UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

School of Biological Sciences

The role of PHYTOCHROME-INTERACTING FACTOR 3 in regulating growth and development in hexaploid wheat

Benjamin Sibbett

Doctor of Philosophy

October 2018

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

School of Biological Sciences

Doctor of Philosophy

THE ROLE OF PHYTOCHROME-INTERACTING FACTOR 3 IN REGULATING GROWTH AND DEVELOPMENT IN HEXAPLOID WHEAT

Benjamin Sibbett

The development of chloroplasts requires tight control to prevent the accumulation of reactive oxygen species and seedling damage. Work with Arabidopsis indicates that gibberellin signalling plays an important role in photo-oxidative stress responses through targeted degradation of growth repressing DELLA proteins. In wheat, mutants of the DELLA protein RHT-1 were the basis of the semi-dwarf lines so crucial to the Green Revolution. These mutants, however, display adverse pleiotropic effects impacting on wheat physiology. Elucidating the downstream targets of RHT-1 could therefore provide a mechanism for targeting specific RHT-1 functions without off-target consequences. In Arabidopsis, one mechanism by which DELLAs function is through repressing the activity of the Phytochrome Interacting Factors (PIFs). PIFs are transcription factors that repress light responses in plants, leading to the promotion of growth and repression of chloroplast development.

To determine the role of PIFs in regulating growth and chloroplast development in wheat we have searched the bread wheat genome for orthologues of known Arabidopsis and rice *PIF* genes. Annotation of wheat *PIF* gene sequences has been achieved through analysis of RNA-seq data and a specific *PIF3* orthologue has been identified. A range of approaches have been used to manipulate TaPIF3 levels in order to study its function. Potential loss-of-function mutations in each of the three *TaPIF3* homoeologues have been identified by TILLING and stacked to generate a *TaPIF3* mutant. A wheat *TaPIF3* overexpression line has also been generated. Phenotypic analysis of the *TaPIF3* triple mutant and overexpression lines indicate a role for TaPIF3 in the regulation of stem elongation and ear length. These results suggest that TaPIF3 is a promising target for downstream regulation of wheat physiology. Furthermore, a heterologous complementation approach has been used to further characterise the function of TaPIF3.

Contents

List of figures

UNIVERSITY OF Southampton

Research Thesis: Declaration of Authorship

Print name:	
Title of thesis:	

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as: [please list references below]:

_____
_____

Signature:		Date:	

ACKNOWLEDGEMENTS:

Firstly, I would like to thank my supervisors, Matthew Terry, Steve Thomas, Lorraine Williams and Peter Hedden for their expert guidance throughout this project. You have all supported me immensely and have always been available when I have needed assistance. This project would not have been as interesting or enjoyable without your continued support

Secondly, I would like to thank all of the past and present members of the lab groups at the University of Southampton and Rothamsted Research. Both groups have made me feel very welcome and have provided much needed help and advice. I have made some good friends working with you.

Last but not least, I would like to thank my family and friends for all of their understanding and support throughout my studies. You have been there when I have needed motivation but perhaps more importantly when I have needed a laugh at the end of the week. I would like to say an especial thank you to my fiancé Sarah, who has endured me through stressful times and has kept a smile on my face.

This work was funded by the Doctoral Training Partnership (DTP) BBSRC studentship programme.

ABBREVIATIONS

3-AT	3-Amino-1,2,4-triazole
ABA	Abscisic acid
AD	Activation domain
ALA	δ-aminolevulinic acid
ANOVA	Analysis of variance
APB	Active phytochrome binding
BAC	Bacterial artificial chromosome
bHLH	Basic helix-loop-helix
BiFC	Bimolecular Fluorescence Complementation
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CIMMYT	Centre for wheat and maize improvement in Mexico
CE	Controlled environment
COP1	Constitutive photomorphogenic 1
CPS	Copalyl diphosphate synthase
CRY	Cryptochrome
CT	Cytokinin
CTD	C terminal domain
D8	Dwarf-8-1
D9	Dwarf-9-1
DNAB	DNA binding
EMS	Ethyl methanesulfonate
ER	Endoplasmic reticulum
ERF	Ethylene-responsive element binding factors
EST	Expressed Sequence Tags
ET	Ethylene
FHL	FHY1-like
FR	Far red
GA	Gibberellin

