drought-responsive. Maintaining iron homeostasis in the cell is important as excess iron can lead to increased ROS production through the Fenton reaction (Kampfenkel *et al.*, 1995; Gellego *et al.*, 1996), leading to ROS imbalance and consequently oxidative stress (Moran *et al.*, 1994; Munné-Bosch and Peñuelas, 2004; Sharma and Dubey, 2005).

Papdi *et al.* (2015) showed that inducible expression of either *RAP2.12* or *RAP2.3* conferred plants with tolerance to osmotic stress, while only the double knockout mutant of these genes was sensitive to osmotic stress. It is therefore likely that *RAP2.12* is

functionally redundant with *RAP2.3*, and possibly *RAP2.2,* (Gibbs *et al.*, 2014) and may explain why the single knockout did not show a drought phenotype.

RAP2.12 is known to be involved in the hypoxic response, particularly during flooding (Licausi *et al.*, 2011). During drought, water deficit restricts diffusion of oxygen to nodules and nodule activity is oxygen-limited (Serraj and Sinclair, 1996). Thus, hypoxia may be sensed in the roots during drought and *RAP2.12* (and homologous genes) may be involved in conveying and sensing this signal in the shoots.

6.4. Comment on the methods used in this work

VBSSM only considers gene expression as a linear event but actually it is non-linear (Bansal *et al.*, 2007). Non-parametric methods like Causal Structure Identification (CSI) model gene expression as a non-linear event and might provide an improved method over VBSSM (Penfold *et al.*, 2012). Even though VBSSM is one of the best performers for modelling gene networks, non-parametric methods outperform dynamic Bayesian models and can handle larger datasets (Penfold and Wild, 2011). However, these methods are computationally more intensive and thus VBSSM provides a balance between accuracy and speed.

Of the four networks tested, only the network for *AGL22* appeared to hold true for a number of the connections. However, only a much smaller number of genes were tested in the other three networks and it could be that this number was too small to make a meaningful statement about the accuracy levels of the predicted models. Also, the accuracy of the prediction may be improved by first verifying connections between hub genes and their child nodes in multiple gene network models i.e., *in silico*, before testing them *in planta*.

In this study, when drought was applied to test mutants of the hub genes relative to the wild-type, it could be seen from the measurements that the plants were not subjected to a severe drought, and in some cases appeared to be better off than the watered plants. This could be because the plants underwent a slow, progressive drought and may have been able to acclimatise to the drought. This was confirmed when the electrolyte leakage data was analysed, indicating that the droughted plants – WT and mutants – had reduced leakage of ions compared to the well-watered plants. Thus, the nature of the drought experienced by the plants in this work is a gradual drought and not the sudden, severe

drought that many studies in literature perform, such as cutting off the rosette and allowing to dehydrate on the bench (Seki *et al.*, 2002).

It is important to make the distinction between mild drought stress and severe drought stress, as it has been shown that plants respond differently to different drought stress intensities (Skirycz *et al.*, 2011). They showed that mutants that were shown to be drought resistant compared to WT plants under severe drought did not show better performance under mild drought. The response of a plant to drought depends on the severity and duration of the drought and the developmental stage of the plants at the time of the drought (Claeys and Inzé, 2013). Any plant experiencing unfavourable conditions that affects its growth and productivity is considered stressed, and hence it is also important to study the effect of mild, but prolonged, drought stress on plants. Claeys and Inzé (2013) suggested that it may be necessary to engineer crop plants that are tolerant to the conditions that they will be expected to grow in, i.e., mild or severe drought conditions, as a 'one size fits all' approach will not work.

The work described in Chapter 4 to test the drought response of mutants of hub genes compared to WT plants and included the evaluation of water status, photosynthetic performance, cell stress status and yield and productivity. This was to obtain a more all-rounded idea of the effect of the mutation on the performance of the plant under drought. Many studies do not assess the performance due to the drought and they particularly give no idea of effect of the drought on productivity (Verslues *et al.* 2006; Lawlor 2013).

To improve on this work, a better way of comparing a mutant of a gene of interest with a WT plant would be to use the analysis proposed by Lawlor (2013) and Claeys *et al.* (2014), to measure the plant's performance over the duration of the stress and to compare between plants with the same rLWC rather than rSWC, as plants of different genotype can take up soil water at different rates. By comparing plants of different genotype at the same rLWC, it enables a better comparison of the effect of the gene of interest on the drought response. Also, performing physiological measurements over the course of the drought would prove much more informative than single-time point measurements, as it would reveal the point during the drought when the gene of interest has an effect on the plant's stress response Claeys *et al.* (2014). Using these methods and changes to phenotyping mutant lines would help accurately assess potential target genes and be more informative in understanding their role in drought stress response.

6.5. Conclusion

- VBSSM was applied to time-series transcriptomics data and used to identify potentially important genes and gene networks involved in drought stress response.
- Two of the seven genes (*AGL22* and *FD*) analysed showed a drought phenotype, while four of the remaining genes (*RAP2.12*, *BHLH038*, *UKTF* and *POZ*) may be functionally redundant and may explain why single knockouts did show a drought phenotype.
- The analysis in this work implies that flowering time and development are important developmental processes linked with drought, which ties in with work done in other studies.
- However, *AGL22* has not been implicated in drought stress response before, but has been shown to induce the expression of stress-responsive genes in this work.
- Thus, VBSSM may have potentially identified genes and gene families that are not known to be drought-responsive.

References

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosakawa, D., and Shinozaki, K. 1997. Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. Plant Cell 9: 1859–1868.

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., & Araki, T., (2005). FD, a bZIP Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. Science, 309, 1052–1056.

Abebe T, Guenzi AC, Martin B, Chushman JC: Tolerance of mannitol-accumulating transgenic wheat to water stress and salinity. Plant Physiol 2003, 131:1748-1755.

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301: 653–657

An Y, Zhang M, Liu G, Han R, Liang Z (2013) Proline Accumulation in Leaves of Periploca sepium via Both Biosynthesis Up-Regulation and Transport during Recovery from Severe Drought. PLoS ONE 8(7): e69942

Aneja, B., Yadav, N.R., Kumar, N. and Yadav, R.C., 2015. Hsp transcript induction is correlated with physiological changes under drought stress in Indian mustard. Physiology and Molecular Biology of Plants, 21(3), pp.305-316.

Apel, K., & Hirt, H., (2004). REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. Annu. Rev. Plant Biol., 55, 373–399.

Apostol I, Heinstein PF, Low PS. 1989. Rapid stimulation of an oxidative burst during elicidation of cultured plant cells. Role in defense and signal transduction. Plant Physiol. 90:106–16

Avonce N, Leyman B, Mascorro-Gallardo JO, Van Dijck P, Thevelein JM, Iturriaga G: The Arabidopsis trehalose-6-P synthase AtTPS1 gene is a regulator of glucose, abscisic acid, and stress signaling. Plant Physiol 2004, 136:3649-3659

Bansal, M., Belcastro, V., Ambesi-Impiombato, A. & di Bernardo, D. 2007 How to infer gene networks from expression profiles. Mol. Syst. Biol. 3, 78.

Bardwell V.J., Treisman R. (1994) The POZ domain: A conserved protein–protein interaction motif. Genes & Dev. 8:1664–1677.

Bartels, D. & Sunkar, R. (2005). Drought and Salt Tolerance in Plants. Crit. Rev. Plant Sci., 24, 23–58.

Beal, M.J., Falciani, F., Ghahramani, Z., Rangel, C. & Wild, D.L. (2005). A Bayesian approach to reconstructing genetic regulatory networks with hidden factors. Bioinformatics, 21, 349–356.

Bechtold U., CA Penfold, DJ Jenkins, R Legaie, JD Moore, T Lawson et al. (2016). Time-series transcriptomics reveals that AGAMOUS-LIKE22 links primary metabolism to developmental processes in drought-stressed Arabidopsis, The Plant Cell. February 3, 2016 TPC2015-00910-LSB

Bechtold, U., Lawson, T., Mejia-Carranza, J., Meyer, R.C., Brown, I.R., Altmann, T., Ton, J. & Mullineaux, P.M. (2010). Constitutive salicylic acid defences do not compromise seed yield, drought tolerance and water productivity in the Arabidopsis accession C24. Plant Cell Environ., 33, 1959–1973.

Bianchi, G., Gamba, A., Murelli, C., Salamini, F. & Bartels, D. (1991). Novel carbohydrate metabolism in the resurrection plant Craterostigma plantagineum. Plant J., 1, 355–359.

Blum A, Ebercon A. 1981. Cell membrane stability as a measure of drought and heat tolerance in wheat. Crop Science 21: 43-47.

Booker, J., Auldridge, M., Wills, S., McCarty, D., Klee, H., & Leyser, O., (2004). MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. Curr. Biol., 14, 1232–1238.

Bouchabke O, Chang F, Simon M, Voisin R, Pelletier G, Durand-Tardif M (2008) Natural variation in Arabidopsis thaliana as a tool for highlighting differential drought responses. PLoS One 3: e1705

Braberg H., Webb B. M., Tjioe E., Pieper U., Sali A., Madhusudhan M. S. (2012) SALIGN: a web server for alignment of multiple protein sequences and structures. Bioinformatics 28, 2072–2073

Bravo, L.A., Gallardo, J., Navarrete, A., Olave, N., Martinez, J., Alberdi, M., Close, T.J. and Corcuera, L.J. (2003) Cryoprotective activity of a cold-induced dehydrin purified from barley. Physiol. Plant. 118, 262–269.

Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y., Penfold, C.A., Jenkins, D., Zhang, C., Morris, K., Jenner, C., Jackson, S., Thomas, B., Tabrett, A., Legaie, R., Moore, J.D., Wild, D.L., Ott, S., Rand, D., Beynon, J., Denby, K., Mead, A., & Buchanan-Wollaston, V., (2011). High-Resolution Temporal Profiling of Transcripts during Arabidopsis Leaf Senescence Reveals a Distinct Chronology of Processes and Regulation. Plant Cell, 23, 873–894.

Bruinsma, J. 2009. The resource outlook to 2050: By how much do land, water use and crop yields need to increase by 2050? 33 pp. Expert Meeting on How to Feed the World in 2050. Rome, FAO and ESDD.

Cabello, J.V. and Chan, R.L. (2012) The homologous homeodomain-leucine zipper transcription factors HaHB1 and AtHB13 confer tolerance to drought and salinity stresses via the induction of proteins that stabilize membranes. Plant Biotechnol. J. 10, 815–825.

Causier, B., Cook, H., & Davies, B., (2003). An Antirrhinum ternary complex factor specifically interacts with C-function and SEPALLATA-like MADS-box factors. Plant Mol. Bio., 52, 1051-1062.

Cha, J.-Y., Kim, J. Y., Jung, I. J., Kim, M. R., Melencion, A., Alam, S. S., et al. (2014). NADPH-dependent thioredoxin reductase A (NTRA) confers elevated tolerance to oxidative stress and drought. Plant Physiol. Biochem. 80, 184–191.

Chalker-Scott, L., (1999). Environmental significance of anthocyanins in plant stress responses. Photochemistry and photobiology, 70, 1–9.

Chaves, M.M., Maroco, J.P. & Pereira, J.S. (2003). Understanding plant responses to drought - from genes to the whole plant. Funct. Plant Biol., 30, 239–264.

Chen, L., Lee, J.H., Weber, H., Tohge, T., Witt, S., Roje, S., Fernie, A.R., and Hellmann, H. (2013). Arabidopsis BPM proteins function as substrate adaptors to a cullin3-based E3 ligase to affect fatty acid metabolism in plants. Plant Cell 25:2253–2264.

Cheng M-C, Hsieh E-J, Chen J-H, Chen H-Y, Lin T-P. 2012. Arabidopsis RGLG2, functioning as a RING E3 ligase, interacts with AtERF53 and negatively regulates the plant drought stress response. Plant Physiology 158: 363–375.

Cho, K.H., Choo, S.M., Jung, S.H., Kim, J.R., Choi, H.S. & Kim, J. (2007). Reverse engineering of gene regulatory networks. IET Syst. Biol., 1, 149–163.

Choi, H. I., Hong, J. H., Ha, J. O., Kang, J. Y., and Kim, S. Y. 2000. ABFs, a family of ABA-responsive element binding factors. J. Biol. Chem. 275: 1723–1730.

Claeys, H. and Inzé, D., 2013. The agony of choice: how plants balance growth and survival under water-limiting conditions. Plant physiology, 162(4), pp.1768-1779.

Claeys, H., Van Landeghem, S., Dubois, M., Maleux, K. and Inzé, D., 2014. What is stress? Dose-response effects in commonly used in vitro stress assays. Plant physiology, 165(2), pp.519-527.

Close, T.J., Kortt, A.A., and Chandler, P.M. (1989). A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. Plant MOI. Biol. 13, 95-108.

Clough, S.J., & Bent, A.F., (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J., 16, 735–743.

Coello, P., Hey, S.J. & Halford, N.G. (2011). The sucrose non-fermenting-1-related (SnRK) family of protein kinases: potential for manipulation to improve stress tolerance and increase yield. J. Exp. Bot., 62, 883–893.

Comprehensive Assessment of Water Management in Agriculture. 2007. Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture. London: Earthscan, and Colombo: International Water Management Institute.

Crevillén P, Ventriglia T, Pinto F, Orea A, Mérida Á, Romero JM (2005) Differential pattern of expression and sugar regulation of Arabidopsis thaliana ADP-glucose pyrophosphorylase-encoding genes. J Biol Chem 280: 8143–8149

Cuin, T.A. and Shabala, S. (2007) Compatible solutes reduce ROS-induced potassium efflux in Arabidopsis roots. Plant Cell Environ. 30, 875–885

de Nadal, E., Alepuz, P. A., and Posas, F. 2002. Dealing with osmostress through MAP kinase activation. EMBO Reports 3: 735–740.

Dixon DP, Edwards R . Enzymes of tyrosine catabolism in Arabidopsis thaliana. Plant Science 2006;171:360-366.

Earley K, Smith M, Weber R, Gregory B, Poethig R (2010) An endogenous F-box protein regulates ARGONAUTE1 in Arabidopsis thaliana. Silence 1:15.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., & Pikaard, C.S., (2006). Gateway-compatible vectors for plant functional genomics and proteomics. Plant J., 45, 616–629.

Edwards, K., Johnstone, C., & Thompson, C., (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res., 19, 1349.

Egea Gutierrez-Cortines M, Davies B. Beyond the ABCs: ternary complex formation in the control of floral organ identity. Trends Plant Sci 2000;5:471–6.

Eisen, M.B. and Brown, P.O. (1999) DNA arrays for analysis of gene expression. Meth. Enzymol. 303, 179-205.

Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER (2004) A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. Plant Cell 16: 2923–2939.

Faith J, Gardner T (2005) Reverse-engineering transcription control networks. Phys Life Rev 2: 65–88

Forsburg SL, Guarente L (1989) Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. Genes Dev 3: 1166–1178.

Fujita Y, Sayama H, Kidokoro S, Maruyama K, Mizoi J, et al. (2010). AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. Plant J.; 61:672-685.

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, et al. 2005. AREB1 is a transcription activator of novel ABRE-dependent ABA-signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell 17:3470–88

Fukao, T., Yeung, E., & Bailey-Serres, J., (2011). The Submergence Tolerance Regulator SUB1A Mediates Crosstalk between Submergence and Drought Tolerance in Rice. The Plant Cell, 23, 412–427.

Galau, G. A, Hughes, D. W., and Dure, III L. 1986. Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. Plant Mol. Biol. 7: 155–170.

Gallego SM, Benavides MP, Tomaro ML.1996. Effect of heavy metal ion excess on sunflower leaves: evidence for involvement of oxidative stress. Plant Science121,151–159

Gao, L. and Xiang, C.B. (2008) The genetic locus At1g73660 encodes a putative MAPKKK and negatively regulates salt tolerance in Arabidopsis. Plant Mol. Biol. 67, 125–134.

Garg AK, Kim J-K, Owens TG, Ranwala AP, Choi YD, Kochian LV, Wu R: Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. Proc Natl Acad Sci USA 2002, 99:15898-15903.

Gibbs DJ, Isa NM, Movahedi M, Lozano-Juste J, Mendiondo GM, Berckhan S, Marín-de la Rosa N, Conde JV, Correia CS, Pearce SP et al. 2014. Nitric oxide sensing in plants mediated by proteolytic control of group VII ERF transcription factors. Molecular Cell 53: 369–379.

Gibson, S.I., Laby, R.J., and Kim, D. (2001). The sugar-insensitive1 (sis1) mutant of Arabidopsis is allelic to ctr1. Biochem. Biophys. Res. Commun. 280, 196–203.

Gingerich, D.J., Gagne, J.M., Salter, D.W., Hellmann, H., Estelle, M., Ma, L., & Vierstra, R.D., (2005). Cullins 3a and 3b assemble with members of the broad complex/tramtrack/bric-a-brac (BTB) protein family to form essential ubiquitin-protein ligases (E3s) in Arabidopsis. J. Biol. Chem., 280, 18810–18845.

Girousse C, Bournoville R, Bonnemain JL (1996) Water deficit-Induced changes in concentrations in proline and some other amino acids in the phloem sap of alfalfa. Plant Physiol 111: 109–113.

Gómez-Mena, C., de Folter, S., Costa, M.M.R., Angenent, G.C., and Sablowski, R. (2005). Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. Development 132, 429–438.

Gosti, F., Beaudoin, N., Serizet, C., Webb, A.A.R., Vartanian, N. & Giraudat, J. (1999). ABI1 Protein Phosphatase 2C Is a Negative Regulator of Abscisic Acid Signalling. Plant Cell, 11, 1897–1909.

Guilbault GG, Brignac PJ, Zimmer M (1967) Homovanillic acid as a fluorometric substrate for oxidative enzymes: analytical applications of the peroxidase, glucose oxidase, and xanthine oxidase systems. Anal Chem 40:190–196.

Gutierrez, R.A. (2005). Systems Biology for the Virtual Plant. Plant Physiol., 138, 550–554.

Haake,V., Cook, D. Riechmann, J. L., Pineda, O., Thomashow, M. F., and Zhang, J. Z. 2002. Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. Plant Physiol. 130: 639–648.

Hakoshima, T. Leucine Zippers John Wiley & Sons, Ltd (2014).

Hara, M., Terashima, S. and Kuboi, T. (2001) Characterization and cryoprotective activity of cold-responsive dehydrin from Citrus unshiu. J. Plant Physiol. 158, 1333–1339.

Harb, A., Krishnan, A., Ambavaram, M.M.R. & Pereira, A. (2010). Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth. Plant Physiol., 154, 1254–1271.

Hartley, J.L., Temple, G.F. and Brasch, M.A. (2000) DNA cloning using in vitro site-specific recombination. Genome Res. 10, 1788–1795.

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

Haüshul, K., Andersson, B., and Adamska, I. 2001. A chloroplast DegP2 protease performs the primary cleavage of the photodamaged D1 protein in plant photosystem II. EMBO J. 20: 713–722.

He J., Gwadz M., Hurwitz D.I. CDD: NCBI's Conserved Domains Database. Nucleic Acids Res. 2015;43:D222-D226.

Hickman, R., C. Hill, C. Penfold, E. Breeze, L. Bowden, J. Moore, P. Zhang, A. Jackson, E. Cooke, F. Bewicke-Copley, A. Mead, J. Beynon, D. Wild, K. Denby, S. Ott, and V. Buchanan-Wollaston (2013): “A local regulatory network around three NAC transcription factors in stress responses and senescence in Arabidopsis leaves,” Plant J., 75, 26–39.

Hill CM, Pearson SA, Smith AJ, Rogers LJ (1985) Inhibition of chlorophyll synthesis in Hordeum vulgare by 3-amino 2,3-dihydrobenzoic acid (gabaculin). Biosci Rep 5:775–781.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. and Zimmermann, P. (2008) Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. Adv. Bioinformatics, 420747.

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009;4(1):44-57.

Huang Y, Li CY, Qi Y, Park S, Gibson SI. SIS8, a putative mitogen-activated protein kinase kinase kinase, regulates sugar-resistant seedling development in Arabidopsis. Plant J. 2014; 77: 577–588.

Ichimura, K., Mizoguchi, T., Yoshida, R., Yuasa, T., and Shinozaki, K. 2000. Various abiotic stresses rapidly activate Arabidopsis MAP kinases AtMAPK4 and AtMPK6. Plant J. 24: 655–665.

Islam T, Manna M, Reddy MK (2015) Glutathione Peroxidase of Pennisetum glaucum (PgGPx) Is a Functional Cd2+ Dependent Peroxiredoxin that Enhances Tolerance against Salinity and Drought Stress. PLoS ONE 10(11): e0143344.

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., Tabata, S., Kakubari, Y., Yamaguchi-Shinozaki, K. & Shinozaki, K. (2001). Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. Plant J., 27, 325–333.

Jiang M, Zhang J. Water stress-induced abscissic acid accumulation triggers the increased generation of reactive oxygen species and up-regulates the activities of antioxidant enzymes in maize-leaves. Journal of Experimental Botany 2002;53:2401-2410.

Jiang Y, Liang G, Yu D (2012) Activated expression of WRKY57 confers drought tolerance in Arabidopsis. Mol Plant 5:1375–1388

Jones DT. 2007. Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics 23: 538–544.

Juenger, T.E., Mckay, J.K., Hausmann, N., Keurentjes, J.J.B., Sen, S., Stowe, K.A., Dawson, T.E., Simms, E.L., & Richards, J.H., (2005). Identification and characterization of QTL underlying whole-plant physiology in Arabidopsis thaliana: 13C, stomatal conductance and transpiration efficiency. Plant Cell Environ., 28, 697–708.

Kallberg, M.; Margaryan, G.; Wang, S.; Ma, J.; Xu, J., RaptorX server: a resource for template-based protein structure modeling. Methods Mol Biol, 2014, 1137, 17-27.

Kampfenkel, K., Van Montagu, M., and Inzé, D. (1995). Effects of iron excess on Nicotiana plumbaginifolia plants. Plant Physiol. 107, 725–735.

Kang SG, Park E, Do KS (2009) Identification of a pathogen-Induced glycine max transcription factor GmWRKY1. Plant Pathol J 25: 381–388.

Kang, J., Choi, H., Im, M., and Kim, S.Y. 2002. Arabidopsis basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. Plant Cell 14: 343–357.

Kavi Kishor, P. B., Hong, Z., Miao, G.-H., Hu, C.-A. A., andVerma, D. P. S. 1995. Over-expression of -pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. Plant Physiol. 108: 1387–1394.

Kelley, L.A. & Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. 4, 363–371 (2009).

Kiddle, S. J., Windram, O. P. F., McHattie, S., Mead, A., Beynon, J., Buchanan-Wollaston, V., Denby, K. J. & Mukherjee, S. 2009 Temporal clustering by affinity propagation reveals transcriptional modules in Arabidopsis thaliana. Bioinformatics 26, 355–362.

Kilian J, Whitehead D, Horak J,Wanke D, Weinl S, Batistic O, et al. 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. The Plant Journal 50: 347–363.

Kim S., Sun H., Tomchick D.R., Yu H., Luo X. (2012) Structure of human Mad1 C-terminal domain reveals its involvement in kinetochore targeting Proc. Natl. Acad. Sci. USA, 109, pp. 6549–6554

Koncz, C., & Schell, J., (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Mol. Gen. Genet., 204, 383–396.

Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper J. 2002. Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. Plant Physiology 130, 2129–2141.

Kubo, H., Peeters, A.J.M., Aarts, M.G.M., Pereira, A., & Koornneef, M., (1999). ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. Plant Cell, 11, 1217–1226.

Lawlor DW (2013) Genetic engineering to improve plant performance under drought: physiological evaluation of achievements, limitations, and possibilities. J Exp Bot 64: 83–108

Lawson T, Weyers JDB.1999. Spatial and temporal variation in gas exchange over the lower surface of Phaseolus vulgaris primary leaves. Journal of Experimental Botany 50, 1381–1391.

Lee, J.H., Yoo, S.J., Park, S.H., Hwang, I., Lee, J.S., & Ahn, J.H., (2007). Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. Genes & Development, 21, 397–402.

Leyva-González MA, Ibarra-Laclette E, Cruz-Ramı´rez A, Herrera-Estrella L (2012) Functional and Transcriptome Analysis Reveals an Acclimatization Strategy for Abiotic Stress Tolerance Mediated by Arabidopsis NF-YA Family Members. PLoS ONE 7(10): e48138.

Li WX, Oono Y, Zhu J, He XJ, Wu JM, et al. (2008) The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and post-transcriptionally to promote drought resistance. Plant Cell 20: 2238–2251.

Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., Helliwell, C.A., Ito, T., Meyerowitz, E., & Yu, H., (2008). A Repressor Complex Governs the Integration of Flowering Signals in Arabidopsis. Dev. Cell, 15, 110–120.

Li, S., Assmann, S.M., & Albert, R., (2006). Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. PLoS Biology, 4, 1732–1748.

Licausi, F., Kosmacz, M., Weits, D.A., Giuntoli, B., Giorgi, F.M., Voesenek, L.A.C.J., Perata, P., & van Dongen, J.T., (2011). Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. Nature, 479, 419–423.

Lin R.-C., Park H.-J., Wang H.-Y., Role of Arabidopsis RAP2.4 in regulating light and ethylene-mediated developmental processes and drought stress tolerance, Mol. Plant 1 (2008) 42–57.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell 10: 1391–1406.

Locke JCW, Millar AJ, Turner MS (2005a) Modelling genetic networks with noisy and varied experimental data: the circadian clock in Arabidopsis thaliana. J Theor Biol 234: 383–393

Locke JCW, Southern MM, Kozma-Bognar L, Hibberd V, Brown PE, Turner MS, Millar AJ (2005b) Extension of a genetic network model by iterative experimentation and mathematical analysis. Mol Syst Biol 1: 13

Locke, J. C. W., Kozma-Bognár, L., Gould, P. D., Fehér, B., Kevei, E., Nagy, F., Turner, M. S., Hall, A. & Millar, A. J. 2006 Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. Mol. Syst. Biol. 2, 59.

Lyzenga WJ, Stone SL (2012) Abiotic stress tolerance mediated by protein ubiquitination. J Exp Bot 63: 599–616

Lyzenga, W.J. and Stone, S.L., 2012. Regulation of ethylene biosynthesis through protein degradation. Plant signaling & behavior, 7(11), pp.1438-1442.

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, 1064–1068.

Mafakheri, A., A. Siosemardeh, B. Bahramnejad, P.C. Struik, and Y. Sohrabi. 2010. Effect of drought stress on yield, proline and chlorophyll contents in three chickpea cultivars. Aust. J. Crop Sci. 4:580-585.

Marchler-Bauer A., Derbyshire M.K., Gonzales N.R., Lu S., Chitsaz F., Geer L.Y., Geer R.C.,

Masiero, S., Li, M.A., Will, I., Hartmann, U., Saedler, H., Huijser, P., Schwarz-Sommer, Z., and Sommer, H. 2004. INCOMPOSITA: A MADS-box gene controlling prophyll development and floral meristem identity in Antirrhinum. Development 131: 5981–5990.

Matsui A, Mizunashi K, Tanaka M, Kaminuma E, Nguyen AH, Nakajima M, Kim J-M, Nguyen DV, Toyoda T, Seki M (2014) tasiRNA-ARF pathway moderates floral architecture in Arabidopsis plants subjected to drought stress. BioMed Res Int 2014: 303451

Matsui, A., Ishida, J., Morosawa, T., Mochizuki, Y., Kaminuma, E., Endo, T.A., Okamoto, M., Nambara, E., Nakajima, M., Kawashima, M., Satou, M., Kim, J.-M., Kobayashi, N., Toyoda, T., Shinozaki, K., & Seki, M., (2008). Arabidopsis Transcriptome Analysis under Drought, Cold, High-Salinity and ABA Treatment Conditions using a Tiling Array. Plant Cell Physiol., 49, 1135–1149.

McKay, J. K., J. H. Richards, and T. Mitchell-Olds. 2003. Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. Mol. Ecol. 12:1137–1151.

Meng, L.-S., Wang, Z.-B., Yao, S.-Q. and Liu, A. (2015). The ARF2–ANT–COR15A gene cascade regulates ABA-signaling-mediated resistance of large seeds to drought in Arabidopsis. J Cell Sci. 128: 3922-3932

Mishra, G., Zhang, W., Deng, F., Zhao, J. & Wang, X. (2006). A bifurcating pathway directs abscisic acid effects on stomatal closure and opening in Arabidopsis. Science, 312, 264–266.

Mizoguchi, T., Irie, K., Hirayama, T., Hayashida, N., Yamaguchi-Shinozaki, K., Matsumoto, K., and Shinozaki, K. 1996. A gene encoding a mitogen-activated protein kinase kinase kinase is induced simultaneously with genes for a mitogen-activated protein kinase and an S6 ribosomal protein kinase by touch, cold, and water stress in Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA 93: 765–769.

Morán JF, Becana M, Iturbe-Ormaetxe I, Frechilla S, Klucas RV, Aparicio-Tejo P (1994) Drought induces oxidative stress in pea plants. Planta 194: 346-352

Mori, I.C. & Schroeder, J.I. (2004). Reactive oxygen species activation of plant Ca2+ channels. A signalling mechanism in polar growth, hormone transduction, stress signalling, and hypothetically mechano-transduction. Plant Physiol., 135, 702–708.

Mukherjee AK, Carp MJ, Zuchman R, Ziv T, Horwitz BA, Gepstein S. Proteomics of the response of Arabidopsis thaliana to infectionwith Alternaria brassicicola. J Proteomics 2010;73:709–20.

Mundy, J., Yamaguchi-Shinozaki, K., and Chua, N. H. 1990. Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice rab gene. Proc. Natl. Acad. Sci. USA 87: 1406–1410.

Munné-Bosch S. & Alegre L. (2000) Changes in carotenoids, tocopherols and diterpenes during drought and recovery, and the biological significance of chlorophyll loss in Rosmarinus officinalis plants. Planta 210, 925–931.

Munné-Bosch, S. and Peñuelas, J. (2004) Drought-induced oxidative stress in strawberry tree (Arbutus unedo L.) growing in Mediterranean field conditions. Plant Science 166, 1105–1110

Munnik T, Vermeer JE. 2010. Osmotic stress-induced phosphoinositide and inositol phosphate signalling in plants. Plant, Cell and Environment 33, 655–669.

Murata, Y., Pei, Z.M., Mori, I.C. & Schroeder, J. (2001). Abscisic acid activation of plasma membrane Ca2+ channels in guard cells requires cytosolic NAD (P) H and is differentially disrupted upstream and downstream of reactive oxygen species production in abi1-1 and abi2-1 protein phosphatase 2C mutants. Plant Cell, 13, 2513–2523.

Mustilli, A.C., Merlot, S., Vavasseur, A., Fenzi, F. & Giraudat, J. (2002). Arabidopsis OST1 protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reactive oxygen species production. Plant Cell, 14, 3089–3099.

Nagata T, Hara H, Saitou K, Kobashi A, Kojima K, Yuasa T, Ueno O (2012) Activation of ADP-Glucose Pyrophosphorylase Gene Promoters by a WRKY Transcription Factor, AtWRKY20, in Arabidopsis thaliana L. and Sweet Potato (Ipomoea batatas Lam.). Plant Prod Sci 15:10–18

Nakashima, K., Kiyosue, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. 1997. A nuclear gene, erd1, encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not only induced by water stress but also developmentally up-regulated during senescence in Arabidopsis thaliana. Plant J. 12: 851–861.

Nakata, M., Mitsuda, N., Herde, M., Koo, A. J., Moreno, J. E., Suzuki, K., Howe, G. A. and Ohme-Takagi, M. (2013). A bHLH-type transcription factor, ABA-INDUCIBLE BHLH-TYPE TRANSCRIPTION FACTOR/JA-ASSOCIATED MYC2-LIKE1, acts as a repressor to negatively regulate jasmonate signaling in Arabidopsis. Plant Cell 25(5): 1641-1656.

Neill Samuel O. & Gould Kevin S. (2003) Anthocyanins in leaves: light attenuators or antioxidants? Functional Plant Biology 30, 865–873.

Nelson DE, Repetti PP, Adams TR, Creelman RA, Wu J, et al. (2007) Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres. Proc Natl Acad Sci USA 104: 16450–16455.

Nishizawa, A., Yabuta, Y. & Shigeoka, S. (2008). Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. Plant Physiol., 147, 1251–1263.

Nugent T., Jones D.T. (2013) Membrane protein orientation and refinement using a knowledge-based statistical potential. BMC Bioinf., 14, 276.

Oberschall, A., Deak, M., T¨or¨ok, K., Saa, L., Vass, I., Kovacs, I., Feher, A., Dudits, D., and Horvath, G. V. 2000. A novel aldose/aldehyde reductase protects transgenic plants against lipid peroxidation under chemical and drought stresses. Plant J. 24: 437–446.

Ogbaga CC, Stepien P, Dyson BC, Rattray NJW, Ellis DI, Goodacre R, et al. (2016) Biochemical Analyses of Sorghum Varieties Reveal Differential Responses to Drought. PLoS ONE 11(5): e0154423.

Okushima Y, Mitina I, Quach HL, Theologis A (2005) AUXIN RESPONSE FACTOR 2 (ARF2): a pleiotropic developmental regulator. Plant J 43: 29–46

Osakabe, Y., Maruyama, K., Seki, M., Satou, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2005). Leucine-Rich Repeat Receptor-Like Kinase1 Is a Key Membrane-Bound Regulator of Abscisic Acid Early Signalling in Arabidopsis. Plant Cell, 17, 1105–1119.

Osakabe, Y., Mizuno, S., Tanaka, H., Maruyama, K., Osakabe, K., Todaka, D., Fujita, Y., Kobayashi, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2010). Overproduction of the Membrane-bound Receptor-like Protein Kinase 1, RPK1, Enhances Abiotic Stress Tolerance in Arabidopsis. J. Biol. Chem., 285, 9190–9201.

Papdi C, Perez-Salamo I, Joseph MP, Giuntoli B, Bogre L, Koncz C, et al. The low oxygen, oxidative and osmotic stress responses synergistically act through the ethylene response factor VII genes RAP2.12, RAP2.2 and RAP2.3. Plant J. 2015;82: 772–784.

Park, S.Y., Fung, P., Nishimura, N., Jensen, D.R., Fujii, H., Zhao, Y., Lumba, S., Santiago, J., Rodrigues, A., Chow, T.F., Alfred, S.E., Bonetta, D., Finkelstein, R., Provart, N.J., Desveaux, D., Rodriguez, P.L., McCourt, P., Zhu, J.-K., Schroeder, J., Volkman, B.F. & Cutler, S.R. (2009). Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. Science, 324, 1068–1071.

Passioura, J. B. (2002). Environmental biology and crop improvement. Funct. Plant Biol. 29, 537–546.

Pazos, F., Pietrosemoli, N., Garcia-Martin, J.A. and Solano, R. (2013) Protein intrinsic disorder in plants. Front. Plant Sci. 4, 363.

Penfold, C. A., V. Buchanan-Wollaston, K. Denby, and D. L. Wild (2012): “Nonparametric Bayesian inference for perturbed and orthologous gene regulatory networks,” Bioinformatics, 28, i233–241.

Penfold, C.A. and Wild, D.L., 2011. How to infer gene networks from expression profiles, revisited. Interface focus, 1(6), pp.857-870.

Peters, S., Mundree, S.G., Thomson, J.A., Farrant, J.M. & Keller, F. (2007). Protection mechanisms in the resurrection plant Xerophyta viscosa (Baker): both sucrose and raffinose family oligosaccharides (RFOs) accumulate in leaves in response to water deficit. J. Exp. Bot., 58, 1947–1956.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–1612.

Pilon-Smits EAH, Ebskamp MJM, Paul MJ, Jeuken MJW, Weisbeek PJ, Smeekens SCM: Improved performance of transgenic fructan-accumulating tobacco under drought stress. Plant Physiol 1995, 107:125-130

Pilon-Smits EAH, Terry N, Sears T, van Dun K: Enhanced droughtresistance in fructan-producing sugar beet. Plant Physiol Biochem 1999, 37:313-317.

Pinheiro C, Chaves MM: Photosynthesis and drought: can we make metabolic connections from available data? J Exp Bot. 2011, 62 (3): 869-882.

Porcel R, Azcón R, Ruiz-Lozano JM. Evaluation of the role of genes encoding for dehydrin proteins (LEA D-11) during drought stress in arbuscular mycorrhizal Glycine max and Lactuca sativa plants. Journal of Experimental Botany 2005;56:1933-1942.

Potter, S., Uknes, S., Lawton, K., Winter, A.M., Chandler, D., Dimaio, J., Novitzky, R., Ward, E., and Ryals, J. (1993). Regulation of a hevein-like protein in Arabidopsis. Mol. Plant-Microbe Interact. 6, 680-681.

Prasad TK, Anderson MD, Martin BA, Stewart CR. 1994. Evidence for chilling-induced oxidative stress in maize seedlings and a regulatory role for hydrogen peroxide. Plant Cell. 6:65–74

Prasch CM, Ott KV, Bauer H, Ache P, Herich R, Sonnewald U (2015) ß-amylase1 mutant Arabidopsis plants show improved drought tolerance due to reduced starch breakdown in guard cells. J Exp Bot 66: 6059–6067

Proietti S, Bertini L, Van der Ent S, Leon-Reyes A, Pieterse CMJ, et al. (2011) Cross activity of orthologous WRKY transcription factors in wheat and Arabidopsis. J Exp Bot 62: 1975–1990

Raghavendra, A.S., Gonugunta, V.K., Christmann, A. & Grill, E. (2010). ABA perception and signalling. Trends Plant Sci., 15, 395–401.

Reiser, V., Raitt, D. C., and Saito, H. 2003. Yeast osmosensor Sln1 and plant cytokinin receptor Cre1 respond to changes in turgor pressure. J. Cell. Biol. 161: 1035–1040.

Ren, G.D., Zhou, Q., Wu, S.X., Zhang, Y.F., Zhang, L.G., Huang, J.R., Sun, Z.F., and Kuai, B.K. (2010). Reverse genetic identification of CRN1 and its distinctive role in chlorophyll degradation in Arabidopsis. J. Integr. Plant Biol. 52: 496–504.

Rentel, M. C., Lecourieux, D., Ouaked, F., Usher, S. L., Petersen, L., Okamoto, H., Knight, H., Peck, S. C., Grierson, C. S., Hirt, H., and Knight, M. R. 2004. OXI1 kinase is necessary for oxidative burst-mediated signalling in Arabidopsis. Nature 427: 858–861.

Rey, P., Pruvot, G., Becuwe, N., Eymery, F., Rumeau, D., and Peltier, G. A. 1998. A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in Solanum tuberosum L. plants. Plant J. 25: 97–107.

Rhodes D, Handa S, Bressan RA. Metabolic changes associated with adaptation of plant-cells to water-stress. Plant Physiology 1986;82:890-903

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 CONSTANS1. Plant Physiol., 162, 1706-1719.

Riboni, M., Test, A.R., Galbiati, M., Tonelli, C. and Conti, L., (2016). ABA-dependent control of GIGANTEA signalling enables drought escape via up-regulation of FLOWERING LOCUS T in Arabidopsis thaliana. Journal of Experimental Botany, p.erw384.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol. Biol. 53, 247–259.

Royal Society (2009). Reaping the benefits: Science and the sustainable intensification of global agriculture. Royal Society: London, U.K.

Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2006. Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. The Plant Cell 18, 1292–309.

Schluttenhofer C, Pattanail S, Patra B, Yuan L. Analyses of Catharanthus roseus and Arabidopsis thaliana WRKY transcription factors reveal involvement in jasmonate signaling. BMC Genomics. 2014;15:502

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

Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D. and Lohmann, J. U. (2003). Dissection of floral induction pathways using global expression analysis. Development 130,6001 -6012.

Schroeder, J.I., Kwak, J.M. and Allen, G.J. (2001) Guard cell abscisic acid signalling and engineering drought hardiness in plants. Nature, 410, 327–330.

Schruff, M.C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R.J. (2006). The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. Development 133, 251–261.

Schumann U, Wanner G, Veenhuis M, Schmid M, Gietl C (2003) AthPEX10, a nuclear gene essential for peroxisome and storage organelle formation during Arabidopsis embryogenesis. Proc Natl Acad Sci USA 100: 9626–9631

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, et al. 2002. Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J. 31:279–92

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y., & Shinozaki, K., (2001). Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell, 13, 61–72.

Seo, E., Lee, H., Jeon, J., Park, H., Kim, J., Noh, Y.-S., & Lee, I., (2009). Crosstalk between Cold Response and Flowering in Arabidopsis Is Mediated through the Flowering-Time Gene SOC1 and Its Upstream Negative Regulator FLC. Plant Cell, 21, 3185–3197.

Serraj, R. and Sinclair, T. R. 2002. Osmolyte accumulation: Can it really help increase in crop yield under drought conditions? Plant Cell Environ. 25: 333–341

Serraj, R. and Sinclair, T.R., 1996. Processes contributing to N2-fixation intensitivity to drought in the soybean cultivar Jackson. Crop Science, 36(4), pp.961-968.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A. 2002. A high-throughput Arabidopsis reverse genetics system. Plant Cell 14: 2985–2994.

Sgherri C. & Navari-Izzo F. (1995) Sunflower seedlings subjected to increasing water deficit stress: oxidative stress and defense mechanisms. Physiologia Plantarum 93, 25–30.

Sharma P, Dubey RS. Drought induces oxidative stress and enhances the activities of antioxidant enzymes in growing rice seedlings. Plant Grow Reg 2005; 46:209-221.

Shen, B., Jensen, R.G. & Bohnert, H.J. (1997). Mannitol protects against oxidation by hydroxyl radicals. Plant Physiol., 115, 527–532.

Sheveleva E, Chmara W, Bohnert HJ, Jensen RG: Increased salt and drought tolerance by D-ononitol production in transgenic Nicotiana tabacum. Plant Physiol 1997, 115:1211-1219.

Shi H, Wang Y, Chen Z, Ye T, Chan Z. 2012a. Analysis of natural variation in bermudagrass (Cynodon dactylon) reveals physiological responses underlying drought tolerance. PLoS One 7, e53422.

Shinozaki K, Yamaguchi-Shinozaki K. 2007. Gene networks involved in drought stress response and tolerance. Journal of Experimental Botany 58: 221–227.

Siefers N, Dang KK, Kumimoto RW, Bynum WE, Tayrose G, et al. (2009) Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. Plant Physiol 149: 625–641.

Simova-Stoilova L., Vassileva V., Petrova T., Tsenov N., Demirevska K., Feller U. (2006): Proteolytic activity in wheat leaves after drought stress and recovery. Gen. Appl. Plant Physiol., Special Issue: 91–100.

Singh D., Laxmi A. (2015). Transcriptional regulation of drought response: a tortuous network of transcriptional factors. Front. Plant Sci. 6:895.

Skirycz, A., Vandenbroucke, K., Clauw, P., Maleux, K., De Meyer, B., Dhondt, S., Pucci, A., Gonzalez, N., Hoeberichts, F., Tognetti, V.B. and Galbiati, M., 2011. Survival and growth of Arabidopsis plants given limited water are not equal. Nature biotechnology, 29(3), pp.212-214.

Skriver, K., and Mundy, J. (1990). Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2, 503-512.

Song CP, Agarwal M, Ohta M, Guo Y, Halfter U, et al. 2005. Role of an Arabidopsis AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. Plant Cell 17:2384–96

Spartz A.K., Lee S.H., Wenger J.P., Gonzalez N., Itoh H., Inzé D., Peer W.A., Murphy A.S., Overvoorde P.J., Gray W.M. (2012). The SAUR19 subfamily of SMALL AUXIN UP RNA genes promote cell expansion. Plant J. 70: 978–990.

Steer, A. 2010. From the Pump Room to the Board Room: Water's Central Role in Climate Change Adaptation. Washington DC, The World Bank.

Stegle, O., Denby, K. J., Cooke, E. J., Wild, D. L., Ghahramani, Z. & Borgwardt, K. M. 2010 A robust Bayesian two-sample test for detecting intervals of differential gene expression in microarray time series. J. Comput. Biol. 17, 355–367.

Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers and David J. Lipman (1990) Journal of Molecular Biology 215:403-410

Stogios, P.J. and Privé, G.G. (2004) The BACK domain in BTB-kelch proteins. Trends Biochem. Sci. 29, 634–637

Su, Z., Ma, X., Guo, H., Sukiran, N.L., Guo, B., Assmann, S.M. and Ma, H., 2013. Flower development under drought stress: morphological and transcriptomic analyses reveal acute responses and long-term acclimation in Arabidopsis. The Plant Cell, 25(10), pp.3785-3807.

Sun, W., Bernard, C., van de Cotte, Van Montagu, M., and Verbruggen, N. 2001. At-HSP17. 6A, encoding a small heat-shock protein in Arabidopsis, can enhance osmotolerance upon overexpression. Plant J. 27: 407–415.

Sunkar, R., Bartels, D., and Kirch, H.-H. 2003. Overexpression of a stress-inducible aldehyde dehydrogenase gene from Arabidopsis thaliana in transgenic plants improves stress tolerance. Plant J. 35: 452–464.

Sunnerhagen M, Pursglove S, Fladvad M (2002) The new MATH: homology suggests shared binding surfaces in meprin tetramers and TRAF trimers. FEBS Lett 530: 1–3

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M.,Kobayashi, M.,Yamaguchi-Shinozaki, K., and Shinozaki, K. 2002. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. Plant J. 29: 417–426.

Tamás, M. J., M. Rep, J. M. Thevelein, and S. Hohmann. 2000. Stimulation of the yeast high osmolarity glycerol (HOG) pathway: evidence for a signal generated by a change in turgor rather than by water stress. FEBS Lett. 472:159-165.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol, 30(12), pp.2725-2729.

Tanz S.K., Castleden I., Hooper C.M., Vacher M., Small I., Millar H.A. (2013). SUBA3: A database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. Nucleic Acids Res. 41: D1185–D1191.

Thalamuthu A, Mukhopadhyay I, Zheng X, Tseng GC: Evaluation and comparison of gene clustering methods in microarray analysis. Bioinformatics 2006, 22(19):2405–2412.

Tsugane K, Kobayashi K, Niwa Y, Ohba Y, Wada K, Kobayashi H. 1999. A recessive Arabidopsis mutant that grows enhanced active oxygen detoxification. Plant Cell. 11:1195–206

United Nations, Department of Economic and Social Affairs, Population Division (2009). World Population Prospects: The 2008 Revision, Highlights, Working Paper No. ESA/P/WP.210

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K., and Yamaguchi-Shinozaki, K. 2000. Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. Proc. Natl. Acad. Sci. USA 97: 11632– 11637.

Urao, T., Yamaguchi-Shinozaki, K., Urao, S., and Shinozaki, K. 1993. An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. Plant Cell 5: 1529–1539.

Valliyodan, B. & Nguyen, H.T. (2006). Understanding regulatory networks and engineering for enhanced drought tolerance in plants. Curr. Opin. Plant Biol., 9, 189–195.

Verbruggen N, Hermans C. Proline accumulation in plants: a review. Amino Acids 2008; 35: 753-759

Verslues, P.E., Agarwal, M., Katiyar-Agarwal, S., Zhu, J. and Zhu, J.K., 2006. Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. The Plant Journal, 45(4), pp.523-539.

Vierstra R.D. (2011). The RPT2 subunit of the 26S proteasome directs complex assembly, histone dynamics, and gametophyte and sporophyte development in Arabidopsis. Plant Cell 23: 4298–4317.

Wang L., Hua D., He J., Duan Y., Chen Z., Hong X., Gong Z. (2011). Auxin Response Factor2 (ARF2) and its regulated homeodomain gene HB33 mediate abscisic acid response in Arabidopsis. PLoS Genet. 7: e1002172

Wang Z, Huang B, Bonos SA, Meyer WA. Abscisic acid accumulation in relation to drought tolerance in Kentucky bluegrass. HortScience 2004;39:1133-1137.

Wang, H.-Y., Klatte, M., Jakoby, M., Bäumlein, H., Weisshaar, B., & Bauer, P., (2007). Iron deficiency-mediated stress regulation of four subgroup Ib BHLH genes in Arabidopsis thaliana. Planta, 226, 897–908.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci. 9: 244–252.

Wass MN, Barton G, Sternberg MJ (2012) CombFunc: predicting protein function using heterogeneous data sources. Nucleic Acids Res 40: W466–470.

Wass, M.N., Kelley, L.A. & Sternberg, M.J. 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Res. 38, W469–W473 (2010).

Weber, H., and Hellmann, H. (2009). Arabidopsis thaliana BTB/POZ-MATH proteins interact with members of the ERF/AP2 transcription factor family. FEBS J. 276:6624–6635.

Wigge, P.A., Kim, M.C., Jaeger, K.E., Busch, W., Schmid, M., Lohmann, J.U., & Weigel, D., (2005). Integration of Spatial and Temporal Information During Floral Induction in Arabidopsis. Science, 309, 1056–1059.