GAI	Gibberellic acid insensitive
GAMT	GA methyl transferases
gDNA	Genomic DNA
GFP	Green fluorescent protein
GGPP	Trans-geranylgeranyl diphosphate
GID1	Gibberellin insensitive dwarf 1
GUS	Glucuronidase reporter gene
h	Hour
HFR1	Long hypocotyl in far-red
HY5	Elongated hypocotyl 5
IWGSC	International Wheat Genome Sequencing Consortium
KAO	*ent*-kaurenoic acid oxidase
KO	*ent*-kaurene oxidase
KS	*ent*-kaurene synthase
LHR	Leucine heptad repeat
LSD	Least significant difference
miRNA	Micro RNA
MYA	Million years ago
NLS	Nuclear localisation signal
NTD	N terminal domain
P	Phosphate
PAC	Paclobutrazol
PCD	Programmed cell death
Pchlide	Protochlorophyllide
PHOT	Phototropin
phyA	Phytochrome A
phyB	Phytochrome B
PIF	Phytochrome Interacting Factor
PIL	PIF3-like
R	Red

RGA	Repressor of ga1-3
RGL	RGA-like
RNAi	RNA interference
SCF	Skp1-cullin-F-box
SCL	SCARECROW-like
SCR	SCARECROW
SED	Standard error of the mean
SLR	Slender rice 1
SLR1	Slender 1
SLY1	Sleepy 1
SNP	Single nucleotide polymorphism
SREBP	Sterol regulatory element-binding protein
TILLING	Targeting Local Induced Lesions IN Genomes
UTR	Untranslated region
WGS	Whole genome shotgun

CHAPTER 1. INTRODUCTION

1.1 Food security and the importance of wheat

1.1.1 Food security

The topic of food security, due to many converging factors has become a majorly important global issue. Food security defined by the World Food Summit of 1996 is considered to exist when all people at all times have access to sufficient, safe, nutritious food to maintain a healthy and active lifestyle (FAO, 1996). The delivery of this concept, however, is challenged by factors such as an increasing global population, limited land availability for agricultural use, limited access to water and declining resources for the production of fertilisers (Cordell et al, 2009; FAO, 2011). During the last half a century food production has seen a substantial growth. This can be demonstrated by the fact that global cereal production more than doubled between 1960 and 2000 (Tilman et al, 2002). This surge in food production increased the global per capita food supply, decreasing the proportion of the world's population that is hungry. This decrease has happened despite the fact that the world's population has doubled in this time (Godfray et al, 2010). The Green Revolution was a significant component to the increase in food production during the last half a century. The Green Revolution refers to the vast increases in grain yields after the 1960s. These increased grain yields were achieved partly by the development and use of fertilisers and pesticides. In addition, the introduction of new varieties of crops such as semi-dwarf wheat and rice varieties to agriculture were essential to the success of the Green Revolution (Hedden, 2003).

Regardless of the progress made to deliver food security to the world, contemporary agriculture faces enormous challenges. One in seven people still do not have access to sufficient energy and protein in their diet (Godfray et al, 2010). An even greater number of people still suffer from micronutrient malnourishment (FAO, 2009). The current rate of progress will not be enough to eradicate hunger by 2030 or even 2050 as reported by the Food and Agriculture Organization (FAO) (FAO, 2017). In addition, the challenges facing global agriculture and food security today are only predicted to intensify in the next 50

years. The global population is predicted to increase from seven billion people in 2011 to more than nine billion in 2050 and more than 11 billion by the end of the century (FAO, 2017). Additional land needed for growing food however faces increased competition from increased bioenergy production. The demand for biofuels in particular has increased dramatically and is projected to keep rising. For instance, by 2024 it is projected that 25 percent of all sugarcane production will be used for the manufacture of ethanol (FAO, 2017), therefore reducing the land available for food production.

The need for more land for the cultivation of crops results in the conversion of forest into farmland contributing to deforestation. Deforestation poses a significant threat to the environment due to the large production of greenhouse gas emissions, soil erosion, destruction of habitats and loss of biodiversity (FAO, 2017). In addition, climate change is predicted to affect every aspect of food production. Increasing variability of precipitation and the increased likelihood of floods and droughts will inevitably have a negative effect on crop yields. Furthermore, increases in day-time temperature beyond a certain threshold will also cause a decline in crop yields (FAO, 2017). Despite the challenges described here, agriculture in 2050 will need to produce almost 50 percent more food, feed and biofuel than it did in 2012 in order to meet demand (FAO, 2017)

1.1.2 The global importance of wheat

Wheat is one of the most important staples in the world with over 700 million tonnes being harvested annually. Wheat is therefore counted among the 'big three' cereal crops alongside maize and rice. In 2014, for example, approximately 728 million tonnes of wheat was harvested world-wide. This is compared with 1,021 tonnes of maize and 740 million tonnes of rice (http://faostat.fao.org/). Wheat however is unrivalled in its range of cultivation spanning from Scandinavia and Russia to Argentina in the South (Shewry, 2009). Therefore, wheat is grown on a greater area of arable land than any other commercial crop (FAO, 2002). Wheat has a very high starch content, about 60-70% of the whole grain, making the crop an invaluable source of calories (Shewry, 2009). A trade-off to such a high starch content is a relatively low protein content (usually 8-15%). Despite this low protein content wheat still provides as much protein to the diet of humans and livestock as the

protein rich crop soybean. World-wide, wheat contributes 20.46 percent of the protein consumed by humans (Shiferaw et al, 2013). In addition, many of the essential amino acids can be found in considerably higher amounts in wheat grain than the recommended levels for adult humans (Shewry, 2009).

1.1.3 Origin and development of wheat cultivation

The cultivation of wheat first occurred approximately 10 000 years ago (Shewry, 2009) as part of a wide-spread change in human culture. During this time, humans began to change their lifestyle of hunting and gathering to one of agriculture and settlements. This transition is termed the Neolithic revolution. In early farming, wild diploid (einkorn) and tetraploid (emmer) wheats were cultivated and selected presumably due to superior yields. During domestication, traits were further selected for that began to separate them from their wild relatives. Two important genetic traits that were selected for during domestication are briefly described. The first is the loss of shattering of the spike at maturity. Shattering of the spike represents an important trait for ensuring seed dispersal in natural populations however causes seed loss at harvesting. The second trait is the transition from hulled forms of wheat in which the glumes adhere tightly to the grain to free-threshing naked forms (Shewry, 2009).

Wild progenitors of cultivated wheat have now been identified and their area of origin determined. The centre of origin is in a mountainous region in southwest Asia that has been termed the 'Fertile Crescent' (Feldman, 2001). The Fertile Crescent ranges from the foothills of the Zagros mountains in south-western Iran, through the Tigris and Euphrates basins in Iraq and south-eastern Turkey and continues south-west over Syria to the Mediterranean (Feldman, 2001). Situated between the sea, the desert and the mountains, the Fertile Crescent is under the influence of several different climates and thus comprises a wide array of plant species. Archaeological data indicates that agriculture first originated in a narrow strip of the Fertile Crescent known as the Levantine Corridor (Feldmen, 2001).

From its site of origin in the Fertile Crescent, cultivated wheat spread across Europe initially via Anatolia to Greece (8000 BP). From Greece, wheat spread north through the Balkans to the Danube (7000 BP) and also across to Italy, France and Spain (7000 BP). Domesticated

wheat finally reached the UK and Scandinavia approximately 5000 BP (Shewry, 2009). In addition, wheat spread into central Asia via Iran reaching China by about 3000 BP and then to Africa via Egypt. Wheat was subsequently introduced to Mexico in 1529 and then to Australia in 1788 (Shewry, 2009).

1.1.4 Origins of the bread wheat genome

Many globally important crop species cultivated in the modern day have originated from ancestral or recent polyploidization events. This results in species containing large, complex genomes consisting of multiple homoeologous gene copies (Edwards et al, 2013). Bread wheat, *Triticum aestivum*, is a prime example of this phenomenon and contains an allohexaploid genome. The genome consists of three sets of each chromosome pair that have originated from three distinct diploid genomes termed the A, B and D genomes. Each diploid ancestor has contributed 7 pairs of chromosomes resulting in a hexaploid genome with 21 pairs of chromosomes (Edwards et al, 2013).