Williams L, Carles CC, Osmont KS, Fletcher JC. 2005. A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. Proc Natl Acad Sci USA 102:9703–9708.

Windram, O., Madhou, P., McHattie, S., Hill, C., Hickman, R., Cooke, E., Jenkins, D.J., Penfold, C.A., Baxter, L., & Breeze, E., (2012). Arabidopsis defense against Botrytis cinerea: chronology and regulation deciphered by high-resolution temporal transcriptomic analysis. Plant Cell, 24, 3530–3557.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. and Provart, N. (2007) An ‘electronic fluorescent pictograph’ browser for exploring and analyzing large-scale biological data sets. PLoS ONE, 2, e718.

Wu, L., Zhang, Z., Zhang, H., Wang, X.-C., and Huang, R. (2008). Transcriptional modulation of ethylene response factor protein JERF3 in the oxidative stress response enhances tolerance of tobacco seedlings to salt, drought, and freezing. Plant Physiol. 148: 1953–1963.

WWAP (United Nations World Water Assessment Programme). 2015. The United Nations World Water Development Report 2015: Water for a Sustainable World. Paris, UNESCO.

Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, et al. (2005) Expression of Arabidopsis MIRNA genes. Plant Physiol 138: 2145–2154.

Xu DP, Duan X, Wang B, Hong B, Ho TD, Wu R.1996. Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiology 110,249–257.

Yamaguchi M., Matoba K., Sawada R., Fujioka Y., Nakatogawa H., Yamamoto H., Kobashigawa Y., Hoshida H., Akada R., Ohsumi Y., Noda N.N., Inagaki F. (2012). Noncanonical recognition and UBL loading of distinct E2s by autophagy-essential Atg7. Nat. Struct. Mol. Biol. 19: 1250–1256.

Yamaguchi-Shinozaki, K. and Shinozaki, K. 1994. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low temperature, or high-salt stress. Plant Cell 6: 251–264.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (1993) Arabidopsis DNA encoding two desiccation-responsive rd29 genes. Plant Physiol. 101, 1119-1120.

Yamaguchi-Shinozaki, K., Koizumi, M., Urao S., and Shinozaki, K. (1992). Molecular cloning and characterization of 9 cDNAs for genes that are responsive to desiccation in Arabidopsis thaliana: Sequence analysis of one cDNA clone that encodes a putative transmembrane channel protein. Plant Cell Physiol. 33, 217-224.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 13:134.

Yona G, Dirks W, Rahman S, Lin DM: Effective similarity measures for expression profiles. Bioinformatics 2006, 22(13):1616–1622.

Yoshiba, Y., Kiyosue, T., Katagiri, T., Ueda, H., Mizoguchi, T., Yamaguchi-Shinozaki, K., Wada, K., Harada, Y. and Shinozaki, K. (1995) Plant J. 7:751-760.

Yoshida T., Fujita Y., Sayama H., Kidokoro S., Maruyama K., Mizoi J., Yamaguchi-Shinozaki K. (2010) AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. The Plant Journal 61, 672–685.

Yuan, J.S., Galbraith, D.W., Dai, S.Y., Griffin, P. & Stewart, C.N. (2008). Plant systems biology comes of age. Trends Plant Sci., 13, 165–171.

Yuan, Y., Wu, H., Wang, N., Li, J., Zhao, W., Du, J., Wang, D., & Ling, H.Q., (2008). FIT interacts with AtbHLH38 and AtbHLH39 in regulating iron uptake gene expression for iron homeostasis in Arabidopsis. Cell Res., 18, 385–397.

Zhang, J., Nguyen, H.T. and Blum, A. (1999) Genetic analysis of osmotic adjustment in crop plants. J. Exp. Bot. 50, 292–302.

Zhang, Y., Zhu, H., Zhang, Q., Li, M., Yan, M., Wang, R., Wang, L., Welti, R., Zhang, W. & Wang, X. (2009). Phospholipase Da1 and phosphatidic acid regulate NADPH oxidase activity and production of reactive oxygen species in ABA-mediated stomatal closure in Arabidopsis. Plant Cell, 21, 2357–2377.

Zhang, Z.J., Li, F., Li, D., Zhang, H., and Huang, R. (2010). Expression of ethylene response factor JERF1 in rice improves tolerance to drought. Planta 232, 765–774.

Zhu, B. C., Su, J., Chan, M. C., Verma, D. P. S., Fan, Y. L., and Wu, R. 1998. Over-expression of a -pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water-stress and salt-stress in transgenic rice. Plant Sci. 139: 41–48.

Zollman, S., Godt, D., Prive, G. G., Couderc, J.-L. & Laski, F. A. (1994) The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila Proc. NatL. Acad. Sci. USA 91, 10717-10721.

Appendix A

The list of 2190 differentially expressed genes selected for further analysis. The locus identifier and gene name for each gene is given in the table below, along with the cluster to which each gene grouped into by TCAP. The 'Networks' column indicates which of the six hub genes was modelled as regulating its expression: *FD*, P (*POZ*), U (*UKTF*), B (*BHLH038*), R (*RAP2.12*) and AN (*ANL2*).

Cluster No.	Locus Identifier	Gene Name	Networks
1	AT1G22370	UDP-glucosyl transferase 85A5	FD, P
1	AT1G30220	inositol transporter 2	FD, P
1	AT1G52180	Aquaporin-like superfamily protein	FD, P
1	AT1G54340	isocitrate dehydrogenase	FD, P
1	AT1G55680	Transducin/WD40 repeat-like superfamily protein	FD, P
1	AT1G55760	BTB/POZ domain-containing protein	FD, P
1	AT1G60600	UbiA prenyltransferase family protein	FD, P
1	AT1G61670	Lung seven transmembrane receptor family protein	FD, P
1	AT1G70360	F-box family protein	FD, P
1	AT1G71720	Nucleic acid-binding proteins superfamily	FD
1	AT1G71980	Protease-associated (PA) RING/U-box zinc finger family protein	FD, P
1	AT1G73110	P-loop containing nucleoside triphosphate hydrolases superfamily protein	FD, P
1	AT2G20140	AAA-type ATPase family protein	FD, P
1	AT2G25940	alpha-vacuolar processing enzyme	FD, P
1	AT2G26350	peroxin 10	FD, P
1	AT2G35060	K uptake permease 11	FD, P
1	AT2G47970	Nuclear pore localization protein NPL4	FD, P
1	AT3G01100	hypothetical protein 1	FD, P
1	AT3G03300	dicer-like 2	FD, P
1	AT3G06170	Serine-domain containing serine and sphingolipid biosynthesis protein	FD, P
1	AT3G07930	DNA glycosylase superfamily protein	FD, P
1	AT3G09280	transmembrane protein	FD, P
1	AT3G12350	F-box family protein	FD, P
1	AT3G12380	actin-related protein 5	FD, P
1	AT3G15730	phospholipase D alpha 1	FD, P
1	AT3G17000	ubiquitin-conjugating enzyme 32	FD, P
1	AT3G18050	GPI-anchored protein	FD, P
1	AT3G20250	pumilio 5	FD, P
1	AT3G24740	cellulose synthase, putative (DUF1644)	FD, P
1	AT3G27060	Ferritin/ribonucleotide reductase-like family protein	FD, P
1	AT3G47680	DNA binding protein	FD, P
1	AT4G00120	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	FD, P
1	AT4G08540	DNA-directed RNA polymerase II protein	FD, P
1	AT4G08980	F-BOX WITH WD-40 2	FD, P
1	AT4G10050	esterase/lipase/thioesterase family protein	FD, P
1	AT4G15540	EamA-like transporter family	FD, P
1	AT4G18120	MEI2-like 3 (ML3)	FD, P

1	AT4G22740	glycine-rich protein	FD, P
1	AT4G22930	pyrimidin 4	FD, P
1	AT4G30960	SOS3-interacting protein 3	FD, P
1	AT4G31860	Protein phosphatase 2C family protein	FD, P
1	AT4G32940	gamma vacuolar processing enzyme	FD, P
1	AT4G37800	xyloglucan endotransglucosylase/hydrolase 7	FD, P
1	AT5G03495	RNA-binding (RRM/RBD/RNP motifs) family protein	FD, P
1	AT5G13800	pheophytinase	FD, P
1	AT5G18130	transmembrane protein	FD, P
1	AT5G22510	alkaline/neutral invertase	FD, P
1	AT5G39570	transmembrane protein	FD, P
1	AT5G39860	basic helix-loop-helix (bHLH) DNA-binding family protein	FD, P
1	AT5G41370	DNA repair helicase XPB1-like protein	FD, P
1	AT5G42240	serine carboxypeptidase-like 42	FD, P
1	AT5G54080	homogentisate 1,2-dioxygenase	FD, P
1	AT5G61500	autophagy 3 (APG3)	FD, P
1	AT5G61510	GroES-like zinc-binding alcohol dehydrogenase family protein	FD, P
1	AT5G63370	Protein kinase superfamily protein	FD, P
2	AT1G07370	proliferating cellular nuclear antigen 1	
2	AT1G09200	Histone superfamily protein	
2	AT1G09750	Eukaryotic aspartyl protease family protein	U
2	AT1G10940	Protein kinase superfamily protein	
2	AT1G13960	WRKY DNA-binding protein 4	U
2	AT1G22690	Gibberellin-regulated family protein	
2	AT1G29070	Ribosomal protein L34	U
2	AT1G29660	GDSL-like Lipase/Acylhydrolase superfamily protein	U
2	AT1G51200	A20/AN1-like zinc finger family protein	
2	AT1G55490	chaperonin 60 beta	
2	AT1G64510	Translation elongation factor EF1B/ribosomal protein S6 family protein	U
2	AT1G68890	2-oxoglutarate decarboxylase/hydro-lyase/magnesium ion-binding protein	
2	AT1G78960	lupeol synthase 2	
2	AT1G80050	adenine phosphoribosyl transferase 2	U
2	AT2G07718	Cytochrome b/b6 protein	
2	AT2G21790	ribonucleotide reductase 1	
2	AT2G28950	expansin A6	U
2	AT2G31730	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	
2	AT3G19490	sodium:hydrogen antiporter 1	
2	AT3G20670	histone H2A 13	
2	AT3G26280	cytochrome P450, family 71, subfamily B, polypeptide 4	
2	AT3G27160	Ribosomal protein S21 family protein	U
2	AT3G27820	monodehydroascorbate reductase 4	
2	AT3G29160	SNF1 kinase homolog 11	
2	AT3G50840	Phototropic-responsive NPH3 family protein	
2	AT3G54560	histone H2A 11	
2	AT3G55330	PsbP-like protein 1	U
2	AT3G60590	cytochrome P450 family protein	

2	AT4G01310	Ribosomal L5P family protein	U
2	AT4G08290	nodulin MtN21 /EamA-like transporter family protein	
2	AT4G09650	F-type H -transporting ATPase subunit delta	
2	AT4G11100	gelsolin protein	U
2	AT4G16980	arabinogalactan-protein family	U
2	AT4G18480	P-loop containing nucleoside triphosphate hydrolases superfamily protein	U
2	AT4G29190	Zinc finger C-x8-C-x5-C-x3-H type family protein	U
2	AT4G34260	1,2-alpha-L-fucosidase	U
2	AT5G03520	RAB GTPase homolog 8C	
2	AT5G10390	Histone superfamily protein	U
2	AT5G14740	carbonic anhydrase 2	U
2	AT5G15230	GAST1 protein homolog 4	
2	AT5G22340	NF-kappa-B inhibitor-like protein	U
2	AT5G22880	histone B2	
2	AT5G23420	high-mobility group box 6	
2	AT5G38430	Ribulose bisphosphate carboxylase (small chain) family protein	U
2	AT5G40440	mitogen-activated protein kinase kinase 3	
2	AT5G43780	Pseudouridine synthase/archaeosine transglycosylase-like family protein	
2	AT5G45680	FK506-binding protein 13	U
2	AT5G45930	magnesium chelatase i2	U
2	AT5G48220	Aldolase-type TIM barrel family protein	U
2	AT5G59870	histone H2A 6	U
2	AT5G63310	nucleoside diphosphate kinase 2	U
2	AT5G63850	amino acid permease 4	
2	AT5G66120	3-dehydroquinate synthase	
2	AT5G66530	Galactose mutarotase-like superfamily protein	U
3	AT1G21400	Thiamin diphosphate-binding fold (THDP-binding) superfamily protein	
3	AT1G29395	COLD REGULATED 314 INNER MEMBRANE 1	
3	AT1G30500	nuclear factor Y, subunit A7	
3	AT1G67300	Major facilitator superfamily protein	
3	AT1G71340	PLC-like phosphodiesterases superfamily protein	
3	AT1G72190	D-isomer specific 2-hydroxyacid dehydrogenase family protein	
3	AT1G78070	Transducin/WD40 repeat-like superfamily protein	
3	AT1G79270	evolutionarily conserved C-terminal region 8	
3	AT2G04240	RING/U-box superfamily protein	
3	AT2G12400	plasma membrane fusion protein	
3	AT2G21970	stress enhanced protein 2	
3	AT2G47890	B-box type zinc finger protein with CCT domain-containing protein	
3	AT3G02910	AIG2-like (avirulence induced gene) family protein	
3	AT3G09580	FAD/NAD(P)-binding oxidoreductase family protein	
3	AT3G13672	TRAF-like superfamily protein	
3	AT3G15350	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein	
3	AT3G24520	heat shock transcription factor C1	
3	AT3G25870	hypothetical protein	
3	AT3G28007	Nodulin MtN3 family protein	

3	AT3G46450	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein
3	AT3G51670	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein
3	AT3G57540	Remorin family protein
3	AT4G01120	G-box binding factor 2
3	AT4G18620	PYR1-like 13
3	AT4G21926	hypothetical protein
3	AT4G23890	NAD(P)H-quinone oxidoreductase subunit S
3	AT4G36840	Galactose oxidase/kelch repeat superfamily protein
3	AT4G38730	magnesium transporter, putative (DUF803)
3	AT5G03210	E3 ubiquitin-protein ligase
3	AT5G04940	SU(VAR)3-9 homolog 1
3	AT5G07920	diacylglycerol kinase1
3	AT5G10540	Zincin-like metalloproteases family protein
3	AT5G12840	nuclear factor Y, subunit A1
3	AT5G13710	sterol methyltransferase 1
3	AT5G15190	hypothetical protein
3	AT5G15500	Ankyrin repeat family protein
3	AT5G17460	glutamyl-tRNA (Gln) amidotransferase subunit C
3	AT5G22290	NAC domain containing protein 89
3	AT5G34780	Thiamin diphosphate-binding fold (THDP-binding) superfamily protein
3	AT5G36670	RING/FYVE/PHD zinc finger superfamily protein
3	AT5G40800	hypothetical protein
4	AT1G04220	3-ketoacyl-CoA synthase 2
4	AT1G09795	ATP phosphoribosyl transferase 2
4	AT1G11210	cotton fiber protein, putative (DUF761)
4	AT1G17280	ubiquitin-conjugating enzyme 34
4	AT1G18710	myb domain protein 47
4	AT1G19550	Glutathione S-transferase family protein
4	AT1G21760	F-box protein 7
4	AT1G22640	myb domain protein 3
4	AT1G27420	Galactose oxidase/kelch repeat superfamily protein
4	AT1G28570	SGNH hydrolase-type esterase superfamily protein
4	AT1G36070	Transducin/WD40 repeat-like superfamily protein
4	AT1G53780	26S proteasome regulatory complex ATPase
4	AT1G71240	chromosome-partitioning protein, putative (DUF639)
4	AT1G75388	
4	AT1G76580	Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein
4	AT1G76590	PLATZ transcription factor family protein
4	AT1G76720	eukaryotic translation initiation factor 2 (eIF-2) family protein
4	AT2G16485	GW repeat- and PHD finger-containing protein NERD
4	AT2G20400	myb-like HTH transcriptional regulator family protein
4	AT2G22910	N-acetyl-l-glutamate synthase 1
4	AT2G31280	transcription factor bHLH155-like protein
4	AT2G32340	TraB family protein
4	AT2G38050	3-oxo-5-alpha-steroid 4-dehydrogenase family protein

4	AT2G41210	phosphatidylinositol- 4-phosphate 5-kinase 5
4	AT2G47670	Plant invertase/pectin methylesterase inhibitor superfamily protein
4	AT2G47780	Rubber elongation factor protein (REF)
4	AT3G22270	Topoisomerase II-associated protein PAT1
4	AT3G22380	time for coffee
4	AT3G26580	Tetratricopeptide repeat (TPR)-like superfamily protein
4	AT3G48710	DEK domain-containing chromatin associated protein
4	AT3G49940	LOB domain-containing protein 38
4	AT3G56650	thylakoid lumenal protein (Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein)
4	AT4G00380	XH/XS domain-containing protein
4	AT4G18530	lysine ketoglutarate reductase trans-splicing-like protein, putative (DUF707)
4	AT4G21910	MATE efflux family protein
4	AT4G22753	
4	AT4G22770	AT hook motif DNA-binding family protein
4	AT4G25020	D111/G-patch domain-containing protein
4	AT4G25080	magnesium-protoporphyrin IX methyltransferase
4	AT4G36900	related to AP2 10
4	AT4G36930	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
4	AT4G38530	phospholipase C1
4	AT4G38860	SAUR-like auxin-responsive protein family
4	AT5G06300	Putative lysine decarboxylase family protein
4	AT5G14960	DP-E2F-like 2
4	AT5G16610	hypothetical protein
4	AT5G19350	RNA-binding (RRM/RBD/RNP motifs) family protein
4	AT5G25220	homeobox protein knotted-1-like 3
4	AT5G27940	WPP domain protein 3
4	AT5G58020	RTF2 RING-finger protein
4	AT5G59010	kinase with tetratricopeptide repeat domain-containing protein
5	AT1G03410	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
5	AT1G09580	emp24/gp25L/p24 family/GOLD family protein
5	AT1G17940	Endosomal targeting BRO1-like domain-containing protein
5	AT1G29050	TRICHOME BIREFRINGENCE-LIKE 38
5	AT1G29540	LOW protein: protein BOBBER-like protein
5	AT1G36160	acetyl-CoA carboxylase 1
5	AT1G50950	protein disulfide-isomerase 5-like protein (DUF1692)
5	AT1G54160	nuclear factor Y, subunit A5
5	AT1G77290	Glutathione S-transferase family protein
5	AT2G19830	SNF7 family protein
5	AT2G21590	Glucose-1-phosphate adenylyltransferase family protein
5	AT2G22270	hematological/neurological-like protein
5	AT2G34850	NAD(P)-binding Rossmann-fold superfamily protein
5	AT2G43580	Chitinase family protein
5	AT2G45600	alpha/beta-Hydrolases superfamily protein
5	AT3G03270	Adenine nucleotide alpha hydrolases-like superfamily protein
5	AT3G12775	ubiquitin-conjugating enzyme family protein

5	AT3G23910	reverse transcriptase-like protein
5	AT3G53980	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
5	AT3G57680	Peptidase S41 family protein
5	AT4G09820	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
5	AT4G12080	AT-hook motif nuclear-localized protein 1
5	AT4G14520	DNA-directed RNA polymerase II-like protein
5	AT4G33550	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
5	AT4G34860	Plant neutral invertase family protein
5	AT5G05870	UDP-glucosyl transferase 76C1
5	AT5G08590	SNF1-related protein kinase 2.1
5	AT5G23530	carboxyesterase 18
5	AT5G23750	Remorin family protein
5	AT5G53710	hypothetical protein
5	AT5G56520	hypothetical protein
5	AT5G59700	Protein kinase superfamily protein
5	AT5G60910	AGAMOUS-like 8
5	AT5G66460	Glycosyl hydrolase superfamily protein
5	AT5G67090	Subtilisin-like serine endopeptidase family protein
6	AT1G03550	Secretory carrier membrane protein (SCAMP) family protein
6	AT1G04390	BTB/POZ domain-containing protein
6	AT1G17680	tetratricopeptide repeat (TPR)-containing protein
6	AT1G24822	hypothetical protein
6	AT1G67310	Calmodulin-binding transcription activator protein with CG-1 and Ankyrin domain
6	AT1G67580	Protein kinase superfamily protein
6	AT1G69540	AGAMOUS-like 94
6	AT1G78270	UDP-glucosyl transferase 85A4
6	AT1G78820	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain-containing protein
6	AT1G80070	Pre-mRNA-processing-splicing factor
6	AT2G02570	nucleic acid binding/RNA binding protein
6	AT2G13100	Major facilitator superfamily protein
6	AT2G23980	cyclic nucleotide-gated channel 6
6	AT2G31960	glucan synthase-like 3
6	AT2G32320	tRNAHis guanylyltransferase
6	AT2G34660	multidrug resistance-associated protein 2
6	AT2G37340	arginine/serine-rich zinc knuckle-containing protein 33
6	AT3G05510	Phospholipid/glycerol acyltransferase family protein
6	AT3G08850	Regulatory-associated protein of TOR 1
6	AT3G12090	tetraspanin6
6	AT3G27540	beta-1,4-N-acetylglucosaminyltransferase family protein
6	AT3G27670	ARM repeat superfamily protein
6	AT3G43520	Transmembrane proteins 14C
6	AT3G50430	golgin
6	AT4G15510	Photosystem II reaction center PsbP family protein
6	AT4G15900	pleiotropic regulatory locus 1
6	AT4G24470	GATA-type zinc finger protein with TIFY domain-containing protein

6	AT4G26690	PLC-like phosphodiesterase family protein
6	AT4G31650	Transcriptional factor B3 family protein
6	AT4G39550	Galactose oxidase/kelch repeat superfamily protein
6	AT5G05580	fatty acid desaturase 8
6	AT5G10370	helicase domain-containing protein / IBR domain-containing protein / zinc finger protein-like protein
6	AT5G12130	integral membrane TerC family protein
6	AT5G38670	Galactose oxidase/kelch repeat superfamily protein
6	AT5G42370	Calcineurin-like metallo-phosphoesterase superfamily protein
6	AT5G47590	Heat shock protein HSP20/alpha crystallin family
6	AT5G47840	adenosine monophosphate kinase
6	AT5G52730	Copper transport protein family
6	AT5G53350	CLP protease regulatory subunit X
7	AT1G01620	plasma membrane intrinsic protein 1C
7	AT1G04690	potassium channel beta subunit 1
7	AT1G07320	ribosomal protein L4
7	AT1G12900	glyceraldehyde 3-phosphate dehydrogenase A subunit 2
7	AT1G14290	sphingoid base hydroxylase 2
7	AT1G19450	Major facilitator superfamily protein
7	AT1G21500	hypothetical protein
7	AT1G28100	hypothetical protein
7	AT1G52240	rop guanine nucleotide exchange factor-like protein
7	AT2G04420	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
7	AT2G05710	aconitase 3
7	AT2G16720	myb domain protein 7
7	AT2G25810	tonoplast intrinsic protein 4;1
7	AT2G34620	Mitochondrial transcription termination factor family protein
7	AT2G41220	glutamate synthase 2
7	AT3G03830	SAUR-like auxin-responsive protein family
7	AT3G03850	SAUR-like auxin-responsive protein family
7	AT3G06750	hydroxyproline-rich glycoprotein family protein
7	AT3G08920	Rhodanese/Cell cycle control phosphatase superfamily protein
7	AT3G28690	Protein kinase superfamily protein
7	AT3G48410	alpha/beta-Hydrolases superfamily protein
7	AT3G54750	downstream neighbor of Son
7	AT3G60290	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
7	AT3G61630	cytokinin response factor 6
7	AT4G14090	UDP-Glycosyltransferase superfamily protein
7	AT4G23820	Pectin lyase-like superfamily protein
7	AT4G24780	Pectin lyase-like superfamily protein
7	AT4G25050	acyl carrier protein 4
7	AT4G33666	hypothetical protein
7	AT5G04760	Duplicated homeodomain-like superfamily protein
7	AT5G08020	RPA70-kDa subunit B
7	AT5G21930	P-type ATPase of Arabidopsis 2
7	AT5G27290	stress regulated protein

7	AT5G44530	Subtilase family protein	
7	AT5G47860	Gut esterase (DUF1350)	
7	AT5G62810	peroxin 14	
8	AT1G03020	Thioredoxin superfamily protein	B
8	AT1G03495	HXXXD-type acyl-transferase family protein	
8	AT1G09140	SERINE-ARGININE PROTEIN 30	B
8	AT1G09780	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	B, R
8	AT1G10070	branched-chain amino acid transaminase 2	
8	AT1G13609	Defensin-like (DEFL) family protein	B
8	AT1G18190	golgin Putative 2	B
8	AT1G34060	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein	B, R
8	AT1G47395	hypothetical protein	B
8	AT1G47400	hypothetical protein	B
8	AT1G53280	Class I glutamine amidotransferase-like superfamily protein	B
8	AT1G53910	related to AP2 12	B, R
8	AT1G63240	hypothetical protein	B, R
8	AT1G79030	Chaperone DnaJ-domain superfamily protein	B, R
8	AT1G80570	RNI-like superfamily protein	B, R
8	AT2G03590	ureide permease 1	B, R
8	AT2G38070	LOW protein: UPF0503-like protein, putative (DUF740)	B, R
8	AT2G40390	neuronal PAS domain protein	B, R
8	AT2G47790	Transducin/WD40 repeat-like superfamily protein	B
8	AT3G05120	alpha/beta-Hydrolases superfamily protein	B
8	AT3G18990	AP2/B3-like transcriptional factor family protein	B, R
8	AT3G21460	Glutaredoxin family protein	B
8	AT3G52180	dual specificity protein phosphatase (DsPTP1) family protein	
8	AT3G56970	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	B, R
8	AT3G62090	phytochrome interacting factor 3-like 2	B, R
8	AT3G63110	isopentenyltransferase 3	B
8	AT4G16150	calmodulin-binding transcription activator 5	B, R
8	AT4G23860	PHD finger protein-like protein	B
8	AT4G25970	phosphatidylserine decarboxylase 3	B, R
8	AT4G27120	DDRGK domain protein	B, R
8	AT4G36540	BR enhanced expression 2	B, R
8	AT4G39770	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein	
8	AT5G08100	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein	B, R
8	AT5G25110	CBL-interacting protein kinase 25	B, R
8	AT5G37370	PRP38 family protein	B, R
8	AT5G57565	Protein kinase superfamily protein	B, R
8	AT5G65010	asparagine synthetase 2	B
9	AT1G07890	ascorbate peroxidase 1	
9	AT1G14490	Putative AT-hook DNA-binding family protein	
9	AT1G20490	AMP-dependent synthetase and ligase family protein	
9	AT1G29390	cold regulated 314 thylakoid membrane 2	
9	AT1G29810	Transcriptional coactivator/pterin dehydratase	

9	AT1G62290	Saposin-like aspartyl protease family protein
9	AT1G62540	flavin-monooxygenase glucosinolate S-oxygenase 2
9	AT1G64740	alpha-1 tubulin
9	AT1G67510	Leucine-rich repeat protein kinase family protein
9	AT1G68220	aerobic coproporphyrinogen-III oxidase (DUF1218)
9	AT1G74810	HCO3- transporter family
9	AT2G01150	RING-H2 finger protein 2B
9	AT2G05185	hypothetical protein
9	AT2G16660	Major facilitator superfamily protein
9	AT2G21820	seed maturation protein
9	AT2G22190	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
9	AT2G22540	K-box region and MADS-box transcription factor family protein
9	AT2G29950	ELF4-like 1
9	AT2G37300	transmembrane protein
9	AT2G38660	Amino acid dehydrogenase family protein
9	AT2G45720	ARM repeat superfamily protein
9	AT3G24760	Galactose oxidase/kelch repeat superfamily protein
9	AT4G01060	CAPRICE-like MYB3
9	AT4G04750	Major facilitator superfamily protein
9	AT4G04760	Major facilitator superfamily protein
9	AT4G16750	Integrase-type DNA-binding superfamily protein
9	AT4G34588	
9	AT4G37980	cinnamyl alcohol dehydrogenase 7
9	AT4G40010	SNF1-related protein kinase 2.7
9	AT5G02560	histone H2A 12
9	AT5G17030	UDP-glucosyl transferase 78D3
9	AT5G25390	Integrase-type DNA-binding superfamily protein
9	AT5G49120	DUF581 family protein, putative (DUF581)
9	AT5G50110	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
9	AT5G59000	RING/FYVE/PHD zinc finger superfamily protein
10	AT1G01290	cofactor of nitrate reductase and xanthine dehydrogenase 3
10	AT1G07530	SCARECROW-like 14
10	AT1G13195	RING/U-box superfamily protein
10	AT1G27460	no pollen germination related 1
10	AT1G47240	NRAMP metal ion transporter 2
10	AT1G64360	hypothetical protein
10	AT1G66100	Plant thionin
10	AT1G69530	expansin A1
10	AT1G79230	mercaptopyruvate sulfurtransferase 1
10	AT2G34060	Peroxidase superfamily protein
10	AT2G35540	DNAJ heat shock N-terminal domain-containing protein
10	AT2G35550	basic pentacysteine 7
10	AT2G40610	expansin A8
10	AT2G43745	jacalin lectin-like protein
10	AT2G46080	BPS1-like protein
10	AT3G03520	non-specific phospholipase C3

10	AT3G12730	Homeodomain-like superfamily protein
10	AT3G16360	HPT phosphotransmitter 4
10	AT3G19990	E3 ubiquitin-protein ligase
10	AT3G26730	RING/U-box superfamily protein
10	AT3G58270	phospholipase-like protein (PEARLI 4) with TRAF-like domain protein
10	AT4G05530	indole-3-butyric acid response 1
10	AT4G12300	cytochrome P450, family 706, subfamily A, polypeptide 4
10	AT4G15093	catalytic LigB subunit of aromatic ring-opening dioxygenase family
10	AT4G26555	FKBP-like peptidyl-prolyl cis-trans isomerase family protein
10	AT4G27180	kinesin 2
10	AT4G36190	Serine carboxypeptidase S28 family protein
10	AT5G41320	stress response NST1-like protein
10	AT5G67160	HXXXD-type acyl-transferase family protein
11	AT1G09530	phytochrome interacting factor 3
11	AT1G17840	white-brown complex-like protein
11	AT1G48635	peroxin 3
11	AT1G56650	production of anthocyanin pigment 1
11	AT1G62710	beta vacuolar processing enzyme
11	AT1G66380	myb domain protein 114
11	AT1G66390	myb domain protein 90
11	AT1G69295	plasmodesmata callose-binding protein 4
11	AT1G69800	Cystathionine beta-synthase (CBS) protein
11	AT1G69820	gamma-glutamyl transpeptidase 3
11	AT1G73040	Mannose-binding lectin superfamily protein
11	AT1G76240	DUF241 domain protein (DUF241)
11	AT2G25625	histone deacetylase-like protein
11	AT2G34730	myosin heavy chain-like protein
11	AT2G38670	phosphorylethanolamine cytidylyltransferase 1
11	AT3G01090	SNF1 kinase homolog 10
11	AT3G09250	Nuclear transport factor 2 (NTF2) family protein
11	AT3G12580	heat shock protein 70
11	AT3G18610	nucleolin like 2
11	AT3G27220	Galactose oxidase/kelch repeat superfamily protein
11	AT3G54320	Integrase-type DNA-binding superfamily protein
11	AT3G61990	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
11	AT4G03110	RNA-binding protein-defense related 1
11	AT4G19390	Uncharacterized protein family (UPF0114)
11	AT4G33040	Thioredoxin superfamily protein
11	AT4G39210	Glucose-1-phosphate adenylyltransferase family protein
11	AT4G39640	gamma-glutamyl transpeptidase 1
11	AT5G04530	3-ketoacyl-CoA synthase 19
11	AT5G05980	DHFS-FPGS homolog B
11	AT5G06950	bZIP transcription factor family protein
11	AT5G16030	mental retardation GTPase activating protein
11	AT5G41870	Pectin lyase-like superfamily protein
11	AT5G45490	P-loop containing nucleoside triphosphate hydrolases superfamily

		protein
11	AT5G56160	Sec14p-like phosphatidylinositol transfer family protein
11	AT5G65730	xyloglucan endotransglucosylase/hydrolase 6
12	AT1G01250	Integrase-type DNA-binding superfamily protein
12	AT1G09932	Phosphoglycerate mutase family protein
12	AT1G13920	Remorin family protein
12	AT1G17290	alanine aminotransferas
12	AT1G18870	isochorismate synthase 2
12	AT1G19190	alpha/beta-Hydrolases superfamily protein
12	AT1G32520	TLDc domain protein
12	AT1G43620	UDP-Glycosyltransferase superfamily protein
12	AT1G45249	abscisic acid responsive elements-binding factor 2
12	AT1G53070	Legume lectin family protein
12	AT1G63420	O-glucosyltransferase-like protein (DUF821)
12	AT1G68870	SOB five-like 2
12	AT1G72920	Toll-Interleukin-Resistance (TIR) domain family protein
12	AT2G17890	calcium-dependent protein kinase 16
12	AT2G19860	hexokinase 2
12	AT2G26530	AR781, pheromone receptor-like protein (DUF1645)
12	AT2G35760	Uncharacterized protein family (UPF0497)
12	AT2G40520	Nucleotidyltransferase family protein
12	AT2G41000	Chaperone DnaJ-domain superfamily protein
12	AT2G44770	ELMO/CED-12 family protein
12	AT3G03690	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
12	AT3G10400	RNA recognition motif and CCHC-type zinc finger domains containing protein
12	AT3G17930	transmembrane protein
12	AT3G19340	aminopeptidase (DUF3754)
12	AT3G47590	alpha/beta-Hydrolases superfamily protein
12	AT3G48700	carboxyesterase 13
12	AT3G52920	transcriptional activator (DUF662)
12	AT3G54000	TIP41-like protein
12	AT4G02520	glutathione S-transferase PHI 2
12	AT4G10360	TRAM, LAG1 and CLN8 (TLC) lipid-sensing domain containing protein
12	AT4G18270	translocase 11
12	AT4G23630	VIRB2-interacting protein 1
12	AT4G25480	dehydration response element B1A
12	AT4G25600	Oxoglutarate/iron-dependent oxygenase
12	AT4G30470	NAD(P)-binding Rossmann-fold superfamily protein
12	AT4G39700	Heavy metal transport/detoxification superfamily protein
12	AT5G17010	Major facilitator superfamily protein
12	AT5G17210	transmembrane protein, putative (DUF1218)
12	AT5G44670	glycosyltransferase family protein (DUF23)
12	AT5G47550	Cystatin/monellin superfamily protein
13	AT1G22790	Low affinity potassium transport system protein
13	AT1G23130	Polyketide cyclase/dehydrase and lipid transport superfamily protein

13	AT1G34340	alpha/beta-Hydrolases superfamily protein
13	AT1G56230	enolase (DUF1399)
13	AT1G67860	transmembrane protein
13	AT2G17390	ankyrin repeat-containing 2B
13	AT2G21940	shikimate kinase 1
13	AT2G28840	hypothetical protein
13	AT2G39570	ACT domain-containing protein
13	AT3G11170	fatty acid desaturase 7
13	AT3G20120	cytochrome P450, family 705, subfamily A, polypeptide 21
13	AT3G21360	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
13	AT3G26300	cytochrome P450, family 71, subfamily B, polypeptide 34
13	AT3G54290	hemerythrin HHE cation-binding domain protein
13	AT3G56720	pre-mRNA-splicing factor
13	AT3G62730	desiccation-like protein
13	AT4G00270	DNA-binding storekeeper protein-related transcriptional regulator
13	AT4G33610	glycine-rich protein
13	AT5G10960	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
13	AT5G11150	vesicle-associated membrane protein 713
13	AT5G16550	voltage-dependent L-type calcium channel subunit
13	AT5G19430	RING/U-box superfamily protein
13	AT5G42780	homeobox protein 27
13	AT5G47720	Thiolase family protein
13	AT5G49990	Xanthine/uracil permease family protein
13	AT5G60800	Heavy metal transport/detoxification superfamily protein
13	AT5G62760	P-loop containing nucleoside triphosphate hydrolases superfamily protein
13	AT5G65760	Serine carboxypeptidase S28 family protein
14	AT1G27400	Ribosomal protein L22p/L17e family protein
14	AT1G70600	Ribosomal protein L18e/L15 superfamily protein
14	AT2G27530	Ribosomal protein L1p/L10e family
14	AT2G33370	Ribosomal protein L14p/L23e family protein
14	AT2G36620	ribosomal protein L24
14	AT2G37270	ribosomal protein 5B
14	AT2G39080	NAD(P)-binding Rossmann-fold superfamily protein
14	AT2G39460	ribosomal protein L23AA
14	AT2G44120	Ribosomal protein L30/L7 family protein
14	AT3G02080	Ribosomal protein S19e family protein
14	AT3G02560	Ribosomal protein S7e family protein
14	AT3G02720	Class I glutamine amidotransferase-like superfamily protein
14	AT3G05560	Ribosomal L22e protein family
14	AT3G07110	Ribosomal protein L13 family protein
14	AT3G08740	elongation factor P (EF-P) family protein
14	AT3G09500	Ribosomal L29 family protein
14	AT3G11940	ribosomal protein 5A
14	AT3G16780	Ribosomal protein L19e family protein
14	AT3G18130	receptor for activated C kinase 1C

14	AT3G28900	Ribosomal protein L34e superfamily protein	
14	AT3G51190	Ribosomal protein L2 family	
14	AT3G62530	ARM repeat superfamily protein	
14	AT4G10480	Nascent polypeptide-associated complex (NAC), alpha subunit family protein	
14	AT4G13720	Inosine triphosphate pyrophosphatase family protein	
14	AT4G15000	Ribosomal L27e protein family	
14	AT4G31700	ribosomal protein S6	
14	AT5G02450	Ribosomal protein L36e family protein	
14	AT5G07090	Ribosomal protein S4 (RPS4A) family protein	
14	AT5G16130	Ribosomal protein S7e family protein	
14	AT5G20290	Ribosomal protein S8e family protein	
14	AT5G23900	Ribosomal protein L13e family protein	
14	AT5G39740	ribosomal protein L5 B	
14	AT5G52650	RNA binding Plectin/S10 domain-containing protein	
14	AT5G53770	Nucleotidyltransferase family protein	
14	AT5G59850	Ribosomal protein S8 family protein	
14	AT5G60670	Ribosomal protein L11 family protein	
14	AT5G61170	Ribosomal protein S19e family protein	
15	AT1G15825	hydroxyproline-rich glycoprotein family protein	
15	AT1G24490	OxaA/YidC-like membrane insertion protein	
15	AT1G31790	Tetratricopeptide repeat (TPR)-like superfamily protein	
15	AT1G58290	Glutamyl-tRNA reductase family protein	
15	AT1G60550	enoyl-CoA hydratase/isomerase D	
15	AT1G67090	ribulose bisphosphate carboxylase small chain 1A	
15	AT1G75690	DnaJ/Hsp40 cysteine-rich domain superfamily protein	
15	AT2G34510	choice-of-anchor C domain protein, putative (Protein of unknown function, DUF642)	
15	AT2G35650	cellulose synthase like	
15	AT2G42470	TRAF-like family protein	
15	AT3G02870	Inositol monophosphatase family protein	
15	AT3G05380	ALWAYS EARLY 2	
15	AT3G50790	esterase/lipase/thioesterase family protein	
15	AT3G53190	Pectin lyase-like superfamily protein	
15	AT4G00400	glycerol-3-phosphate acyltransferase 8	
15	AT4G34220	Leucine-rich repeat protein kinase family protein	
15	AT5G11420	transmembrane protein, putative (Protein of unknown function, DUF642)	
15	AT5G38410	Ribulose bisphosphate carboxylase (small chain) family protein	
15	AT5G38860	BES1-interacting Myc-like protein 3	
15	AT5G41761	hypothetical protein	
15	AT5G64470	trichome birefringence-like protein (DUF828)	
15	AT5G67330	natural resistance associated macrophage protein 4	
16	AT1G09390	GDSL-like Lipase/Acylhydrolase superfamily protein	AN
16	AT1G75960	AMP-dependent synthetase and ligase family protein	
16	AT2G21200	SAUR-like auxin-responsive protein family	AN
16	AT2G32880	TRAF-like family protein	AN
16	AT2G37260	WRKY family transcription factor family protein	AN

16	AT3G03820	SAUR-like auxin-responsive protein family	AN
16	AT3G13403	Defensin-like (DEFL) family protein	AN
16	AT3G56680	Single-stranded nucleic acid binding R3H protein	AN
16	AT3G62930	Thioredoxin superfamily protein	AN
16	AT4G04330	Chaperonin-like RbcX protein	AN
16	AT4G22570	adenine phosphoribosyl transferase 3	AN
16	AT4G26980	RNI-like superfamily protein	AN
16	AT4G38840	SAUR-like auxin-responsive protein family	AN
16	AT4G38850	SAUR-like auxin-responsive protein family	AN
16	AT5G12040	Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase family protein	AN
16	AT5G18050	SAUR-like auxin-responsive protein family	AN
16	AT5G19190	hypothetical protein	AN
16	AT5G19470	nudix hydrolase homolog 24	AN
16	AT5G35120	MADS-box family protein	AN
16	AT5G44316	Stabilizer of iron transporter SufD superfamily protein	AN
16	AT5G53420	CCT motif family protein	AN
16	AT5G61412	hypothetical protein	AN
16	AT5G62650	Tic22-like family protein	AN
17	AT1G08230	Transmembrane amino acid transporter family protein	
17	AT1G31830	Amino acid permease family protein	
17	AT1G36150	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	
17	AT1G53730	STRUBBELIG-receptor family 6	
17	AT1G53785		
17	AT1G59610	dynamin-like 3	
17	AT1G80840	WRKY DNA-binding protein 40	
17	AT2G21620	Adenine nucleotide alpha hydrolases-like superfamily protein	
17	AT2G29140	pumilio 3	
17	AT2G45790	phosphomannomutase	
17	AT3G07730	hypothetical protein	
17	AT3G19290	ABRE binding factor 4	
17	AT3G21800	UDP-glucosyl transferase 71B8	
17	AT3G47540	Chitinase family protein	
17	AT3G54140	peptide transporter 1	
17	AT3G59210	F-box/RNI-like superfamily protein	
17	AT4G15660	Thioredoxin superfamily protein	
17	AT4G17780	F-box and associated interaction domains-containing protein	
17	AT4G23670	Polyketide cyclase/dehydrase and lipid transport superfamily protein	
17	AT4G27080	PDI-like 5-4	
17	AT4G29890	choline monooxygenase, putative (CMO-like)	
17	AT4G37200	Thioredoxin superfamily protein	
17	AT4G38350	Patched family protein	
17	AT5G05780	RP non-ATPase subunit 8A	
17	AT5G20840	Phosphoinositide phosphatase family protein	
17	AT5G22520	hypothetical protein	
17	AT5G45810	CBL-interacting protein kinase 19	

17	AT5G62000	auxin response factor 2
18	AT1G23750	Nucleic acid-binding, OB-fold-like protein
18	AT1G32990	plastid ribosomal protein l11
18	AT1G35680	Ribosomal protein L21
18	AT1G51550	Kelch repeat-containing F-box family protein
18	AT1G55280	Lipase/lipooxygenase, PLAT/LH2 family protein
18	AT1G62780	dimethylallyl, adenosine tRNA methylthiotransferase
18	AT1G68238	transmembrane protein
18	AT1G74230	glycine-rich RNA-binding protein 5
18	AT1G74970	ribosomal protein S9
18	AT1G79850	ribosomal protein S17
18	AT2G22430	homeobox protein 6
18	AT2G24490	replicon protein A2
18	AT2G29630	thiaminC
18	AT2G33450	Ribosomal L28 family
18	AT2G38140	plastid-specific ribosomal protein 4
18	AT2G43620	Chitinase family protein
18	AT3G02820	zinc knuckle (CCHC-type) family protein
18	AT3G03640	beta glucosidase 25
18	AT3G05545	RING/U-box superfamily protein
18	AT3G15190	chloroplast 30S ribosomal protein S20
18	AT3G53900	uracil phosphoribosyltransferase
18	AT3G54210	Ribosomal protein L17 family protein
18	AT4G13010	Oxidoreductase, zinc-binding dehydrogenase family protein
18	AT4G22800	
18	AT4G23710	vacuolar ATP synthase subunit G2
18	AT4G24930	thylakoid lumenal 17.9 kDa protein, chloroplast
18	AT5G01790	hypothetical protein
18	AT5G15350	early nodulin-like protein 17
18	AT5G21920	YGGT family protein
18	AT5G25630	Tetratricopeptide repeat (TPR)-like superfamily protein
18	AT5G27390	tagatose-6-phosphate ketose/aldose isomerase, putative (Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein)
18	AT5G48490	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
18	AT5G58250	YCF54
18	AT5G65360	Histone superfamily protein
18	AT5G67260	CYCLIN D3;2
19	AT1G07500	hypothetical protein
19	AT1G07720	3-ketoacyl-CoA synthase 3
19	AT1G12000	Phosphofructokinase family protein
19	AT1G28140	integral membrane family protein
19	AT1G44960	SNARE associated Golgi protein family
19	AT1G45150	alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
19	AT1G65660	Pre-mRNA splicing Prp18-interacting factor
19	AT1G74410	RING/U-box superfamily protein

19	AT1G78600	light-regulated zinc finger protein 1
19	AT2G40490	Uroporphyrinogen decarboxylase
19	AT2G47940	DEGP protease 2
19	AT3G01480	cyclophilin 38
19	AT3G05165	Major facilitator superfamily protein
19	AT3G10720	Plant invertase/pectin methylesterase inhibitor superfamily
19	AT3G28290	transmembrane protein, putative (DUF677)
19	AT3G48720	HXXXD-type acyl-transferase family protein
19	AT3G52800	A20/AN1-like zinc finger family protein
19	AT4G14020	Rapid alkalinization factor (RALF) family protein
19	AT4G25580	CAP160 protein
19	AT5G02790	Glutathione S-transferase family protein
19	AT5G05860	UDP-glucosyl transferase 76C2
19	AT5G08280	hydroxymethylbilane synthase
19	AT5G10930	CBL-interacting protein kinase 5
19	AT5G54840	Ras-related small GTP-binding family protein
19	AT5G60580	RING/U-box superfamily protein
20	AT1G07510	FTSH protease 10
20	AT1G17870	ethylene-dependent gravitropism-deficient and yellow-green-like 3
20	AT1G20450	Dehydrin family protein
20	AT1G28960	nudix hydrolase homolog 15
20	AT1G51620	Protein kinase superfamily protein
20	AT1G57590	Pectinacetylesterase family protein
20	AT1G64970	gamma-tocopherol methyltransferase
20	AT1G69260	ABI five binding protein
20	AT1G71140	MATE efflux family protein
20	AT1G73390	Endosomal targeting BRO1-like domain-containing protein
20	AT2G41190	Transmembrane amino acid transporter family protein
20	AT3G22370	alternative oxidase 1A
20	AT3G27250	hypothetical protein
20	AT3G29575	ABI five binding protein 3
20	AT3G61890	homeobox 12
20	AT4G29070	Phospholipase A2 family protein
20	AT4G29740	cytokinin oxidase 4
20	AT4G34000	abscisic acid responsive elements-binding factor 3
20	AT4G37320	cytochrome P450, family 81, subfamily D, polypeptide 5
20	AT4G37470	alpha/beta-Hydrolases superfamily protein
20	AT5G01260	Carbohydrate-binding-like fold
20	AT5G10410	ENTH/ANTH/VHS superfamily protein
20	AT5G47640	nuclear factor Y, subunit B2
20	AT5G49280	hydroxyproline-rich glycoprotein family protein
20	AT5G64230	1,8-cineole synthase
21	AT1G08390	recQ-mediated instability-like protein
21	AT1G15010	mediator of RNA polymerase II transcription subunit
21	AT1G26770	expansin A10
21	AT1G30520	acyl-activating enzyme 14