The hexaploid genome of *Triticum aestivum* is thought to have originated from two independent polyploidization events (International Wheat Genome Sequencing, 2014). The first hybridisation event, estimated to have occurred several hundred thousand years ago, brought together the diploid genomes of *Triticum urartu* (AA) and an unknown grass of the section Sitopsis thought to be closely related to *Aegilops speltoides* (BB). This initial hybridization event produced tetraploid wild emmer wheat, *Triticum turgidum* (AABB), an ancestor of wild emmer wheat cultivated by early farmers and *Triticum turgidum* ssp. *Durum* that is grown for pasta today (Chantret et al, 2005; Edwards et al, 2013; International Wheat Genome Sequencing, 2014). A second interspecies hybridization event between *Triticum turgidum* and the diploid species *Aegilops tauschii* (DD) created the ancestral allohexaploid *Triticum aestivum* (AABBDD) (see figure 1.1).

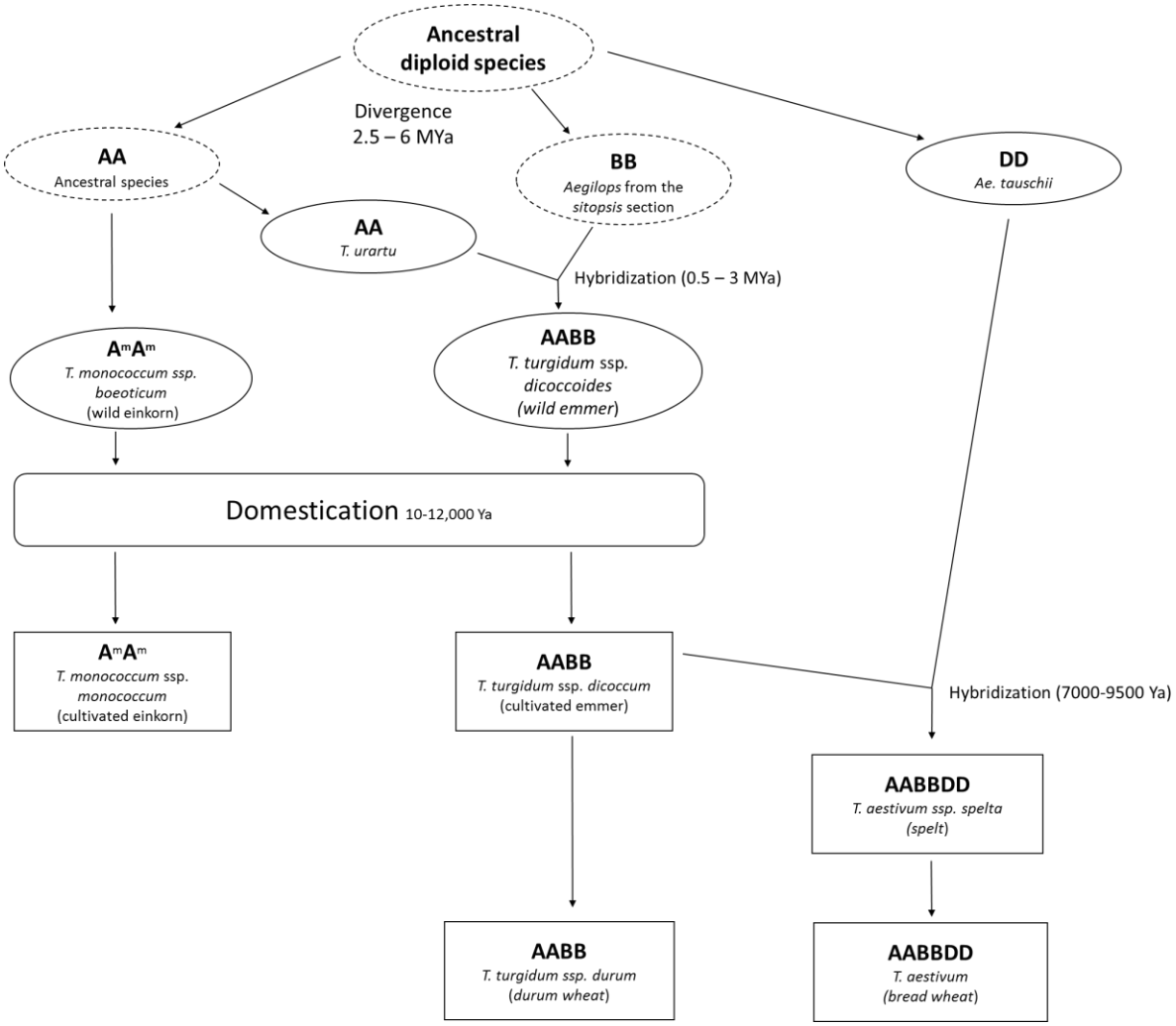


Figure 1.1 Schematic representation of the evolution and domestication of wheat species. Wild and domesticated species are represented in ovals and rectangles respectively. Ancestral species or species that are unknown are surrounded by a dotted line. Actual species are surrounded in a solid circle. Figure adapted from Chantret et al, 2005.

1.1.5 The Green Revolution

The introduction of dwarfing genes into wheat varieties during the Green Revolution represented a major breakthrough in wheat agriculture. Previously the tall stems of high yielding wheat varieties did not have the strength to support the plant. This meant that in adverse weather conditions such as wind and rain the plant was likely to fall over. This process is known as lodging and was responsible for causing significant yield losses (Hedden, 2003). The origin of dwarfing genes in wheat can be found in Japan with the production of

the wheat variety Norin 10. Norin 10 was the progeny of a cross between a semi-dwarf wheat variety and an American high-yielding variety (Hedden, 2003). Subsequently, Norin 10 was used in breeding programmes in the USA. One crucial cross, Norin 10xBrevor 14 was bred with varieties adapted to grow in tropical and sub-tropical climates by Norman Borlaug at the Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT) in Mexico. These varieties were rapidly adopted in Latin America and South Asia causing