21	AT1G47380	Protein phosphatase 2C family protein
21	AT1G52880	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
21	AT1G55330	arabinogalactan protein 21
21	AT1G74070	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein
21	AT1G76160	SKU5 similar 5
21	AT2G19780	Leucine-rich repeat (LRR) family protein
21	AT2G28740	histone H4
21	AT2G35190	Putative plant snare 11
21	AT3G26380	Melibiase family protein
21	AT3G46320	Histone superfamily protein
21	AT3G52870	IQ calmodulin-binding motif family protein
21	AT3G54270	sucrose-6F-phosphate phosphohydrolase family protein
21	AT4G12380	hypothetical protein
21	AT4G15180	SET domain protein 2
21	AT5G04160	Nucleotide-sugar transporter family protein
21	AT5G06530	ABC-2 type transporter family protein
21	AT5G11790	N-MYC downregulated-like 2
21	AT5G25460	transmembrane protein, putative (Protein of unknown function, DUF642)
21	AT5G59970	Histone superfamily protein
21	AT5G64480	hypothetical protein
22	AT1G09030	nuclear factor Y, subunit B4
22	AT1G13390	translocase subunit seca
22	AT1G16310	Cation efflux family protein
22	AT1G43700	VIRE2-interacting protein 1
22	AT1G48480	receptor-like kinase 1
22	AT1G72610	germin-like protein 1
22	AT2G07696	Ribosomal protein S7p/S5e family protein
22	AT2G33270	atypical CYS HIS rich thioredoxin 3
22	AT2G36340	DNA-binding storekeeper protein-related transcriptional regulator
22	AT2G40500	Protein kinase superfamily protein
22	AT2G47750	putative indole-3-acetic acid-amido synthetase GH3.9
22	AT3G01550	phosphoenolpyruvate (pep)/phosphate translocator 2
22	AT3G06890	transmembrane protein
22	AT3G11630	Thioredoxin superfamily protein
22	AT3G14930	Uroporphyrinogen decarboxylase
22	AT3G21055	photosystem II subunit T
22	AT3G24570	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein
22	AT3G26690	nudix hydrolase homolog 13
22	AT3G28180	Cellulose-synthase-like C4
22	AT3G28580	P-loop containing nucleoside triphosphate hydrolases superfamily protein
22	AT3G51880	high mobility group B1
22	AT3G60440	Phosphoglycerate mutase family protein
22	AT3G62030	rotamase CYP 4
22	AT4G26810	SWIB/MDM2 domain superfamily protein
22	AT5G01015	transmembrane protein

22	AT5G01820	serine/threonine protein kinase 1
22	AT5G06290	2-cysteine peroxiredoxin B
22	AT5G27350	Major facilitator superfamily protein
22	AT5G47770	farnesyl diphosphate synthase 1
22	AT5G57700	BNR/Asp-box repeat family protein
22	AT5G59540	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
23	AT1G16010	magnesium transporter 2
23	AT1G22400	UDP-Glycosyltransferase superfamily protein
23	AT1G35550	elongation factor Tu C-terminal domain-containing protein
23	AT1G44750	purine permease 11
23	AT1G58270	TRAF-like family protein
23	AT1G62420	DUF506 family protein (DUF506)
23	AT1G66260	RNA-binding (RRM/RBD/RNP motifs) family protein
23	AT1G67340	HCP-like superfamily protein with MYND-type zinc finger
23	AT1G76670	Nucleotide-sugar transporter family protein
23	AT2G16580	SAUR-like auxin-responsive protein family
23	AT2G17650	AMP-dependent synthetase and ligase family protein
23	AT2G23780	RING/U-box superfamily protein
23	AT2G46470	inner membrane protein OXA1-like protein
23	AT3G20650	mRNA capping enzyme family protein
23	AT3G26085	CAAX amino terminal protease family protein
23	AT3G55310	NAD(P)-binding Rossmann-fold superfamily protein
23	AT3G56630	cytochrome P450, family 94, subfamily D, polypeptide 2
23	AT3G59350	Protein kinase superfamily protein
23	AT4G30460	glycine-rich protein
23	AT4G36890	Nucleotide-diphospho-sugar transferases superfamily protein
23	AT5G16760	Inositol 1,3,4-trisphosphate 5/6-kinase family protein
23	AT5G44585	hypothetical protein
23	AT5G60870	Regulator of chromosome condensation (RCC1) family protein
23	AT5G63830	HIT-type Zinc finger family protein
23	AT5G66420	TIM-barrel signal transduction protein
23	AT5G67360	Subtilase family protein
24	AT1G09415	NIM1-interacting 3
24	AT1G10650	SBP (S-ribonuclease binding protein) family protein
24	AT1G20650	Protein kinase superfamily protein
24	AT1G61930	senescence regulator (Protein of unknown function, DUF584)
24	AT1G73700	MATE efflux family protein
24	AT2G22900	Galactosyl transferase GMA12/MNN10 family protein
24	AT2G46330	arabinogalactan protein 16
24	AT3G16530	Legume lectin family protein
24	AT3G55760	hypothetical protein
24	AT4G17770	trehalose phosphatase/synthase 5
24	AT4G18700	CBL-interacting protein kinase 12
24	AT4G23210	cysteine-rich RLK (RECEPTOR-like protein kinase) 13
24	AT4G33490	Eukaryotic aspartyl protease family protein
24	AT4G36010	Pathogenesis-related thaumatin superfamily protein

24	AT5G06510	nuclear factor Y, subunit A10
24	AT5G39410	Saccharopine dehydrogenase
24	AT5G64550	loricrin-like protein
25	AT1G06475	transmembrane protein
25	AT1G22160	senescence-associated family protein (DUF581)
25	AT1G33590	Leucine-rich repeat (LRR) family protein
25	AT1G33600	Leucine-rich repeat (LRR) family protein
25	AT1G35580	cytosolic invertase 1
25	AT1G65590	beta-hexosaminidase 3
25	AT1G72030	Acyl-CoA N-acyltransferases (NAT) superfamily protein
25	AT1G74910	ADP-glucose pyrophosphorylase family protein
25	AT2G14860	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein
25	AT2G25900	Zinc finger C-x8-C-x5-C-x3-H type family protein
25	AT2G36310	uridine-ribohydrolase 1
25	AT3G06080	trichome birefringence-like protein (DUF828)
25	AT3G10330	Cyclin-like family protein
25	AT3G12150	alpha/beta hydrolase family protein
25	AT3G15650	alpha/beta-Hydrolases superfamily protein
25	AT3G23630	isopentenyltransferase 7
25	AT3G25730	ethylene response DNA binding factor 3
25	AT3G27570	Sucrase/ferredoxin-like family protein
25	AT3G54680	proteophosphoglycan-like protein
25	AT3G55120	Chalcone-flavanone isomerase family protein
25	AT4G05010	F-box family protein
25	AT4G33905	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein
25	AT4G37295	hypothetical protein
25	AT5G47060	hypothetical protein (DUF581)
25	AT5G49730	ferric reduction oxidase 6
25	AT5G65300	hypothetical protein
26	AT1G08030	tyrosylprotein sulfotransferase
26	AT1G20340	Cupredoxin superfamily protein
26	AT1G23780	F-box family protein
26	AT1G33270	Acyl transferase/acyl hydrolase/lysophospholipase superfamily protein
26	AT1G50570	Calcium-dependent lipid-binding (CaLB domain) family protein
26	AT1G56610	Protein with RNI-like/FBD-like domain
26	AT1G71960	ATP-binding casette family G25
26	AT1G74730	transmembrane protein, putative (DUF1118)
26	AT1G75800	Pathogenesis-related thaumatin superfamily protein
26	AT1G80230	Rubredoxin-like superfamily protein
26	AT2G02250	phloem protein 2-B2
26	AT2G03350	DUF538 family protein (Protein of unknown function, DUF538)
26	AT2G34380	Putative adipose-regulatory protein (Seipin)
26	AT2G37650	GRAS family transcription factor
26	AT3G02630	Plant stearoyl-acyl-carrier-protein desaturase family protein
26	AT3G12345	FKBP-type peptidyl-prolyl cis-trans isomerase
26	AT3G30180	brassinosteroid-6-oxidase 2

26	AT3G42630	Pentatricopeptide repeat (PPR) superfamily protein
26	AT3G51630	with no lysine (K) kinase 5
26	AT4G24120	YELLOW STRIPE like 1
26	AT5G48330	Regulator of chromosome condensation (RCC1) family protein
26	AT5G50850	Transketolase family protein
26	AT5G51840	junctophilin-like protein
26	AT5G52100	Dihydrodipicolinate reductase, bacterial/plant
27	AT1G12820	auxin signaling F-box 3
27	AT1G64180	intracellular protein transport protein USO1-like protein
27	AT1G78130	Major facilitator superfamily protein
27	AT2G14910	MAR-binding filament-like protein
27	AT2G19670	protein arginine methyltransferase 1A
27	AT2G41110	calmodulin 2
27	AT3G01400	ARM repeat superfamily protein
27	AT3G23640	heteroglycan glucosidase 1
27	AT3G56910	plastid-specific 50S ribosomal protein 5
27	AT4G01840	Ca2 activated outward rectifying K channel 5
27	AT4G11330	MAP kinase 5
27	AT4G13395	ROTUNDIFOLIA like 12
27	AT4G22340	cytidinediphosphate diacylglycerol synthase 2
27	AT4G38640	Plasma-membrane choline transporter family protein
27	AT4G39180	Sec14p-like phosphatidylinositol transfer family protein
27	AT5G21020	transmembrane protein
27	AT5G22640	MORN (Membrane Occupation and Recognition Nexus) repeat-containing protein
27	AT5G54095	proteoglycan-like protein
27	AT5G63510	gamma carbonic anhydrase like 1
27	AT5G64280	dicarboxylate transporter 2.2
28	AT1G08040	trimethylguanosine synthase (DUF707)
28	AT1G13930	oleosin-B3-like protein
28	AT1G55450	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
28	AT1G56520	Disease resistance protein (TIR-NBS-LRR class) family
28	AT1G66090	Disease resistance protein (TIR-NBS class)
28	AT1G72930	toll/interleukin-1 receptor-like protein
28	AT1G74710	ADC synthase superfamily protein
28	AT2G19500	cytokinin oxidase 2
28	AT2G23000	serine carboxypeptidase-like 10
28	AT2G36220	hypothetical protein
28	AT2G41100	Calcium-binding EF hand family protein
28	AT3G04210	Disease resistance protein (TIR-NBS class)
28	AT3G14840	Leucine-rich repeat transmembrane protein kinase
28	AT3G15790	methyl-CPG-binding domain 11
28	AT3G45600	tetraspanin3
28	AT3G48640	transmembrane protein
28	AT3G61420	BSD domain (BTF2-like transcription factors, Synapse-associated proteins and DOS2-like proteins)
28	AT3G62370	heme binding protein

28	AT4G24010	cellulose synthase like G1
28	AT4G24130	DUF538 family protein (Protein of unknown function, DUF538)
28	AT4G33360	NAD(P)-binding Rossmann-fold superfamily protein
28	AT4G33700	CBS domain protein (DUF21)
28	AT4G37970	cinnamyl alcohol dehydrogenase 6
28	AT5G03120	transmembrane protein
28	AT5G03350	Legume lectin family protein
28	AT5G43850	RmlC-like cupins superfamily protein
28	AT5G50580	SUMO-activating enzyme 1B
29	AT1G67360	Rubber elongation factor protein (REF)
29	AT1G74670	Gibberellin-regulated family protein
29	AT1G80120	LURP-one-like protein (DUF567)
29	AT2G20560	DNAJ heat shock family protein
29	AT2G20670	sugar phosphate exchanger, putative (DUF506)
29	AT2G30420	Homeodomain-like superfamily protein
29	AT2G45180	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
29	AT2G46420	helicase with zinc finger protein
29	AT2G48020	Major facilitator superfamily protein
29	AT3G15356	Legume lectin family protein
29	AT3G62950	Thioredoxin superfamily protein
29	AT4G13572	hypothetical protein
29	AT4G17810	C2H2 and C2HC zinc fingers superfamily protein
29	AT4G19410	Pectinacetylesterase family protein
29	AT4G32840	phosphofructokinase 6
29	AT5G14640	shaggy-like kinase 13
29	AT5G18600	Thioredoxin superfamily protein
29	AT5G25190	Integrase-type DNA-binding superfamily protein
29	AT5G44020	HAD superfamily, subfamily IIIB acid phosphatase
29	AT5G58690	phosphatidylinositol-speciwc phospholipase C5
29	AT5G59080	hypothetical protein
30	AT1G05350	NAD(P)-binding Rossmann-fold superfamily protein
30	AT1G06460	alpha-crystallin domain 32.1
30	AT1G29310	SecY protein transport family protein
30	AT1G31200	phloem protein 2-A9
30	AT1G35720	annexin 1
30	AT1G49720	abscisic acid responsive element-binding factor 1
30	AT1G53625	hypothetical protein
30	AT1G62810	Copper amine oxidase family protein
30	AT1G63010	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein
30	AT1G69740	Aldolase superfamily protein
30	AT1G73250	GDP-4-keto-6-deoxymannose-3,5-epimerase-4-reductase 1
30	AT1G76180	Dehydrin family protein
30	AT1G79630	Protein phosphatase 2C family protein
30	AT2G01670	nudix hydrolase homolog 17
30	AT2G16990	Major facilitator superfamily protein

30	AT2G19690	phospholipase A2-beta
30	AT2G32380	Transmembrane protein 97, Putative
30	AT2G46800	zinc transporter
30	AT3G20300	extracellular ligand-gated ion channel protein (DUF3537)
30	AT3G62800	double-stranded-RNA-binding protein 4
30	AT4G27020	inositol-1,4,5-trisphosphate 5-phosphatase
30	AT5G40950	ribosomal protein large subunit 27
30	AT5G47560	tonoplast dicarboxylate transporter
30	AT5G52550	stress response NST1-like protein
30	AT5G63180	Pectin lyase-like superfamily protein
31	AT1G16590	DNA-binding HORMA family protein
31	AT1G26110	decapping 5
31	AT1G28440	HAESA-like 1
31	AT1G54080	oligouridylate-binding protein 1A
31	AT2G28070	ABC-2 type transporter family protein
31	AT2G29200	pumilio 1
31	AT2G39950	flocculation protein
31	AT3G08800	ARM repeat superfamily protein
31	AT3G08840	D-alanine-D-alanine ligase family
31	AT3G11960	Cleavage and polyadenylation specificity factor (CPSF) A subunit protein
31	AT3G13340	Transducin/WD40 repeat-like superfamily protein
31	AT3G22210	transmembrane protein
31	AT4G18830	ovate family protein 5
31	AT4G24450	phosphoglucan, water dikinase
31	AT4G32620	Enhancer of polycomb-like transcription factor protein
31	AT4G32850	nuclear poly(a) polymerase
31	AT4G38600	HECT ubiquitin protein ligase family protein KAK
31	AT5G03280	NRAMP metal ion transporter family protein
31	AT5G03780	TRF-like 10
31	AT5G04670	Enhancer of polycomb-like transcription factor protein
31	AT5G08720	polyketide cyclase/dehydrase/lipid transporter
31	AT5G16270	sister chromatid cohesion 1 protein 4
31	AT5G19500	Tryptophan/tyrosine permease
31	AT5G22620	phosphoglycerate/bisphosphoglycerate mutase family protein
31	AT5G42220	Ubiquitin-like superfamily protein
31	AT5G60640	PDI-like 1-4
32	AT1G09430	ATP-citrate lyase A-3
32	AT1G10820	hypothetical protein (DUF3755)
32	AT1G33610	Leucine-rich repeat (LRR) family protein
32	AT1G45191	Glycosyl hydrolase superfamily protein
32	AT1G70530	cysteine-rich RLK (RECEPTOR-like protein kinase) 3
32	AT1G80510	Transmembrane amino acid transporter family protein
32	AT2G16770	Basic-leucine zipper (bZIP) transcription factor family protein
32	AT2G25950	PITH domain protein (DUF1000)
32	AT2G35930	plant U-box 23
32	AT3G01210	RNA-binding (RRM/RBD/RNP motifs) family protein

32	AT3G03900	adenosine-5'-phosphosulfate (APS) kinase 3
32	AT3G27750	PPR containing protein
32	AT3G32090	WRKY family transcription factor
32	AT3G60910	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
32	AT4G15450	Senescence/dehydration-associated protein-like protein
32	AT4G27657	hypothetical protein
32	AT4G27920	PYR1-like 10
32	AT5G01075	Glycosyl hydrolase family 35 protein
32	AT5G07790	hypothetical protein
32	AT5G13190	GSH-induced LITAF domain protein
32	AT5G37550	hypothetical protein
32	AT5G45110	NPR1-like protein 3
32	AT5G52420	transmembrane protein
33	AT1G16515	transmembrane protein
33	AT1G64930	cytochrome P450, family 87, subfamily A, polypeptide 7
33	AT1G78480	Prenyltransferase family protein
33	AT2G34720	nuclear factor Y, subunit A4
33	AT3G48240	Octicosapeptide/Phox/Bem1p family protein
33	AT4G03070	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
33	AT4G32295	histone acetyltransferase
33	AT4G33950	Protein kinase superfamily protein
33	AT5G02930	F-box/RNI-like superfamily protein
33	AT5G03230	senescence regulator (Protein of unknown function, DUF584)
33	AT5G04440	RAP release 2, galactose-binding-like domain protein, putative (DUF1997)
33	AT5G17450	Heavy metal transport/detoxification superfamily protein
33	AT5G50240	protein-l-isoaspartate methyltransferase 2
34	AT1G03130	photosystem I subunit D-2
34	AT1G04530	Tetratricopeptide repeat (TPR)-like superfamily protein
34	AT1G15820	light harvesting complex photosystem II subunit 6
34	AT1G24793	UDP-3-O-acyl N-acetylglycosamine deacetylase family protein
34	AT1G67370	DNA-binding HORMA family protein
34	AT2G02955	maternal effect embryo arrest 12
34	AT2G03670	cell division cycle 48B
34	AT2G34430	light-harvesting chlorophyll-protein complex II subunit B1
34	AT2G37400	Tetratricopeptide repeat (TPR)-like superfamily protein
34	AT2G38185	RING/U-box superfamily protein
34	AT3G25430	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
34	AT3G30730	hypothetical protein
34	AT3G48850	phosphate transporter 3;2
34	AT4G00690	UB-like protease 1B
34	AT4G08580	microfibrillar-associated protein-like protein
34	AT4G12750	Homeodomain-like transcriptional regulator
34	AT4G30935	WRKY DNA-binding protein 32
34	AT4G38440	RPAP1-like, carboxy-terminal protein
34	AT5G04000	hypothetical protein

34	AT5G07950	hypothetical protein
34	AT5G11840	YCF36, putative (DUF1230)
34	AT5G14770	PPR repeat protein
34	AT5G58980	Neutral/alkaline non-lysosomal ceramidase
35	AT1G32400	tobamovirus multiplication 2A
35	AT2G34180	CBL-interacting protein kinase 13
35	AT2G45630	D-isomer specific 2-hydroxyacid dehydrogenase family protein
35	AT2G47140	NAD(P)-binding Rossmann-fold superfamily protein
35	AT3G10910	RING/U-box superfamily protein
35	AT3G11420	beta-1,3-N-acetylglucosaminyltransferase lunatic protein, putative (DUF604)
35	AT3G11690	hypothetical protein
35	AT3G22730	F-box and associated interaction domains-containing protein
35	AT3G26670	magnesium transporter, putative (DUF803)
35	AT3G48690	alpha/beta-Hydrolases superfamily protein
35	AT4G25670	stress response NST1-like protein
35	AT4G28350	Concanavalin A-like lectin protein kinase family protein
35	AT4G38060	hypothetical protein
35	AT5G01520	RING/U-box superfamily protein
35	AT5G10840	Endomembrane protein 70 protein family
35	AT5G40790	hypothetical protein
35	AT5G43260	chaperone protein dnaJ-like protein
35	AT5G57040	Lactoylglutathione lyase / glyoxalase I family protein
35	AT5G62100	BCL-2-associated athanogene 2
35	AT5G65830	receptor like protein 57
36	AT1G06210	ENTH/VHS/GAT family protein
36	AT1G24265	bZIP transcription factor, putative (DUF1664)
36	AT1G27030	hypothetical protein
36	AT1G64690	branchless trichome
36	AT1G68550	Integrase-type DNA-binding superfamily protein
36	AT1G80300	nucleotide transporter 1
36	AT3G49220	Plant invertase/pectin methylesterase inhibitor superfamily
36	AT3G59300	Pentatricopeptide repeat (PPR) superfamily protein
36	AT4G18250	receptor Serine/Threonine kinase-like protein
36	AT4G27560	UDP-Glycosyltransferase superfamily protein
36	AT4G27570	UDP-Glycosyltransferase superfamily protein
36	AT5G01880	RING/U-box superfamily protein
36	AT5G19230	Glycoprotein membrane precursor GPI-anchored
36	AT5G48540	receptor-like protein kinase-related family protein
36	AT5G62630	hipl2 protein precursor
36	AT5G63020	Disease resistance protein (CC-NBS-LRR class) family
36	AT5G65280	GCR2-like 1
37	AT1G08630	threonine aldolase 1
37	AT1G12790	DNA ligase-like protein
37	AT1G27080	nitrate transporter 1.6
37	AT1G28600	GDSL-like Lipase/Acylhydrolase superfamily protein
37	AT1G66400	calmodulin like 23

37	AT2G17430	Seven transmembrane MLO family protein
37	AT2G46940	fold protein
37	AT3G01060	lysine-tRNA ligase
37	AT3G29090	pectin methylesterase 31
37	AT3G57190	peptide chain release factor
37	AT4G23000	Calcineurin-like metallo-phosphoesterase superfamily protein
37	AT4G30830	myosin-like protein (Protein of unknown function, DUF593)
37	AT4G37140	alpha/beta-Hydrolases superfamily protein
37	AT5G03240	polyubiquitin 3
37	AT5G12860	dicarboxylate transporter 1
37	AT5G49448	
37	AT5G57790	hypothetical protein
37	AT5G58760	damaged DNA binding 2
38	AT1G08060	ATP-dependent helicase family protein
38	AT1G08650	phosphoenolpyruvate carboxylase kinase 1
38	AT1G18620	LONGIFOLIA protein
38	AT1G22910	RNA-binding (RRM/RBD/RNP motifs) family protein
38	AT1G33340	ENTH/ANTH/VHS superfamily protein
38	AT1G70260	nodulin MtN21 /EamA-like transporter family protein
38	AT1G70560	tryptophan aminotransferase of Arabidopsis 1
38	AT2G05440	GLYCINE RICH PROTEIN 9
38	AT3G14940	phosphoenolpyruvate carboxylase 3
38	AT3G18520	histone deacetylase 15
38	AT3G18773	RING/U-box superfamily protein
38	AT3G61700	helicase with zinc finger protein
38	AT3G63450	RNA-binding (RRM/RBD/RNP motifs) family protein
38	AT4G00730	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
38	AT4G32340	Tetratricopeptide repeat (TPR)-like superfamily protein
38	AT4G37540	LOB domain-containing protein 39
38	AT5G51770	Protein kinase superfamily protein
39	AT1G27520	Glycosyl hydrolase family 47 protein
39	AT1G59750	auxin response factor 1
39	AT2G26200	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
39	AT2G37470	Histone superfamily protein
39	AT2G44920	Tetratricopeptide repeat (TPR)-like superfamily protein
39	AT3G18190	TCP-1/cpn60 chaperonin family protein
39	AT3G20680	plant/protein (DUF1995)
39	AT3G50670	U1 small nuclear ribonucleoprotein-70K
39	AT3G59870	hypothetical protein
39	AT4G02720	Ras-induced vulval development antagonist protein
39	AT4G16143	importin alpha isoform 2
39	AT4G32208	heat shock protein 70 (Hsp 70) family protein
39	AT4G38340	Plant regulator RWP-RK family protein
39	AT5G10250	Phototropic-responsive NPH3 family protein
39	AT5G12170	CRT (chloroquine-resistance transporter)-like transporter 3

39	AT5G27320	alpha/beta-Hydrolases superfamily protein
39	AT5G63840	Glycosyl hydrolases family 31 protein
40	AT1G73230	Nascent polypeptide-associated complex NAC
40	AT1G74060	Ribosomal protein L6 family protein
40	AT2G07698	ATPase, F1 complex, alpha subunit protein
40	AT2G19730	Ribosomal L28e protein family
40	AT2G20780	Major facilitator superfamily protein
40	AT2G36170	60S ribosomal protein L40-1
40	AT3G47370	Ribosomal protein S10p/S20e family protein
40	AT3G56340	Ribosomal protein S26e family protein
40	AT3G61830	auxin response factor 18
40	AT4G02580	NADH-ubiquinone oxidoreductase 24 kDa subunit
40	AT4G04770	ATP binding cassette protein 1
40	AT5G02610	Ribosomal L29 family protein
40	AT5G20320	dicer-like 4
40	AT5G63400	adenylate kinase 1
40	AT5G64130	cAMP-regulated phosphoprotein 19-related protein
41	AT1G07130	Nucleic acid-binding, OB-fold-like protein
41	AT1G18750	AGAMOUS-like 65
41	AT1G21000	PLATZ transcription factor family protein
41	AT2G03340	WRKY DNA-binding protein 3
41	AT2G17710	Big1
41	AT2G21160	Translocon-associated protein (TRAP), alpha subunit
41	AT2G26980	CBL-interacting protein kinase 3
41	AT2G27420	Cysteine proteinases superfamily protein
41	AT2G28470	beta-galactosidase 8
41	AT2G37180	Aquaporin-like superfamily protein
41	AT2G40711	hypothetical protein
41	AT2G46460	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
41	AT3G08030	DNA-directed RNA polymerase subunit beta (Protein of unknown function, DUF642)
41	AT3G14600	Ribosomal protein L18ae/LX family protein
41	AT3G21390	Mitochondrial substrate carrier family protein
41	AT3G46540	ENTH/VHS family protein
41	AT4G34610	BEL1-like homeodomain 6
41	AT5G06850	C2 calcium/lipid-binding plant phosphoribosyltransferase family protein
41	AT5G59480	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
41	AT5G66160	receptor homology region transmembrane domain ring H2 motif protein 1
42	AT1G07440	NAD(P)-binding Rossmann-fold superfamily protein
42	AT1G23860	RS-containing zinc finger protein 21
42	AT1G53460	craniofacial development protein
42	AT1G67440	Minichromosome maintenance (MCM2/3/5) family protein
42	AT2G15580	RING/U-box superfamily protein
42	AT2G22640	BRICK1
42	AT3G09980	ankyrin repeat 30A-like protein (DUF662)
42	AT3G10210	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein

42	AT3G16390	nitrile specifier protein 3
42	AT3G16400	nitrile specifier protein 1
42	AT4G01690	Flavin containing amine oxidoreductase family
43	AT1G07770	ribosomal protein S15A
43	AT1G17500	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
43	AT2G40130	Double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases superfamily protein
43	AT2G43560	FKBP-like peptidyl-prolyl cis-trans isomerase family protein
43	AT3G10405	vacuolar acid trehalase
43	AT3G16940	calmodulin-binding transcription activator
43	AT3G17350	wall-associated receptor kinase carboxy-terminal protein
43	AT3G17780	B-cell receptor-associated-like protein
43	AT4G30620	Putative BCR, YbaB family COG0718
43	AT4G31290	ChaC-like family protein
43	AT4G34680	GATA transcription factor 3
43	AT5G03100	F-box/RNI-like superfamily protein
43	AT5G13430	Ubiquinol-cytochrome C reductase iron-sulfur subunit
43	AT5G20060	alpha/beta-Hydrolases superfamily protein
43	AT5G28060	Ribosomal protein S24e family protein
43	AT5G60600	4-hydroxy-3-methylbut-2-enyl diphosphate synthase
43	AT5G66770	GRAS family transcription factor
44	AT1G04350	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
44	AT1G23480	cellulose synthase-like A3
44	AT1G27210	ARM repeat superfamily protein
44	AT1G49130	B-box type zinc finger protein with CCT domain-containing protein
44	AT1G49730	Protein kinase superfamily protein
44	AT1G52080	actin binding protein family
44	AT1G53840	pectin methylesterase 1
44	AT1G63260	tetraspanin10
44	AT1G64370	filaggrin-like protein
44	AT1G72620	alpha/beta-Hydrolases superfamily protein
44	AT2G04350	AMP-dependent synthetase and ligase family protein
44	AT3G06470	GNS1/SUR4 membrane protein family
44	AT3G28270	transmembrane protein, putative (DUF677)
44	AT4G01460	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
44	AT4G30410	sequence-specific DNA binding transcription factor
44	AT4G30630	hypothetical protein
44	AT5G07870	HXXXD-type acyl-transferase family protein
44	AT5G44340	tubulin beta chain 4
45	AT1G07430	highly ABA-induced PP2C protein 2
45	AT1G17210	IAP-like protein 1
45	AT1G27200	glycosyltransferase family protein (DUF23)
45	AT1G29430	SAUR-like auxin-responsive protein family
45	AT1G70300	K uptake permease 6
45	AT1G76490	hydroxy methylglutaryl CoA reductase 1
45	AT2G33280	Major facilitator superfamily protein

45	AT3G24130	Pectin lyase-like superfamily protein
45	AT4G16620	nodulin MtN21 /EamA-like transporter family protein
45	AT4G32920	glycine-rich protein
45	AT4G34980	subtilisin-like serine protease 2
45	AT4G35300	tonoplast monosaccharide transporter2
45	AT5G14280	DNA-binding storekeeper-like protein
45	AT5G19900	PRLI-interacting factor
45	AT5G27270	Tetratricopeptide repeat (TPR)-like superfamily protein
46	AT1G19400	Erythronate-4-phosphate dehydrogenase family protein
46	AT1G51450	TRAUCO
46	AT1G57630	Toll-Interleukin-Resistance (TIR) domain family protein
46	AT1G79970	hypothetical protein
46	AT2G23320	WRKY DNA-binding protein 15
46	AT2G41870	Remorin family protein
46	AT3G01290	SPFH/Band 7/PHB domain-containing membrane-associated protein family
46	AT4G00340	receptor-like protein kinase 4
46	AT4G00630	K efflux antiporter 2
46	AT4G13150	transmembrane protein
46	AT4G21630	Subtilase family protein
46	AT5G08300	Succinyl-CoA ligase, alpha subunit
46	AT5G14690	transmembrane protein
46	AT5G46230	hypothetical protein (Protein of unknown function, DUF538)
46	AT5G63860	Regulator of chromosome condensation (RCC1) family protein
47	AT1G32540	lsd one like 1
47	AT1G64660	methionine gamma-lyase
47	AT1G75250	RAD-like 6
47	AT2G18300	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
47	AT2G30490	cinnamate-4-hydroxylase
47	AT3G21700	Ras-related small GTP-binding family protein
47	AT3G22840	Chlorophyll A-B binding family protein
47	AT3G51240	flavanone 3-hydroxylase
47	AT4G09750	NAD(P)-binding Rossmann-fold superfamily protein
47	AT4G23990	cellulose synthase like G3
47	AT5G01600	ferretin 1
48	AT1G22780	Ribosomal protein S13/S18 family
48	AT1G22830	Tetratricopeptide repeat (TPR)-like superfamily protein
48	AT2G40010	Ribosomal protein L10 family protein
48	AT2G42740	ribosomal protein large subunit 16A
48	AT2G44040	Dihydrodipicolinate reductase, bacterial/plant
48	AT3G04840	Ribosomal protein S3Ae
48	AT3G11510	Ribosomal protein S11 family protein
48	AT3G47800	Galactose mutarotase-like superfamily protein
48	AT4G18100	Ribosomal protein L32e
48	AT4G19100	PAM68-like protein (DUF3464)
48	AT4G34290	SWIB/MDM2 domain superfamily protein
48	AT4G39040	RNA-binding CRS1 / YhbY (CRM) domain protein

48	AT4G39200	Ribosomal protein S25 family protein
48	AT5G62300	Ribosomal protein S10p/S20e family protein
49	AT1G18740	ROH1, putative (DUF793)
49	AT1G23390	Kelch repeat-containing F-box family protein
49	AT1G32900	UDP-Glycosyltransferase superfamily protein
49	AT1G52310	protein kinase family protein / C-type lectin domain-containing protein
49	AT1G66080	hikeshi-like protein
49	AT3G06070	hypothetical protein
49	AT3G13750	beta galactosidase 1
49	AT3G14720	MAP kinase 19
49	AT3G17130	Plant invertase/pectin methylesterase inhibitor superfamily protein
49	AT3G24150	hypothetical protein
49	AT4G18130	phytochrome E
49	AT4G33140	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
49	AT4G35030	Protein kinase superfamily protein
49	AT4G39800	myo-inositol-1-phosphate synthase 1
49	AT5G37050	sorting nexin
49	AT5G59430	telomeric repeat binding protein 1
50	AT1G10700	phosphoribosyl pyrophosphate (PRPP) synthase 3
50	AT1G14530	tobamovirus multiplication-like protein (DUF1084)
50	AT1G79110	zinc ion binding protein
50	AT2G19810	CCCH-type zinc finger family protein
50	AT3G02300	Regulator of chromosome condensation (RCC1) family protein
50	AT3G21000	Gag-Pol-related retrotransposon family protein
50	AT3G28050	nodulin MtN21 /EamA-like transporter family protein
50	AT4G00250	DNA-binding storekeeper protein-related transcriptional regulator
50	AT4G01595	Protein kinase superfamily protein
50	AT4G34265	hypothetical protein
50	AT5G23240	DNAJ heat shock N-terminal domain-containing protein
50	AT5G28300	Duplicated homeodomain-like superfamily protein
50	AT5G58960	glucose-6-phosphate isomerase, putative (DUF641)
50	AT5G62090	SEUSS-like 2
51	AT1G28610	GDSL-like Lipase/Acylhydrolase superfamily protein
51	AT1G65930	cytosolic NADP -dependent isocitrate dehydrogenase
51	AT2G44500	O-fucosyltransferase family protein
51	AT3G09550	Ankyrin repeat family protein
51	AT3G17998	
51	AT3G21080	ABC transporter-like protein
51	AT3G25710	basic helix-loop-helix 32
51	AT3G62690	AtL5
51	AT4G09500	UDP-Glycosyltransferase superfamily protein
51	AT4G14440	3-hydroxyacyl-CoA dehydratase 1
51	AT4G29950	Ypt/Rab-GAP domain of gyp1p superfamily protein
51	AT4G30280	xyloglucan endotransglucosylase/hydrolase 18
51	AT4G36240	GATA transcription factor 7
51	AT5G20820	SAUR-like auxin-responsive protein family

51	AT5G48050	Copia-like polyprotein/retrotransposon
52	AT1G01430	TRICHOME BIREFRINGENCE-LIKE 25
52	AT1G15410	aspartate-glutamate racemase family
52	AT1G54010	GDSL-like Lipase/Acylhydrolase superfamily protein
52	AT1G64080	membrane-associated kinase regulator
52	AT1G72430	SAUR-like auxin-responsive protein family
52	AT1G73270	serine carboxypeptidase-like 6
52	AT2G32100	ovate family protein 16
52	AT4G01026	PYR1-like 7
52	AT4G08570	Heavy metal transport/detoxification superfamily protein
52	AT5G02480	HSP20-like chaperones superfamily protein
52	AT5G56170	LORELEI-LIKE-GPI-ANCHORED PROTEIN 1
52	AT5G63960	DNA polymerase delta subunit 1
52	AT5G67420	LOB domain-containing protein 37
53	AT1G03040	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
53	AT1G07610	metallothionein 1C
53	AT1G28230	purine permease 1
53	AT1G29600	Zinc finger C-x8-C-x5-C-x3-H type family protein
53	AT1G80130	Tetratricopeptide repeat (TPR)-like superfamily protein
53	AT1G80530	Major facilitator superfamily protein
53	AT2G32070	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
53	AT3G15060	RAB GTPase homolog A1G
53	AT3G20270	lipid-binding serum glycoprotein family protein
53	AT3G52710	hypothetical protein
53	AT3G57120	Protein kinase superfamily protein
53	AT4G26640	WRKY family transcription factor family protein
53	AT5G16570	glutamine synthetase 1;4
53	AT5G35490	mto 1 responding up 1
53	AT5G66820	transmembrane protein
54	AT1G22540	Major facilitator superfamily protein
54	AT1G23200	Plant invertase/pectin methylesterase inhibitor superfamily
54	AT1G28120	ubiquitin thioesterase otubain-like protein
54	AT1G76570	Chlorophyll A-B binding family protein
54	AT3G05930	germin-like protein 8
54	AT3G09630	Ribosomal protein L4/L1 family
54	AT3G10410	SERINE CARBOXYPEPTIDASE-LIKE 49
54	AT3G23000	CBL-interacting protein kinase 7
54	AT3G49010	breast basic conserved 1
54	AT3G52610	GATA zinc finger protein
54	AT4G13590	Uncharacterized protein family (UPF0016)
54	AT4G16490	ARM repeat superfamily protein
54	AT5G25610	BURP domain-containing protein
54	AT5G53140	Protein phosphatase 2C family protein
54	AT5G53540	P-loop containing nucleoside triphosphate hydrolases superfamily protein
55	AT1G15930	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
55	AT1G28350	Nucleotidylyl transferase superfamily protein

55	AT1G41880	Ribosomal protein L35Ae family protein
55	AT1G53250	histone-lysine N-methyltransferase, H3 lysine-79 specific-like protein
55	AT1G77930	Chaperone DnaJ-domain superfamily protein
55	AT2G01850	endoxyloglucan transferase A3
55	AT2G02450	NAC domain containing protein 35
55	AT2G39030	Acyl-CoA N-acyltransferases (NAT) superfamily protein
55	AT3G12930	Lojap-related protein
55	AT3G24830	Ribosomal protein L13 family protein
55	AT3G25520	ribosomal protein L5
55	AT4G17390	Ribosomal protein L23/L15e family protein
55	AT5G01750	LURP-one-like protein (DUF567)
55	AT5G03200	RING/U-box superfamily protein
56	AT1G05575	transmembrane protein
56	AT1G49660	carboxyesterase 5
56	AT2G05920	Subtilase family protein
56	AT2G21300	ATP binding microtubule motor family protein
56	AT2G26560	phospholipase A 2A
56	AT2G31865	poly(ADP-ribose) glycohydrolase 2
56	AT2G41250	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
56	AT3G09540	Pectin lyase-like superfamily protein
56	AT3G59080	Eukaryotic aspartyl protease family protein
56	AT4G18205	Nucleotide-sugar transporter family protein
56	AT4G19660	NPR1-like protein 4
56	AT4G26950	senescence regulator (Protein of unknown function, DUF584)
56	AT5G25250	SPFH/Band 7/PHB domain-containing membrane-associated protein family
56	AT5G62770	membrane-associated kinase regulator, putative (DUF1645)
57	AT1G19380	sugar, putative (DUF1195)
57	AT1G20970	calponin-like domain protein
57	AT1G22360	UDP-glucosyl transferase 85A2
57	AT1G78830	Curculin-like (mannose-binding) lectin family protein
57	AT2G22990	sinapoylglucose 1
57	AT2G44670	senescence-associated family protein (DUF581)
57	AT3G03990	alpha/beta-Hydrolases superfamily protein
57	AT3G04720	pathogenesis-related 4
57	AT3G18160	peroxin 3-1
57	AT5G03650	starch branching enzyme 2.2
57	AT5G06860	polygalacturonase inhibiting protein 1
57	AT5G06870	polygalacturonase inhibiting protein 2
57	AT5G39080	HXXXD-type acyl-transferase family protein
57	AT5G54880	DTW domain-containing protein
57	AT5G57550	xyloglucan endotransglucosylase/hydrolase 25
57	AT5G65660	hydroxyproline-rich glycoprotein family protein
58	AT1G56570	Tetratricopeptide repeat (TPR)-like superfamily protein
58	AT2G36250	Tubulin/FtsZ family protein
58	AT3G07660	flocculation protein (DUF1296)
58	AT3G26520	tonoplast intrinsic protein 2

58	AT4G27610	intracellular protein transporter
58	AT4G37390	Auxin-responsive GH3 family protein
58	AT5G03435	Ca2 dependent plant phosphoribosyltransferase family protein
58	AT5G06680	spindle pole body component 98
58	AT5G19740	Peptidase M28 family protein
58	AT5G22020	Calcium-dependent phosphotriesterase superfamily protein
58	AT5G27860	hypothetical protein
58	AT5G53660	growth-regulating factor 7
58	AT5G64120	Peroxidase superfamily protein
59	AT1G23020	ferric reduction oxidase 3
59	AT1G55740	seed imbibition 1
59	AT1G66760	MATE efflux family protein
59	AT1G71870	MATE efflux family protein
59	AT1G75220	Major facilitator superfamily protein
59	AT2G01660	plasmodesmata-located protein 6
59	AT3G08770	lipid transfer protein 6
59	AT3G17070	Peroxidase family protein
59	AT3G26790	AP2/B3-like transcriptional factor family protein
59	AT3G59420	crinkly4
59	AT4G01390	TRAF-like family protein
59	AT4G12470	azelaic acid induced 1
59	AT4G15620	Uncharacterized protein family (UPF0497)
59	AT4G17030	expansin-like B1
59	AT4G24480	Protein kinase superfamily protein
60	AT1G06980	6,7-dimethyl-8-ribityllumazine synthase
60	AT1G19650	Sec14p-like phosphatidylinositol transfer family protein
60	AT1G51690	protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform
60	AT2G01860	Tetratricopeptide repeat (TPR)-like superfamily protein
60	AT2G16380	Sec14p-like phosphatidylinositol transfer family protein
60	AT2G17280	Phosphoglycerate mutase family protein
60	AT2G22930	UDP-Glycosyltransferase superfamily protein
60	AT2G30300	Major facilitator superfamily protein
60	AT3G17550	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
60	AT3G56260	hypothetical protein
60	AT4G33440	Pectin lyase-like superfamily protein
60	AT4G36500	hypothetical protein
60	AT5G11460	hypothetical protein (DUF581)
60	AT5G19120	Eukaryotic aspartyl protease family protein
60	AT5G25265	Hyp O-arabinosyltransferase-like protein
61	AT1G12320	ankyrin repeat/KH domain protein (DUF1442)
61	AT1G51140	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
61	AT2G03140	alpha/beta-Hydrolases superfamily protein
61	AT2G12190	Cytochrome P450 superfamily protein
61	AT3G20620	F-box family protein-like protein
61	AT4G18810	NAD(P)-binding Rossmann-fold superfamily protein
61	AT5G17050	UDP-glucosyl transferase 78D2

61	AT5G20280	sucrose phosphate synthase 1F
61	AT5G43380	type one serine/threonine protein phosphatase 6
61	AT5G53970	Tyrosine transaminase family protein
61	AT5G59050	G patch domain protein
62	AT1G03380	yeast autophagy 18 G-like protein
62	AT1G07380	Neutral/alkaline non-lysosomal ceramidase
62	AT1G30530	UDP-glucosyl transferase 78D1
62	AT1G47330	methyltransferase, putative (DUF21)
62	AT2G01820	Leucine-rich repeat protein kinase family protein
62	AT2G15695	peptide methionine sulfoxide reductase (Protein of unknown function DUF829, transmembrane 53)
62	AT2G32290	beta-amylase 6
62	AT2G35940	BEL1-like homeodomain 1
62	AT2G38800	Plant calmodulin-binding protein-like protein
62	AT2G39850	Subtilisin-like serine endopeptidase family protein
62	AT2G40840	disproportionating enzyme 2
62	AT3G22440	FRIGIDA-like protein
62	AT4G13800	magnesium transporter NIPA (DUF803)
62	AT4G33740	myb-like protein X
62	AT5G03430	phosphoadenosine phosphosulfate (PAPS) reductase family protein
62	AT5G43600	ureidoglycolate amidohydrolase
63	AT1G01580	ferric reduction oxidase 2
63	AT1G24580	RING/U-box superfamily protein
63	AT1G47271	Cystathionine beta-synthase (CBS) family protein
63	AT1G68520	B-box type zinc finger protein with CCT domain-containing protein
63	AT2G13360	alanine:glyoxylate aminotransferase
63	AT3G24750	hypothetical protein
63	AT3G46820	type one serine/threonine protein phosphatase 5
63	AT3G50270	HXXXD-type acyl-transferase family protein
63	AT4G17530	RAB GTPase homolog 1C
63	AT5G05440	Polyketide cyclase/dehydrase and lipid transport superfamily protein
63	AT5G24880	chromo domain cec-like protein
63	AT5G63480	mediator of RNA polymerase II transcription subunit
64	AT1G01420	UDP-glucosyl transferase 72B3
64	AT1G66750	CDK-activating kinase 4
64	AT1G70090	glucosyl transferase family 8
64	AT2G17470	aluminum activated malate transporter family protein
64	AT3G06300	P4H isoform 2
64	AT3G26800	transmembrane protein
64	AT4G12110	sterol-4alpha-methyl oxidase 1-1
64	AT4G19450	Major facilitator superfamily protein
64	AT5G41700	ubiquitin conjugating enzyme 8
64	AT5G58140	phototropin 2
65	AT1G07420	sterol 4-alpha-methyl-oxidase 2-1
65	AT1G13360	hypothetical protein
65	AT1G62050	Ankyrin repeat family protein
65	AT1G73170	P-loop containing nucleoside triphosphate hydrolases superfamily

		protein
65	AT1G75810	transmembrane protein
65	AT2G36390	starch branching enzyme 2.1
65	AT3G14595	Ribosomal protein L18ae family
65	AT4G12000	SNARE associated Golgi protein family
65	AT4G12990	transmembrane protein
65	AT5G06060	NAD(P)-binding Rossmann-fold superfamily protein
65	AT5G38344	Toll-Interleukin-Resistance (TIR) domain family protein
65	AT5G39340	histidine-containing phosphotransmitter 3
65	AT5G50100	Putative thiol-disulfide oxidoreductase DCC
65	AT5G54855	Pollen Ole e 1 allergen and extensin family protein
66	AT1G05170	Galactosyltransferase family protein
66	AT1G05870	hypothetical protein (DUF1685)
66	AT1G19490	Basic-leucine zipper (bZIP) transcription factor family protein
66	AT1G32870	NAC domain protein 13
66	AT1G60190	ARM repeat superfamily protein
66	AT1G75490	Integrase-type DNA-binding superfamily protein
66	AT2G17450	RING-H2 finger A3A
66	AT2G27050	ETHYLENE-INSENSITIVE3-like 1
66	AT3G17611	RHOMBOID-like protein 14
66	AT3G22420	with no lysine (K) kinase 2
66	AT4G38470	ACT-like protein tyrosine kinase family protein
66	AT5G18830	squamosa promoter binding protein-like 7
66	AT5G28770	bZIP transcription factor family protein
66	AT5G53160	regulatory components of ABA receptor 3
67	AT1G06149	
67	AT1G19020	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase
67	AT1G26500	Pentatricopeptide repeat (PPR) superfamily protein
67	AT2G18670	RING/U-box superfamily protein
67	AT2G18690	transmembrane protein
67	AT2G20142	Toll-Interleukin-Resistance (TIR) domain family protein
67	AT2G23810	tetraspanin8
67	AT3G13437	transmembrane protein
67	AT3G46080	C2H2-type zinc finger family protein
67	AT3G51480	glutamate receptor 3.6
67	AT4G01770	rhamnogalacturonan xylosyltransferase 1
67	AT4G11890	Protein kinase superfamily protein
67	AT4G14365	hypothetical protein
68	AT1G04440	casein kinase like 13
68	AT1G18570	myb domain protein 51
68	AT1G26650	Son of sevenless protein
68	AT1G61740	Sulfite exporter TauE/SafE family protein
68	AT1G68610	PLANT CADMIUM RESISTANCE 11
68	AT1G74520	HVA22 homologue A
68	AT1G80460	Actin-like ATPase superfamily protein
68	AT2G03730	ACT domain repeat 5

68	AT2G17300	hypothetical protein
68	AT3G04670	WRKY DNA-binding protein 39
68	AT3G12950	Trypsin family protein
68	AT3G13275	transmembrane protein
68	AT3G63530	RING/U-box superfamily protein
68	AT4G30120	heavy metal atpase 3
68	AT5G17660	tRNA (guanine-N-7) methyltransferase
68	AT5G20950	Glycosyl hydrolase family protein
68	AT5G55950	Nucleotide/sugar transporter family protein
69	AT2G23470	root UVB sensitive protein (Protein of unknown function, DUF647)
69	AT3G29590	HXXXD-type acyl-transferase family protein
69	AT4G04510	cysteine-rich RLK (RECEPTOR-like protein kinase) 38
69	AT4G22870	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
69	AT4G22880	leucoanthocyanidin dioxygenase
69	AT5G07990	Cytochrome P450 superfamily protein
69	AT5G17220	glutathione S-transferase phi 12
69	AT5G18930	Adenosylmethionine decarboxylase family protein
69	AT5G42800	dihydroflavonol 4-reductase
69	AT5G54060	UDP-glucose:flavonoid 3-o-glucosyltransferase
70	AT1G25230	Calcineurin-like metallo-phosphoesterase superfamily protein
70	AT1G73290	serine carboxypeptidase-like 5
70	AT3G02170	longifolia2
70	AT3G21560	UDP-Glycosyltransferase superfamily protein
70	AT3G48280	cytochrome P450, family 71, subfamily A, polypeptide 25
70	AT4G03210	xyloglucan endotransglucosylase/hydrolase 9
70	AT4G15480	UDP-Glycosyltransferase superfamily protein
70	AT4G27710	cytochrome P450, family 709, subfamily B, polypeptide 3
70	AT5G40890	chloride channel A
70	AT5G43830	aluminum induced protein with YGL and LRDR motifs
70	AT5G49740	ferric reduction oxidase 7
71	AT1G13210	autoinhibited Ca2 /ATPase II
71	AT1G18660	zinc finger (C3HC4-type RING finger) family protein
71	AT1G66540	Cytochrome P450 superfamily protein
71	AT1G73630	EF hand calcium-binding protein family
71	AT1G76800	Vacuolar iron transporter (VIT) family protein
71	AT3G09650	Tetratricopeptide repeat (TPR)-like superfamily protein
71	AT3G56290	potassium transporter
71	AT4G09350	Chaperone DnaJ-domain superfamily protein
71	AT4G32770	tocopherol cyclase, chloroplast / vitamin E deficient 1 (VTE1) / sucrose export defective 1 (SXD1)
71	AT4G35320	hypothetical protein
71	AT4G36660	polyol transporter, putative (DUF1195)
71	AT5G20070	nudix hydrolase homolog 19
71	AT5G42146	transmembrane protein
71	AT5G59350	transmembrane protein
71	AT5G65600	Concanavalin A-like lectin protein kinase family protein