significant increases in wheat yields in these countries (Hedden, 2003). Of importance, semi-dwarf wheat varieties possessed a shorter stronger stalk that conferred resistance to lodging. In addition, the decrease in stem stature resulted in a greater proportion of assimilate being partitioned into the grain (Hedden, 2003). Wheat yields therefore increased dramatically during the green revolution.

Evidence is now suggesting however that growth rates in wheat yields are either stagnating or declining (Ladha et al, 2003, Brisson et al, 2010). For instance, between 2000 and 2008 wheat production fell by 5.5 percent, a decrease attributable to climatic trends (Lobell et al, 2011). Additionally, in 3 of the past 10 years, worldwide wheat production was not sufficient to meet demand (FAO, 2017). In the face of a global population that is estimated to reach nine billion people by 2050 (United Nations, 2009) these findings are of great concern. It is therefore important to invest time and resources into finding new ways to increase wheat yields.

1.2 Light regulation of plant development

Plants are sessile organisms that transform solar energy into chemical energy through the process of photosynthesis. During a plants life cycle, light acts as a vital informational cue that triggers a multitude of developmental and physiological responses. These responses include seed germination, de-etiolation, directional growth, organelle movement, flowering and senescence (Kong and Okajima, 2016). To enable the plant to respond appropriately to light as an external stimulus, plants are required to perceive a number of parameters relating to the ambient light conditions. Light quality (spectral composition), quantity, direction and duration change depending on the season will all influence how the plant responds (Kami et al, 2010). Changes in light conditions are detected by photoreceptors. In Arabidopsis, four classes of photoreceptors have been identified. Phytochromes are red (R)/far-red (FR) light photoreceptors of which there are five in Arabidopsis, PHYA-PHYE. In addition three classes of UV-A/blue light receptors are known: cryptochromes (cry1 and cry2), phototropins (phot1 and phot2) and zeitlupes (ZTL, FKF1 and LKP2) (Demarsy and Fankhauser, 2009; Franklin et al, 2005). These four classes of photoreceptors are present in all higher plants examined to date (Kami et al, 2010).

1.2.1 The effect of light on monocot species

Light has been shown to regulate the growth and development of monocot species including wheat. Changes in the R/FR light ratio can trigger dramatic changes to the morphology of the plant including changes to tiller number, stem development and leaf elongation (Barnes and Bugbee, 1991, Casal et al, 1985, Deregibus et al, 1983) (for wheat morphology see figure 1.2). In *Lolium multiflorum* (ryegrass) for example, a greater R/FR light ratio significantly increases the number of tillers produced (Deregibus et al, 1983). In addition, above and belowground biomass, number of leaves per plant and leaf area are all increased in plants exposed to a higher R/FR light ratio.

In wheat (Triticum aestivum), plants grown under increased light radiation were shown to have a greater number of tillers and subsequently increased dry mass accumulation (Barnes and Bugbee, 1991). However a low R/FR light treatment was sufficient to reduce the number of tillers by 50%. A low R/FR treatment was also shown to increase leaf sheath

length of early developing leaves and to an even greater extent in later developing leaves. In addition, the rate of development of wheat plants was affected by a low R/FR ratio. Reducing R/FR causes a small but consistent increase in haun stage as early as day 16 after emergence. Plants with an increased development rate due to a low R/FR ratio produce less tillers but accumulate the same amount of dry mass. These findings suggests young wheat plants subjected to low R/FR light divert resources away from producing additional tillers to more rapid main culm development to allow the plant to compete for light (Barnes and Bugbee, 1991). A similar influence of a low R/FR light ratio was observed in ryegrass. A low R/FR caused earlier floral development and increased the number of fertile tillers on each plant. In addition, plants had longer leaf sheaths, blades and reproductive shoots (Casal et al, 1985). These results further indicate that a low R/FR ratio caused by shading represses tillering and promotes growth of taller, reproductive tillers with longer leaves to compete for light.

Irradiance of red light alone has an inhibitory effect on wheat growth and development. Exposure of dark-grown wheat seedlings to R has a marked negative effect on coleoptile elongation (Smith and Jackson, 1987). A 5 minute treatment with R caused a substantial depression of growth and treatment with FR after illumination with R was not sufficient to reverse the R-mediated depression of growth. However, shortening the period of R light to just 1 s followed by immediate 5 min FR caused complete reversibility. In addition to the R light effect on coleoptile elongation, R light has been shown to regulate the greening process of wheat seedlings. Seedlings that have their shoot bases exposed to R light do not accumulate chlorophyll or carotenoids and remain yellow (Sood et al, 2005). The absence of chlorophyll correlates with a reduction in chlorophyll precursors such as δ-aminolevulinic acid (ALA) and protochlorophyllide (Pchlide). The block of greening by R light has been observed in additional monocot species such as barley and rice. Key enzymes in the chlorophyll biosynthetic pathway such as glutamyl-tRNA reductase have also been shown to be negatively regulated by R light perceived by the shoot base in wheat (Sood et al, 2005). Of interest, is that seedlings exposed to R light that have their shoot bases covered retain the ability to green. The shoot bottom contains the meristematic layer, a tissue known to be rich in phytochromes. Phytochromes in these cells are therefore likely be important in perceiving red light. Far-red light and blue light are able to reverse the red light effect on

greening. Treatment with FR light partially rescues the levels of chlorophyll and carotenoids in young seedlings. Seedlings grown in red and blue light however accumulate substantial levels of chlorophyll and carotenoids and demonstrate complete reversibility of the red light effect (Sood et al, 2005). It is possible that this result is caused by a phytochrome-mediated blue light response. Alternatively, an interaction between phytochrome and cryptochrome may be responsible for the reversal of the red light effect (Sood et al, 2005).