72	AT1G16730	hypothetical protein
72	AT1G17100	SOUL heme-binding family protein
72	AT1G20470	SAUR-like auxin-responsive protein family
72	AT2G33250	transmembrane protein
72	AT2G42975	myosin-G heavy chain-like protein
72	AT3G63120	cyclin p1;1
72	AT4G17050	ureidoglycine aminohydrolase
72	AT4G37925	subunit NDH-M of NAD(P)H:plastoquinone dehydrogenase complex
72	AT5G23920	transmembrane protein
73	AT1G01360	regulatory component of ABA receptor 1
73	AT1G07080	Thioredoxin superfamily protein
73	AT1G12610	Integrase-type DNA-binding superfamily protein
73	AT3G05360	receptor like protein 30
73	AT3G50340	hypothetical protein
73	AT3G55400	methionyl-tRNA synthetase / methionine-tRNA ligase / MetRS (cpMetRS)
73	AT4G02640	bZIP transcription factor family protein
73	AT4G14200	Pentatricopeptide repeat (PPR) superfamily protein
73	AT4G16500	Cystatin/monellin superfamily protein
73	AT4G27990	YGGT family protein
73	AT5G62920	response regulator 6
74	AT1G08980	amidase 1
74	AT1G15890	Disease resistance protein (CC-NBS-LRR class) family
74	AT1G70290	trehalose-6-phosphatase synthase S8
74	AT2G17550	RB1-inducible coiled-coil protein
74	AT2G25200	hypothetical protein (DUF868)
74	AT2G40475	hypothetical protein
74	AT3G21890	B-box type zinc finger family protein
74	AT3G30460	RING/U-box superfamily protein
74	AT3G46870	Pentatricopeptide repeat (PPR) superfamily protein
74	AT4G34550	F-box protein
74	AT4G36670	Major facilitator superfamily protein
74	AT5G09230	sirtuin 2
74	AT5G21170	5'-AMP-activated protein kinase beta-2 subunit protein
75	AT1G16750	GPI-anchored adhesin-like protein, putative (Protein of unknown function, DUF547)
75	AT2G14247	Expressed protein
75	AT2G20725	CAAX amino terminal protease family protein
75	AT2G30630	P-loop containing nucleoside triphosphate hydrolases superfamily protein
75	AT3G05030	sodium hydrogen exchanger 2
75	AT3G27010	TEOSINTE BRANCHED 1, cycloidea, PCF (TCP)-domain family protein 20
75	AT4G19200	proline-rich family protein
75	AT4G24510	HXXXD-type acyl-transferase family protein
75	AT5G47070	Protein kinase superfamily protein
75	AT5G55530	Calcium-dependent lipid-binding (CaLB domain) family protein
75	AT5G58930	hypothetical protein (DUF740)
75	AT5G67290	FAD-dependent oxidoreductase family protein

75	AT5G67370	DUF1230 family protein (DUF1230)
76	AT1G13110	cytochrome P450, family 71 subfamily B, polypeptide 7
76	AT1G51700	DOF zinc finger protein 1
76	AT2G46735	death domain associated protein
76	AT3G08570	Phototropic-responsive NPH3 family protein
76	AT3G16280	Integrase-type DNA-binding superfamily protein
76	AT3G54390	sequence-specific DNA binding transcription factor
76	AT4G00050	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
76	AT5G10730	NAD(P)-binding Rossmann-fold superfamily protein
76	AT5G43330	Lactate/malate dehydrogenase family protein
76	AT5G43400	plant/protein
76	AT5G47610	RING/U-box superfamily protein
76	AT5G63160	BTB and TAZ domain protein 1
76	AT5G64410	oligopeptide transporter 4
77	AT1G12480	C4-dicarboxylate transporter/malic acid transport protein
77	AT1G61770	Chaperone DnaJ-domain superfamily protein
77	AT1G68470	Exostosin family protein
77	AT2G26860	FBD, F-box and Leucine Rich Repeat domains containing protein
77	AT2G35640	Homeodomain-like superfamily protein
77	AT3G06270	Protein phosphatase 2C family protein
77	AT3G26618	eukaryotic release factor 1-3
77	AT4G28740	LOW PSII ACCUMULATION-like protein
77	AT4G36050	endonuclease/exonuclease/phosphatase family protein
77	AT5G14550	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
77	AT5G45275	Major facilitator superfamily protein
78	AT1G16320	plant/protein (DUF2358)
78	AT1G19580	gamma carbonic anhydrase 1
78	AT1G72070	Chaperone DnaJ-domain superfamily protein
78	AT1G79160	filamentous hemagglutinin transporter
78	AT1G80490	TOPLESS-related 1
78	AT3G20130	cytochrome P450, family 705, subfamily A, polypeptide 22
78	AT4G11410	NAD(P)-binding Rossmann-fold superfamily protein
78	AT4G36830	GNS1/SUR4 membrane protein family
78	AT5G43880	methyl-coenzyme M reductase II subunit gamma, putative (DUF3741)
78	AT5G55250	IAA carboxylmethyltransferase 1
78	AT5G55960	transmembrane protein C9orf5 protein
79	AT1G68190	B-box zinc finger family protein
79	AT2G02980	Pentatricopeptide repeat (PPR) superfamily protein
79	AT2G18210	hypothetical protein
79	AT3G19170	presequence protease 1
79	AT3G23080	Polyketide cyclase/dehydrase and lipid transport superfamily protein
79	AT4G00300	receptor-like kinase
79	AT4G00710	BR-signaling kinase 3
79	AT5G02160	transmembrane protein
79	AT5G03550	MATH domain/coiled-coil protein
79	AT5G24490	30S ribosomal protein

80	AT1G13180	Actin-like ATPase superfamily protein
80	AT1G14890	Plant invertase/pectin methylesterase inhibitor superfamily protein
80	AT1G18210	Calcium-binding EF-hand family protein
80	AT1G58520	GDSL-like lipase/acylhydrolase superfamily protein
80	AT2G04850	Auxin-responsive family protein
80	AT2G25735	hypothetical protein
80	AT3G17800	alanine-tRNA ligase, putative (DUF760)
80	AT3G61220	NAD(P)-binding Rossmann-fold superfamily protein
80	AT4G27940	manganese tracking factor for mitochondrial SOD2
80	AT4G31875	hypothetical protein
80	AT5G47800	Phototropic-responsive NPH3 family protein
81	AT1G06800	alpha/beta-Hydrolases superfamily protein
81	AT1G07135	glycine-rich protein
81	AT1G12300	Tetratricopeptide repeat (TPR)-like superfamily protein
81	AT1G28760	guanosine-3',5'-bis (diphosphate) 3'-pyrophosphohydrolase, putative (DUF2215)
81	AT1G56660	MAEBL domain protein
81	AT1G77110	Auxin efflux carrier family protein
81	AT3G15095	Serine/Threonine-kinase pakA-like protein
81	AT4G08691	hypothetical protein
81	AT5G26600	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
81	AT5G54490	pinoid-binding protein 1
81	AT5G54530	serine protease, putative (Protein of unknown function, DUF538)
82	AT1G15170	MATE efflux family protein
82	AT1G17990	FMN-linked oxidoreductases superfamily protein
82	AT1G80980	stress response NST1-like protein
82	AT2G17360	Ribosomal protein S4 (RPS4A) family protein
82	AT2G22860	phytosulfokine 2 precursor
82	AT2G47270	transcription factor UPBEAT protein
82	AT2G47610	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
82	AT3G47380	Plant invertase/pectin methylesterase inhibitor superfamily protein
82	AT3G52340	sucrose-6F-phosphate phosphohydrolase 2
82	AT4G10380	NOD26-like intrinsic protein 5;1
82	AT4G30680	Initiation factor eIF-4 gamma, MA3
82	AT5G07250	RHOMBOID-like protein 3
82	AT5G10360	Ribosomal protein S6e
82	AT5G11740	arabinogalactan protein 15
82	AT5G58420	Ribosomal protein S4 (RPS4A) family protein
83	AT1G16090	wall associated kinase-like 7
83	AT2G05632	hypothetical protein
83	AT2G32950	Transducin/WD40 repeat-like superfamily protein
83	AT3G07380	glycosyltransferase family protein (DUF23)
83	AT3G13620	Amino acid permease family protein
83	AT4G10400	F-box/RNI-like/FBD-like domains-containing protein
83	AT4G19180	GDA1/CD39 nucleoside phosphatase family protein
83	AT4G20170	glycosyltransferase family protein (DUF23)

83	AT4G24040	trehalase 1
83	AT4G36530	alpha/beta-Hydrolases superfamily protein
83	AT5G08170	porphyromonas-type peptidyl-arginine deiminase family protein
83	AT5G20250	Raffinose synthase family protein
84	AT1G01740	kinase with tetratricopeptide repeat domain-containing protein
84	AT1G04360	RING/U-box superfamily protein
84	AT1G71480	Nuclear transport factor 2 (NTF2) family protein
84	AT1G78680	gamma-glutamyl hydrolase 2
84	AT3G10740	alpha-L-arabinofuranosidase 1
84	AT3G58070	C2H2 and C2HC zinc fingers superfamily protein
84	AT3G61160	Protein kinase superfamily protein
84	AT4G25300	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
84	AT5G02230	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
84	AT5G22530	hypothetical protein
85	AT1G20560	acyl activating enzyme 1
85	AT1G53380	hypothetical protein (DUF641)
85	AT1G78020	senescence-associated family protein, putative (DUF581)
85	AT2G22240	myo-inositol-1-phosphate synthase 2
85	AT2G30390	ferrochelatase 2
85	AT4G01070	UDP-Glycosyltransferase superfamily protein
85	AT4G28700	ammonium transporter 1;4
85	AT5G20990	molybdopterin biosynthesis CNX1 protein / molybdenum cofactor biosynthesis enzyme CNX1 (CNX1)
85	AT5G23170	Protein kinase superfamily protein
85	AT5G48150	GRAS family transcription factor
86	AT1G56600	galactinol synthase 2
86	AT1G72800	RNA-binding (RRM/RBD/RNP motifs) family protein
86	AT3G53400	peptide upstream protein
86	AT4G16000	hypothetical protein
86	AT5G03190	peptide upstream protein
86	AT5G17520	root cap 1 (RCP1)
87	AT1G09440	Protein kinase superfamily protein
87	AT1G65060	4-coumarate:CoA ligase 3
87	AT1G65560	Zinc-binding dehydrogenase family protein
87	AT3G06500	Plant neutral invertase family protein
87	AT3G12240	serine carboxypeptidase-like 15
87	AT4G33180	alpha/beta-Hydrolases superfamily protein
87	AT5G05270	Chalcone-flavanone isomerase family protein
87	AT5G08640	flavonol synthase 1
87	AT5G13930	Chalcone and stilbene synthase family protein
88	AT1G24807	Glutamine amidotransferase type 1 family protein
88	AT2G38240	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
88	AT3G23250	myb domain protein 15
88	AT3G28930	AIG2-like (avirulence induced gene) family protein
88	AT3G57020	Calcium-dependent phosphotriesterase superfamily protein
88	AT5G03406	Class II aaRS and biotin synthetases superfamily protein

88	AT5G05840	replication factor C subunit, putative (DUF620)
88	AT5G62130	Per1-like family protein
89	AT1G59500	Auxin-responsive GH3 family protein
89	AT2G18600	Ubiquitin-conjugating enzyme family protein
89	AT2G22080	transmembrane protein
89	AT3G22160	VQ motif-containing protein
89	AT4G13820	Leucine-rich repeat (LRR) family protein
89	AT4G16442	Uncharacterized protein family (UPF0497)
89	AT4G20780	calmodulin like 42
89	AT4G21180	DnaJ / Sec63 Brl domains-containing protein
89	AT4G31800	WRKY DNA-binding protein 18
89	AT4G38740	rotamase CYP 1
89	AT5G08430	SWIB/MDM2 and Plus-3 and GYF domain-containing protein
90	AT1G17190	glutathione S-transferase tau 26
90	AT1G51340	MATE efflux family protein
90	AT2G17695	outer envelope protein
90	AT2G18230	pyrophosphorylase 2
90	AT2G21960	transmembrane protein
90	AT2G35470	ribosome maturation factor
90	AT3G10940	dual specificity protein phosphatase (DsPTP1) family protein
90	AT3G24030	hydroxyethylthiazole kinase family protein
90	AT4G25830	Uncharacterized protein family (UPF0497)
90	AT5G05090	Homeodomain-like superfamily protein
91	AT1G28280	VQ motif-containing protein
91	AT1G77760	nitrate reductase 1
91	AT3G59000	F-box/RNI-like superfamily protein
91	AT4G09670	Oxidoreductase family protein
91	AT4G09900	methyl esterase 12
91	AT4G13100	RING/U-box superfamily protein
91	AT4G29120	6-phosphogluconate dehydrogenase family protein
91	AT5G24380	YELLOW STRIPE like 2
91	AT5G43950	vacuolar protein sorting-associated protein (DUF946)
92	AT1G25400	transmembrane protein
92	AT3G02250	O-fucosyltransferase family protein
92	AT3G50260	cooperatively regulated by ethylene and jasmonate 1
92	AT3G61750	Cytochrome b561/ferric reductase transmembrane with DOMON related domain-containing protein
92	AT4G17900	PLATZ transcription factor family protein
92	AT4G19880	Glutathione S-transferase family protein
92	AT5G02150	Fes1C
92	AT5G15740	O-fucosyltransferase family protein
92	AT5G22740	cellulose synthase-like A02
93	AT1G35420	alpha/beta-Hydrolases superfamily protein
93	AT1G70580	alanine-2-oxoglutarate aminotransferase 2
93	AT2G05380	glycine-rich protein 3 short isoform
93	AT2G05400	Ubiquitin-specific protease family C19-related protein
93	AT2G05530	Glycine-rich protein family

93	AT4G13430	isopropyl malate isomerase large subunit 1
93	AT4G27440	protochlorophyllide oxidoreductase B
93	AT5G10180	slufate transporter 2;1
93	AT5G23010	methylthioalkylmalate synthase 1
93	AT5G65480	hypothetical protein
94	AT1G08360	Ribosomal protein L1p/L10e family
94	AT1G11410	S-locus lectin protein kinase family protein
94	AT1G72370	40s ribosomal protein SA
94	AT2G34480	Ribosomal protein L18ae/LX family protein
94	AT2G41040	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
94	AT3G53130	Cytochrome P450 superfamily protein
94	AT3G60130	beta glucosidase 16
94	AT3G62870	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
94	AT5G37790	Protein kinase superfamily protein
94	AT5G46710	PLATZ transcription factor family protein
95	AT1G62560	flavin-monooxygenase glucosinolate S-oxygenase 3
95	AT1G64140	WRKY transcription factor
95	AT3G02020	aspartate kinase 3
95	AT3G22060	Receptor-like protein kinase-related family protein
95	AT5G49660	Leucine-rich repeat transmembrane protein kinase family protein
95	AT5G54390	HAL2-like protein
96	AT1G21460	Nodulin MtN3 family protein
96	AT1G33790	jacalin lectin family protein
96	AT1G52200	PLAC8 family protein
96	AT1G70000	myb-like transcription factor family protein
96	AT3G03860	APR-like 5
96	AT5G52900	membrane-associated kinase regulator
96	AT5G57100	Nucleotide/sugar transporter family protein
96	AT5G57660	CONSTANS-like 5
97	AT1G02660	alpha/beta-Hydrolases superfamily protein
97	AT2G23950	Leucine-rich repeat protein kinase family protein
97	AT2G27450	nitrilase-like protein 1
97	AT2G28120	Major facilitator superfamily protein
97	AT3G32980	Peroxidase superfamily protein
97	AT3G58750	citrate synthase 2
97	AT4G15610	Uncharacterized protein family (UPF0497)
97	AT5G16120	alpha/beta-Hydrolases superfamily protein
98	AT2G44940	Integrase-type DNA-binding superfamily protein
98	AT3G17770	Dihydroxyacetone kinase
98	AT3G18980	EIN2 targeting protein1
98	AT3G19370	filament-like protein (DUF869)
98	AT4G23020	hypothetical protein
98	AT5G03760	Nucleotide-diphospho-sugar transferases superfamily protein
98	AT5G20220	zinc knuckle (CCHC-type) family protein
98	AT5G41460	transferring glycosyl group transferase (DUF604)
99	AT1G27770	autoinhibited Ca2 -ATPase 1

99	AT2G37130	Peroxidase superfamily protein
99	AT3G61210	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
99	AT4G23750	cytokinin response factor 2
99	AT5G11810	rhomboid family protein
99	AT5G17170	rubredoxin family protein
99	AT5G49330	myb domain protein 111
100	AT1G23740	Oxidoreductase, zinc-binding dehydrogenase family protein
100	AT1G33230	TMPIT-like protein
100	AT1G61100	disease resistance protein (TIR class)
100	AT2G44160	methylenetetrahydrofolate reductase 2
100	AT3G04940	cysteine synthase D1
100	AT3G53920	RNApolymerase sigma-subunit C
100	AT4G11450	bromo-adjacent domain protein, putative (DUF3527)
100	AT5G13460	IQ-domain 11
101	AT1G21440	Phosphoenolpyruvate carboxylase family protein
101	AT1G65860	flavin-monooxygenase glucosinolate S-oxygenase 1
101	AT2G45220	Plant invertase/pectin methylesterase inhibitor superfamily
101	AT3G10815	RING/U-box superfamily protein
101	AT3G28130	nodulin MtN21 /EamA-like transporter family protein
101	AT3G53720	cation/H exchanger 20
101	AT3G53800	Fes1B
101	AT4G12490	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
101	AT4G22290	Ubiquitin-specific protease family C19-related protein
101	AT4G30210	P450 reductase 2
101	AT5G15450	casein lytic proteinase B3
102	AT2G21560	nucleolar-like protein
102	AT2G24270	aldehyde dehydrogenase 11A3
102	AT2G24540	Galactose oxidase/kelch repeat superfamily protein
102	AT2G34460	NAD(P)-binding Rossmann-fold superfamily protein
102	AT3G03980	NAD(P)-binding Rossmann-fold superfamily protein
102	AT3G04030	Homeodomain-like superfamily protein
102	AT3G12600	nudix hydrolase homolog 16
102	AT3G63420	Ggamma-subunit 1
102	AT4G32600	C3H4 type zinc finger protein
102	AT4G38960	B-box type zinc finger family protein
102	AT5G24120	sigma factor E
103	AT1G15960	NRAMP metal ion transporter 6
103	AT1G16400	cytochrome P450, family 79, subfamily F, polypeptide 2
103	AT1G16410	cytochrome p450 79f1
103	AT1G18590	sulfotransferase 17
103	AT3G13110	serine acetyltransferase 2;2
103	AT3G56000	cellulose synthase like A14
103	AT4G12030	bile acid transporter 5
103	AT4G14040	selenium-binding protein 2
103	AT4G39940	APS-kinase 2

103	AT5G62520	similar to RCD one 5
104	AT1G09940	Glutamyl-tRNA reductase family protein
104	AT1G16060	ARIA-interacting double AP2 domain protein
104	AT1G62800	aspartate aminotransferase 4
104	AT3G03190	glutathione S-transferase F11
104	AT3G54600	Class I glutamine amidotransferase-like superfamily protein
104	AT3G58990	isopropylmalate isomerase 1
104	AT5G33320	Glucose-6-phosphate/phosphate translocator-like protein
104	AT5G44380	FAD-binding Berberine family protein
104	AT5G44720	Molybdenum cofactor sulfurase family protein
105	AT1G51940	protein kinase family protein / peptidoglycan-binding LysM domain-containing protein
105	AT1G64780	ammonium transporter 1;2
105	AT1G80190	partner of SLD five 1
105	AT2G15890	maternal effect embryo arrest 14
105	AT2G40460	Major facilitator superfamily protein
105	AT3G04140	Ankyrin repeat family protein
105	AT3G27350	transcriptional regulator ATRX-like protein
105	AT4G18390	TEOSINTE BRANCHED 1, cycloidea and PCF transcription factor 2
106	AT1G12440	A20/AN1-like zinc finger family protein
106	AT1G13700	6-phosphogluconolactonase 1
106	AT2G18290	anaphase promoting complex 10
106	AT3G15450	aluminum induced protein with YGL and LRDR motifs
106	AT3G15630	plant/protein
106	AT3G48590	nuclear factor Y, subunit C1
106	AT4G29130	hexokinase 1
106	AT5G11070	hypothetical protein
106	AT5G22920	CHY-type/CTCHY-type/RING-type Zinc finger protein
107	AT1G27100	Actin cross-linking protein
107	AT2G45170	AUTOPHAGY 8E
107	AT3G03250	UDP-GLUCOSE PYROPHOSPHORYLASE 1
107	AT4G17880	Basic helix-loop-helix (bHLH) DNA-binding family protein
107	AT4G27450	aluminum induced protein with YGL and LRDR motifs
107	AT5G10695	methionyl-tRNA synthetase
108	AT2G14620	xyloglucan endotransglucosylase/hydrolase 10
108	AT2G33310	auxin-induced protein 13
108	AT3G19620	Glycosyl hydrolase family protein
108	AT3G53970	proteasome inhibitor-like protein
108	AT4G31310	AIG2-like (avirulence induced gene) family protein
108	AT5G16360	NC domain-containing protein-like protein
108	AT5G20500	Glutaredoxin family protein
108	AT5G27760	Hypoxia-responsive family protein
108	AT5G57510	cotton fiber protein
109	AT1G78100	F-box family protein
109	AT2G37580	RING/U-box superfamily protein
109	AT2G40316	autophagy-like protein
109	AT4G02630	Protein kinase superfamily protein

109	AT4G21970	senescence regulator (Protein of unknown function, DUF584)
109	AT5G05830	RING/FYVE/PHD zinc finger superfamily protein
109	AT5G25930	kinase family with leucine-rich repeat domain-containing protein
109	AT5G50800	Nodulin MtN3 family protein
110	AT1G04985	triacylglycerol lipase-like protein
110	AT1G48280	hydroxyproline-rich glycoprotein family protein
110	AT1G66580	senescence associated gene 24
110	AT2G36970	UDP-Glycosyltransferase superfamily protein
110	AT3G10760	Homeodomain-like superfamily protein
110	AT3G51660	Tautomerase/MIF superfamily protein
110	AT4G01895	systemic acquired resistance (SAR) regulator protein NIMIN-1-like protein
110	AT4G20090	Pentatricopeptide repeat (PPR) superfamily protein
110	AT5G23760	Copper transport protein family
111	AT1G29120	Hydrolase-like protein family
111	AT1G60710	NAD(P)-linked oxidoreductase superfamily protein
111	AT1G73660	protein tyrosine kinase family protein
111	AT3G04590	AT hook motif DNA-binding family protein
111	AT3G09270	glutathione S-transferase TAU 8
111	AT3G14067	Subtilase family protein
111	AT3G46900	copper transporter 2
111	AT3G51895	sulfate transporter 3;1
111	AT5G58870	FTSH protease 9
112	AT1G05620	uridine-ribohydrolase 2
112	AT1G64900	cytochrome P450, family 89, subfamily A, polypeptide 2
112	AT2G15020	hypothetical protein
112	AT5G03555	permease, cytosine/purines, uracil, thiamine, allantoin family protein
112	AT5G24150	FAD/NAD(P)-binding oxidoreductase family protein
112	AT5G41900	alpha/beta-Hydrolases superfamily protein
112	AT5G58770	Undecaprenyl pyrophosphate synthetase family protein
113	AT1G09560	germin-like protein 5
113	AT1G17860	Kunitz family trypsin and protease inhibitor protein
113	AT1G32940	Subtilase family protein
113	AT1G74790	catalytics
113	AT4G08300	nodulin MtN21 /EamA-like transporter family protein
113	AT4G34131	UDP-glucosyl transferase 73B3
113	AT4G39670	Glycolipid transfer protein (GLTP) family protein
114	AT1G72040	P-loop containing nucleoside triphosphate hydrolases superfamily protein
114	AT3G06200	P-loop containing nucleoside triphosphate hydrolases superfamily protein
114	AT3G44190	FAD/NAD(P)-binding oxidoreductase family protein
114	AT3G48250	Pentatricopeptide repeat (PPR) superfamily protein
114	AT4G32520	serine hydroxymethyltransferase 3
114	AT4G34910	P-loop containing nucleoside triphosphate hydrolases superfamily protein
115	AT1G04100	indoleacetic acid-induced protein 10
115	AT1G78850	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain-containing protein