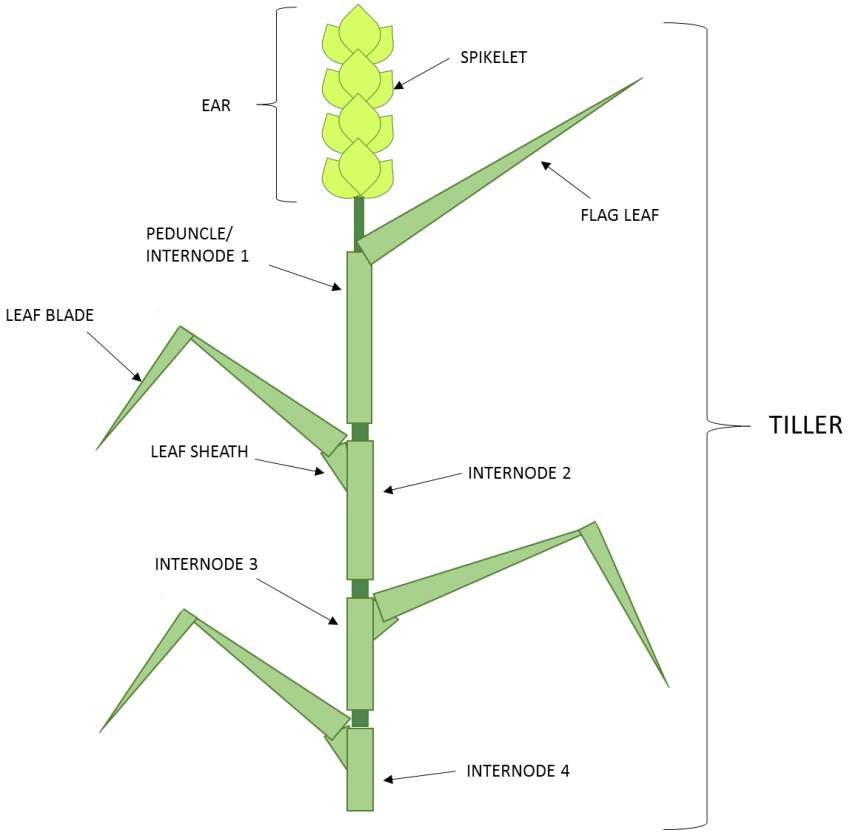


Figure 1.2 Wheat anatomy. A schematic diagram of the morphology of a wheat tiller. The different aspects that make up the morphology of a tiller are labelled individually.

1.2.2 Phytochromes

Arabidopsis, due to the development of molecular genetics in this species, became the model organism to study the mechanisms and functions of photoreceptors in plants (Kami et al, 2010). The Phytochromes are a family of photoreceptors that are required for developmental responses to red and far-red light. Evolutionary studies have suggested that phytochromes in seed plants are encoded by three main clades of genes; *PHYA*, *PHYB* and *PHYC* (Li et al, 2015). Gene duplication events within the *PHYB* clade has given rise to *PHYD* and *PHYE* genes in dicot species such as Arabidopsis. Most monocot species however including wheat, rice and barley contain just a single copy of the three phytochrome genes (*PHYA* to *PHYC*) (Pearce et al, 2016).

Phytochromes consist of an apoprotein that is covalently linked to the tetrapyrrole chromophore, phytochromobilin (Rockwell et al, 2006). Phytochromes regulate gene expression by being able to exist in two distinct but photoreversible conformations. The Pr form absorbs red light maximally at 660 nm whereas the Pfr form absorbs far-red light maximally at 730 nm (Rockwell et al, 2006). In addition to their maximal absorptions of R and FR, it should be noted that phytochromes also weakly absorb blue (B) light (Li et al, 2011). The Pr conformer that is generally considered to be the inactive form can be converted to its biologically active form by the absorption of R. A pulse of FR light reverses this transition. In the dark, phytochromes in their Pr conformation are localised in the cytoplasm. Activation of phytochrome by light acts as a trigger for the translocation of phytochrome to the nucleus to enable regulation of gene expression (Nagatani, 2004). phyB nuclear accumulation is initiated by continuous R light, and to a lesser extent by B light and is reliant on a nuclear localization signal (NLS) located at the C terminal of the photoreceptor. In contrast, phyA nuclear accumulation is efficiently initiated by R, FR and B light. Furthermore, phyA does not contain a typical NLS (Li et al, 2011). It has been shown that two small plant-specific proteins FAR-RED ELONGATED HYPOCOTYL 1 (FHY1) and FHY1-LIKE (FHL) are required for phyA transport into the nucleus (Zhou et al, 2005).