115	AT2G40900	nodulin MtN21 /EamA-like transporter family protein
115	AT4G33420	Peroxidase superfamily protein
115	AT4G34540	NmrA-like negative transcriptional regulator family protein
115	AT4G35900	Basic-leucine zipper (bZIP) transcription factor family protein
115	AT5G02970	alpha/beta-Hydrolases superfamily protein
116	AT1G11090	alpha/beta-Hydrolases superfamily protein
116	AT2G27500	Glycosyl hydrolase superfamily protein
116	AT2G29680	cell division control 6
116	AT3G12750	zinc transporter 1 precursor
116	AT3G15090	GroES-like zinc-binding alcohol dehydrogenase family protein
116	AT3G56170	Ca-2 dependent nuclease
116	AT4G15760	monooxygenase 1
116	AT5G02590	Tetratricopeptide repeat (TPR)-like superfamily protein
116	AT5G16530	Auxin efflux carrier family protein
117	AT1G10020	formin-like protein (DUF1005)
117	AT1G10970	zinc transporter
117	AT1G11700	senescence regulator (Protein of unknown function, DUF584)
117	AT3G47340	glutamine-dependent asparagine synthase 1
117	AT3G58720	RING/U-box superfamily protein
117	AT4G05030	Copper transport protein family
117	AT5G40630	Ubiquitin-like superfamily protein
118	AT1G09310	plant/protein (Protein of unknown function, DUF538)
118	AT1G75280	NmrA-like negative transcriptional regulator family protein
118	AT2G15440	polysaccharide biosynthesis protein (DUF579)
118	AT3G18850	lysophosphatidyl acyltransferase 5
118	AT4G00360	cytochrome P450, family 86, subfamily A, polypeptide 2
119	AT1G22770	gigantea protein (GI)
119	AT1G77260	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
119	AT3G01470	homeobox 1
119	AT3G24190	Protein kinase superfamily protein
119	AT5G41400	RING/U-box superfamily protein
119	AT5G57220	cytochrome P450, family 81, subfamily F, polypeptide 2
120	AT1G03220	Eukaryotic aspartyl protease family protein
120	AT1G56140	Leucine-rich repeat transmembrane protein kinase
120	AT1G78580	trehalose-6-phosphate synthase
120	AT2G17660	RPM1-interacting protein 4 (RIN4) family protein
120	AT4G32480	sugar phosphate exchanger, putative (DUF506)
120	AT4G33640	costars family protein
120	AT5G54500	flavodoxin-like quinone reductase 1
121	AT1G25550	myb-like transcription factor family protein
121	AT1G26380	FAD-binding Berberine family protein
121	AT1G36622	transmembrane protein
121	AT2G20570	GBF's pro-rich region-interacting factor 1
121	AT3G43600	aldehyde oxidase 2
121	AT4G02500	UDP-xylosyltransferase 2
121	AT5G01090	Concanavalin A-like lectin family protein

121	AT5G67140	F-box/RNI-like superfamily protein
122	AT1G49500	transcription initiation factor TFIID subunit 1b-like protein
122	AT1G67980	caffeoyl-CoA 3-O-methyltransferase
122	AT2G22980	serine carboxypeptidase-like 13
122	AT2G31070	TCP domain protein 10
122	AT3G07350	sulfate/thiosulfate import ATP-binding protein, putative (DUF506)
122	AT4G02840	Small nuclear ribonucleoprotein family protein
122	AT4G15820	ABC subfamily C protein
122	AT4G38210	expansin A20
123	AT1G22890	transmembrane protein
123	AT2G39360	Protein kinase superfamily protein
123	AT2G39930	isoamylase 1
123	AT3G61250	myb domain protein 17
123	AT4G01450	nodulin MtN21 /EamA-like transporter family protein
123	AT4G25835	P-loop containing nucleoside triphosphate hydrolases superfamily protein
123	AT4G35985	Senescence/dehydration-associated protein-like protein
123	AT5G14970	seed maturation-like protein
124	AT1G62790	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
124	AT2G42200	squamosa promoter binding protein-like 9
124	AT2G47510	fumarase 1
124	AT3G45060	high affinity nitrate transporter 2.6
124	AT4G22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
124	AT5G20230	blue-copper-binding protein
125	AT2G04650	ADP-glucose pyrophosphorylase family protein
125	AT2G06960	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
125	AT2G20240	GPI-anchored adhesin-like protein, putative (DUF3741)
125	AT2G21320	B-box zinc finger family protein
125	AT4G17090	chloroplast beta-amylase
125	AT4G25810	xyloglucan endotransglycosylase 6
125	AT4G39660	alanine:glyoxylate aminotransferase 2
126	AT1G03230	Eukaryotic aspartyl protease family protein
126	AT2G18700	trehalose phosphatase/synthase 11
126	AT2G43820	UDP-glucosyltransferase 74F2
126	AT4G37520	Peroxidase superfamily protein
127	AT2G15090	3-ketoacyl-CoA synthase 8
127	AT4G13440	Calcium-binding EF-hand family protein
127	AT4G33010	glycine decarboxylase P-protein 1
127	AT4G35090	catalase 2
127	AT5G02600	Heavy metal transport/detoxification superfamily protein
127	AT5G14230	ankyrin
128	AT1G01480	1-amino-cyclopropane-1-carboxylate synthase 2
128	AT1G57990	purine permease 18
128	AT2G35290	hypothetical protein
128	AT3G18390	CRS1 / YhbY (CRM) domain-containing protein
128	AT3G28210	zinc finger (AN1-like) family protein

128	AT4G01360	BPS1-like protein
129	AT1G06135	transmembrane protein
129	AT1G15060	alpha/beta hydrolase family protein
129	AT4G14630	germin-like protein 9
129	AT5G16600	myb domain protein 43
129	AT5G64170	dentin sialophosphoprotein-like protein
130	AT1G05460	P-loop containing nucleoside triphosphate hydrolases superfamily protein
130	AT1G19610	defensin-like protein
130	AT1G32170	xyloglucan endotransglucosylase/hydrolase 30
130	AT2G14520	CBS domain protein (DUF21)
130	AT2G29290	NAD(P)-binding Rossmann-fold superfamily protein
130	AT3G45970	expansin-like A1
131	AT2G02950	phytochrome kinase substrate 1
131	AT4G05330	ARF-GAP domain 13
131	AT5G45750	RAB GTPase homolog A1C
131	AT5G51910	TCP family transcription factor
132	AT1G12560	expansin A7
132	AT1G73600	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
132	AT2G17265	homoserine kinase
132	AT3G25180	cytochrome P450, family 82, subfamily G, polypeptide 1
132	AT4G29680	Alkaline-phosphatase-like family protein
132	AT5G11200	DEAD/DEAH box RNA helicase family protein
133	AT1G06040	B-box zinc finger family protein
133	AT1G22900	Disease resistance-responsive (dirigent-like protein) family protein
133	AT2G24190	NAD(P)-binding Rossmann-fold superfamily protein
133	AT2G32295	EXS (ERD1/XPR1/SYG1) family protein
133	AT3G21520	transmembrane protein, putative (DUF679 domain membrane protein 1)
133	AT4G20070	allantoate amidohydrolase
134	AT1G62180	5'adenylylphosphosulfate reductase 2
134	AT3G15210	ethylene responsive element binding factor 4
134	AT4G16060	hypothetical protein
134	AT5G35970	P-loop containing nucleoside triphosphate hydrolases superfamily protein
135	AT1G32960	Subtilase family protein
135	AT1G60730	NAD(P)-linked oxidoreductase superfamily protein
135	AT4G20830	FAD-binding Berberine family protein
135	AT4G37530	Peroxidase superfamily protein
135	AT5G04020	calmodulin binding protein
135	AT5G57010	calmodulin-binding family protein
136	AT1G26400	FAD-binding Berberine family protein
136	AT2G35980	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
136	AT3G13450	Transketolase family protein
136	AT3G15770	hypothetical protein
136	AT5G61420	myb domain protein 28
137	AT1G19330	histone deacetylase complex subunit

137	AT1G66840	PLASTID MOVEMENT IMPAIRED protein (DUF827)
137	AT1G76650	calmodulin-like 38
137	AT3G55970	jasmonate-regulated gene 21
137	AT4G37290	transmembrane protein
137	AT5G01100	O-fucosyltransferase family protein
138	AT1G18440	Peptidyl-tRNA hydrolase family protein
138	AT2G44195	pre-mRNA splicing factor domain-containing protein
138	AT3G01180	starch synthase 2
138	AT3G01830	Calcium-binding EF-hand family protein
138	AT5G07690	myb domain protein 29
139	AT2G23180	cytochrome P450, family 96, subfamily A, polypeptide 1
139	AT2G29460	glutathione S-transferase tau 4
139	AT3G22942	G-protein gamma subunit 2
139	AT3G25250	AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kinase family protein
139	AT4G17680	SBP (S-ribonuclease binding protein) family protein
140	AT2G37240	Thioredoxin superfamily protein
140	AT3G07330	Cellulose-synthase-like C6
140	AT4G13400	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
140	AT4G14640	calmodulin 8
141	AT2G41890	curculin-like (mannose-binding) lectin family protein / PAN domain-containing protein
141	AT2G46640	NAD-dependent protein deacetylase HST1-like protein
141	AT3G09770	RING/U-box superfamily protein
141	AT5G20700	senescence-associated family protein, putative (DUF581)
142	AT1G73590	Auxin efflux carrier family protein
142	AT1G78590	NAD(H) kinase 3
142	AT2G41560	autoinhibited Ca(2)-ATPase, isoform 4
142	AT3G52720	alpha carbonic anhydrase 1
143	AT1G18300	nudix hydrolase homolog 4
143	AT1G24520	homolog of Brassica campestris pollen protein 1
143	AT3G25585	aminoalcoholphosphotransferase
143	AT4G09180	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
143	AT4G17350	auxin canalization protein (DUF828)
144	AT1G19630	cytochrome P450, family 722, subfamily A, polypeptide 1
144	AT1G30370	alpha/beta-Hydrolases superfamily protein
144	AT1G63030	Integrase-type DNA-binding superfamily protein
144	AT3G11800	Expp1 protein
144	AT3G25780	allene oxide cyclase 3
145	AT4G12130	Glycine cleavage T-protein family
145	AT4G15810	P-loop containing nucleoside triphosphate hydrolases superfamily protein
145	AT4G26400	RING/U-box superfamily protein
145	AT5G04110	DNA GYRASE B3
146	AT3G21750	UDP-glucosyl transferase 71B1
146	AT3G21760	UDP-Glycosyltransferase superfamily protein
146	AT4G37000	accelerated cell death 2 (ACD2)
146	AT5G16170	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family

		protein
147	AT1G17340	Phosphoinositide phosphatase family protein
147	AT2G39450	Cation efflux family protein
147	AT4G23496	SPIRAL1-like5
148	AT1G71696	carboxypeptidase D
148	AT3G19710	branched-chain aminotransferase4
148	AT3G60530	GATA transcription factor 4
149	AT1G35560	TCP family transcription factor
149	AT1G60960	iron regulated transporter 3
149	AT2G31790	UDP-Glycosyltransferase superfamily protein
149	AT3G03780	methionine synthase 2
150	AT3G04310	transmembrane protein
150	AT3G61190	BON association protein 1
150	AT4G30700	Pentatricopeptide repeat (PPR) superfamily protein
150	AT5G53030	hypothetical protein
151	AT1G28480	Thioredoxin superfamily protein
151	AT1G56060	cysteine-rich/transmembrane domain protein B
151	AT3G46070	C2H2-type zinc finger family protein
152	AT1G06550	ATP-dependent caseinolytic (Clp) protease/crotonase family protein
152	AT5G08415	Radical SAM superfamily protein
152	AT5G57300	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
153	AT1G14205	Ribosomal L18p/L5e family protein

The list of lowering time genes extracted from the list above – the genes highlighted in bold are those that are seen in the GRN of flowering genes.

AT1G01480	**AT1G63030**	**AT3G05120**	AT4G03210	**AT5G39860**
AT1G04390	**AT1G65060**	**AT3G18520**	**AT4G04760**	**AT5G40630**
AT1G08230	**AT1G66400**	**AT3G18990**	AT4G14640	**AT5G49730**
AT1G12610	AT1G70560	AT3G19170	**AT4G15900**	**AT5G49740**
AT1G12820	AT1G73250	AT3G22160	**AT4G28700**	**AT5G54840**
AT1G13260	**AT1G73660**	AT3G22370	**AT4G35900**	AT5G55950
AT1G16750	**AT2G19670**	**AT3G22380**	**AT4G36050**	AT5G58930
AT1G17680	**AT2G22540**	**AT3G24130**	AT4G36930	**AT5G62000**
AT1G19190	AT2G22930	AT3G25730	AT4G39210	**AT5G63960**
AT1G22770	AT2G23950	AT3G46070	**AT5G01600**	
AT1G23020	AT2G32950	AT3G46820	**AT5G04000**	
AT1G23860	**AT2G34060**	**AT3G51630**	**AT5G06850**	
AT1G27460	**AT2G34850**	AT3G54560	AT5G08020	
AT1G35720	**AT2G37340**	AT3G58070	**AT5G19430**	
AT1G36150	AT2G41100	AT3G59420	**AT5G20320**	
AT1G51140	**AT2G41250**	AT3G61700	AT5G22290	
AT1G51200	AT2G41890	AT4G00360	**AT5G25220**	
AT1G52310	AT2G47890	AT4G01060	**AT5G27320**	

Appendix B

The list of T-DNA insertional lines obtained from NASC and screened to identify loss-of-function mutants.

Line name	Gene Locus Identified	Mutant line	NASC No.
rap2.12-1	At1g53910	GK-503A11	N448203
rap2.12-2	At1g53910	SAIL_1215_H10	N878557
rap2.12-3	At1g53910	GK-137C12	N413092
rap2.12-4	At1g53910	SALK_019187	N519187
rap2.12-5	At1g53910	SALK_070755	N570755
rap2.12-6	At1g53910	SALK_106201	N606201
rap2.12-7	At1g53910	SALK_106203	N606203
rap2.12-8	At1g53910	SALK_152421	N652421
uktf-1	At1g16750	SALK_061668	N561668
uktf-2	At1g16750	SALK_089535	N589535
uktf-3	At1g16750	SALK_144830C	N658962
uktf-4	At1g16750	SAIL_327_D01	N815201
uktf-5	At1g16750	SAIL_744_D11	N833296
uktf-6	At1g16750	SAIL_802_F11	N863293
fd-2	At4g35900	SALK_013288	N513288
fd-3	At4g35900	SALK_054421c	N678498
fd-4	At4g35900	SALK_118487c	N676655
fd-5	At4g35900	SALK_150991	N650991
fd-6	At4g35900	SALK_013272	N513272
poz-1	At1g55760	SALK_038930c	N654605
poz-2	At1g55760	SALK_127778c	N676860
poz-3	At1g55760	SALK_075267c	N680500
poz-4	At1g55760	SAIL_672_G04	N862962
agl22-1	At2g22540	SALK_141694	N641694
agl22-2	At2g22540	SALK_141683	N641683
agl22-3	At2g22540	SALK_141674	N641674
agl22-4	At2g22540	SAIL_583_C08	N824840
agl22-5	At2g22540	SALK_141675	N641675
svp-31	At2g22540	SALK_026551C	N660785
svp-32	At2g22540	SALK_072930C	N666411
bhlh038-1	At3g56970	SALK_020183c	N664471
bhlh038-2	At3g56970	SALK_108159c	N655819
bhlh038-3	At3g56970	SAIL_106_F03	N805131
bhlh038-4	At3g56970	SAIL_447_H01	N874278
bhlh038-5	At3g56970	GK-047G03	N404491
anl2-2	At4g00730	SALK_000196c	N661242
anl2-3	At4g00730	SALK_007529	N507529
anl2-4	At4g00730	SAIL_418_C10	N819318
anl2-5	At4g00730	SAIL_604_C01	N825790
anl2-6	At4g00730	SALK_103392	N603392
anl2-7	At4g00730	SAIL_1158_C12	N842731

Appendix C

Primers for screening T-DNA insertional mutants

rap2.12-1	scr_GK1_RAP2.12_F	TGGCTACTCCTGAATGCAAAC
	scr_GK1_RAP2.12_R	CTCAGCTGTCTTGAACGTTCC
rap2.12-3	scr_GK2_RAP2.12_F	AAAGTTACTGGCTTGGATGGG
	scr_GK2_RAP2.12_R	AATTTACGAGAACCGGTTTGG
UKTF-3, -4	UKTF_F_screen	CTTTGGCTCGCTCCACTTATCTGA
	UKTF_R_screen	TCTCTTGTTCGAGCTTAGAGACCGTA
fd-5	FD_F_screen	AGTCACACATGAAAGTTGAAAGTGA
	FD R stop	TCAAAATGGAGCTGTGG
poz-2, -3	POZ F start	caccATGACTGATTCTGCTTACAGAGTAG
	POZ R stop	CTAAAATCCTTTCCAGGTACTG
agl22-3	AGL22 F start	caccATGGCGAGAGAAAAGATTC
	AGL22 R stop	CTAACCACCATACGGTAAGCTGCACAACAC
agl22-4	scr3_AGL22_F	TCAGCGAACTTCAGAAAAAGG
	AGL22 R stop	CTAACCACCATACGGTAAGCTGCACAACAC
bhlh038-2	BHLH038 F start	caccATGTGTGCATTAGTCCCTTC
	scr_BHLH038_R	ATCAGCCCCTCGTCTCTCTAG
bhlh038-4	BHLH038 F start	caccATGTGTGCATTAGTCCCTTC
	BHLH038 R stop	CTAGTTAAACGAGTTTTCACATTTC
anl2-2	ANL2 F start	caccATGAACTTCGGTAGTCTCTTCG
	ANL2 R stop	TCAGCTTTCGCATTGTAAAG
anl2-4	scr2_ANL2_SAIL_F	GCATCCAGACGAGAAACAAAG
	scr2_ANL2_SAIL_R	GCAGAACCGGAGGAAATTAAC

Primers for Gateway® cloning and RT-PCR

RAP2.12	RAP F cloning	caccATGTGTGGAGGAGCTATAATA
	RAP R stop	TCAGAAGACTCCTCCAATC
UKTF	UKTF F start	caccATGTCTGGAGACTCTCTGCT
	UKTF R stop	CTACGGAAGAAACAAATATCG
FD	FD F start	caccATGTTGTCATCAGCTAAGCAT
	FD R stop	TCAAAATGGAGCTGTGG
POZ	POZ F start	caccATGACTGATTCTGCTTACAGAGTAG
	POZ R stop	CTAAAATCCTTTCCAGGTACTG
AGL22	AGL22 F start	caccATGGCGAGAGAAAAGATTC
	AGL22 R stop	CTAACCACCATACGGTAAGCTGCACAACAC
BHLH038	BHLH038 F start	caccATGTGTGCATTAGTCCCTTC
	BHLH038 R stop	CTAGTTAAACGAGTTTTCACATTTC
ANL2	ANL2 F start	caccATGAACTTCGGTAGTCTCTTCG
	ANL2 R stop	TCAGCTTTCGCATTGTAAAG

Primers to verify overexpressors using qPCR

rap2.12-1	clo_RAP2.12_F_start	ATGTGTGGAGGAGCTATAATATCCGA
	qPCR_RAP_5_R	ATTCTTCTTCAGATCCGGCCAAA
rap2.12-3	qPCR_RAP_6_F	CAGCTCTGACCAGGGTAGTAATTC
	qPCR_RAP_6_R	CGCAGAAGAGATGTCGGGAG
UKTF-3, -4	qPCR_UK_F	CGGGATGCACTCATCTCCTG
	qPCR_UK_R	CTTGGGCTCACTCCCACTGA
fd-5	qPCR_FD_F_II	ACCTTGCTTCCATCCACCAC
	qPCR_FD_R_II	TTTTGGCCCCTGAACCTTGG
poz-2, -3	qPCR_POZ_2,3_F	ACTGGGAAAATCATCATCGACGT
	qPCR_POZ_2,3_R	ACCGTTGGCCCAGATAGAGT
agl22-3	qPCR_AGL22_3_F	CTCTCCGTTCTCTGCGACG
	qPCR_AGL22_3_R	GGGCGTGATCACTGTTCTCA
agl22-4	qPCR_AGL22_4_F	CGGAAGAGAACGAGCGACTT
	qPCR_AGL22_4_R	TTCTCCGATTCAGCACCACC
bhlh038-2, -4	qPCR_BHLH038_1,4_F	AAGGCGGTCGCGAGTTATCTCTCA
	qPCR_BHLH038_1,4_R	TGCAAAGTGTAGAAGAGCCTCTC
anl2-2	qPCR_ANL2_1_F	CCGGTGAAGATCAAGACGCT
	qPCR_ANL2_1_R	GTTTCTCGTCTGGATGCGGA
anl2-4	qPCR_ANL2_4_F	GCTTGGACGGTGAGAGAGAC
	qPCR_ANL2_4_R	CTTCAGTAGCTAAACCGGTTGGT

Primers to test network connections using qPCR

qPCR_RSZ21_F	GCTCTGTGAGGCGTAGAAGC
qPCR_RSZ21_R	AGTAACACTGCGACGCCTTG
qPCR_PRL1_F	CGGAAGCCAAGCAGATTCGT
qPCR_PRL1_R	AGGTCCTGGCTGCTCATTGA
qPCR_DCL4_F	GACTTTGACCACGCCGATGT
qPCR_DCL4_R	CAGCCGCGGAGAACAAAGAA
qPCR_PROXD_F	TCGTCGGTTCCGTCACTTCT
qPCR_PROXD_R	ACCCATCACAACCCTCAACG
qPCR_FBH3_F	AGACGTCCTCCTTTGGCACA
qPCR_FBH3_R	CCCGCTTGGCTCTGATCTTAC
qPCR_NUC-PHOS_F	CCAACAAGGTGGAAAGGTGGA
qPCR_NUC-PHOS_R	GCTGTGTTCCGGGATGTCTG
qPCR_DREB4A_chr1_F	GCCAACGCCGTGAAGATCAT
qPCR_DREB4A_chr1_R	CCACCCACCTCCACTCATCA
qPCR_ARF2_F	CCGCTTGTGACGGTTCCTAG
qPCR_ARF2_R	ATCTGTTGTTCTGCCGCCTG
qPCR_DREB4A_chr2_F	TCAAGAATCTGGCTCGGGACT
qPCR_DREB4A_chr2_R	CGCCGGCTAACTTGGGAAAA
qPCR_DREB1A_F	TGAGATGTGTGATGCGACGAC
qPCR_DREB1A_R	AAACATCGCCTCATCGTGCA

qPCR_WRKY20_F	CTGGCCGAGATGAGAAGGGA
qPCR_WRKY20_R	CGCTTCTCCACCATCGTCAG
qPCR_PIL2_F	TTGGGGCCGAACTACTCTCA
qPCR_PIL2_R	GCAAGCCAGCTCCTAGAACA
qPCR_BZO2H1_F	CTGCTCCCATGACGACGAAG
qPCR_BZO2H1_R	TCACAGACCCAACCCGAAGA
qPCR_XPB1_F	GGCTCTACAAAAGGCGAGACG
qPCR_XPB1_R	TGCTGACCGGATTTGTGACG
qPCR_BPC7_F	ACCCACTACCCATGAGCACA
qPCR_BPC7_R	GTCATAGCCTTCGTCCGCAA
qPCR_WNK5_F	GGACACGAGAACACAGGTGC
qPCR_WNK5_R	GTGGACGACGGTTGAGGTTG
qPCR_GID1C_F	GGCCCGAGAAGCAAGAGTCT
qPCR_GID1C_R	TTCTTGAGCCCTTCCGCGTA
qPCR_SIS8_F	GTTGGGGCAGTTGGGTTTCA
qPCR_SIS8_R	TTTGAATCCGTCTGCCAGCA
qPCR_NF-YA10_F	TCGCCACCACAGACTCCTT
qPCR_NF-YA10_R	TGCCTCCGTAGTAACTGCCT
qPCR_DEL2_F	TCTCGCTCCCCAGGTTTACA
qPCR_DEL2_R	CGGCATCATCGAGCCCAAAT