Phytochromes can be classified into two distinct groups according to their stability in light (Sharrock and Quail, 1989). phyA is a type I phytochrome due to being light labile and phyB to phyE are all type II, light stable phytochromes (Li et al, 2011). In dark-grown etiolated seedlings, phyA is the most abundant member of the phytochrome family but is rapidly

degraded upon transfer to light (Franklin and Quail, 2010, Sharrock and Quail, 1989). Light triggers the downregulation of phyA abundance at the transcriptional and post-transcriptional levels. Due to this mode of action, phyA initially functions as a highly sensitive light 'antenna' enabling rapid response to low quantities of light upon soil emergence (Casal and Sellaro, 2014, Franklin and Quail, 2010). This response is termed a Very Low Fluence Response (VLFR). Type II phytochromes (phyB-E) are responsible for mediating photoreversible responses and display relative stability in the light. This mode of action is termed Low Fluence Response (LFR). Phytochrome B is the most abundant phytochrome in plants grown in the light (Franklin and Quail, 2010). Analysis of phytochrome mutants in Arabidopsis has elucidated the role of phytochrome signalling throughout plant development, a brief overview is provided here.

1.2.3 Seedling development

Phytochromes are vital to the plant during seedling establishment. The involvement of phytochrome in mediating seed germination was first demonstrated in lettuce (Borthwick et al, 1952). Lettuce seed was subjected to an alternating R and FR light treatment. Seed that received a R light treatment last achieved an almost 100% germination rate. This rate was markedly decreased in seed subjected to FR last. Multiple mutant studies have shown at least three phytochromes are involved in the control of Arabidopsis seed germination (Franklin and Quail, 2010, Li et al, 2011). phyB has a prominent role in regulating germination in R light as a LFR whilst phyA is responsible for the VLFRs in R and FR. In addition, both phyA and phyE have been found to play a role in promoting germination in continuous FR light.

The transition from dark-grown development (skotomorphogenesis) to light-grown development (photomorphogeneis) is an important stage to seedling establishment. An example of this transition would be a seedling breaking through the soil for the first time. Skotomorphogenic growth (etiolation) is characterised by long hypocotyls, closed cotyledons and apical hook and the development of proplastids to etioplasts. Photomorphogenesis (de-etiolation) however is characterised by short hypocotyls, open and expanded cotyledons and apical hook and the development of mature chloroplasts

(Franklin and Quail, 2010, Li et al, 2011). Phytochromes perform overlapping functions in regulating seedling de-etiolation. phyA has been shown to have a unique role in mediating de-etiolation under FR light. Analysing transcriptional profiles between etiolated WT and *phyA* mutants subject to FR has revealed more than 800 genes under direct or indirect regulation by phyA (Tepperman et al, 2001). Phytochrome B however has a prominent role in regulating de-etiolation under W and R light. This can be demonstrated using a *phyB* mutant that is characterised by elongated hypocotyls, reduced cotyledon expansion and reduced chlorophyll synthesis (Franklin and Quail, 2010). phyA also functions in R light and is implicated in regulating rapid gene expression responses upon R light illumination (Tepperman et al, 2006).

1.2.4 Shade avoidance response

A major function of phytochrome signalling is mediating the shade avoidance response. One of the greatest threats to a plants development and survival is the limitation of light due to shading by neighbouring plants. Shading by neighbouring vegetation is prevalent in environments such as densely planted fields or beneath a forest canopy. Phytochromes are crucial to detecting neigbouring vegetation and initiating an escape response before canopy closure (Franklin and Quail, 2010). Light transmitted through or reflected by a vegetative canopy is deficient in R and B light. In contrast FR and green (G) wavebands are enriched in transmitted or reflected light. This results in a reduction in the R/FR ratio and therefore modifies phytochrome activity. The change in R/FR elicits the shade avoidance syndrome (SAS) (Franklin and Quail, 2010, Li et al, 2011). This includes elongation of stems and petioles. In addition, the plant prioritises the growth of the main central stem over side branches in a phenomenon that is termed apical dominance. If the above physiological responses are unsuccessful in foraging light, flowering is then accelerated enabling seed production in unfavourable conditions (Franklin and Quail, 2010). phyB has a predominant role in suppressing SAS in high R/FR. *phyB* mutants are often described as ‘constitutive shade avoiders’ due to their dramatically elongated, early flowering phenotypes. Other important phytochrome functions involve regulating stomatal development, integration of light signals into the circadian clock and regulation of floral transition (Franklin and Quail, 2010).

1.2.5 Phytochrome regulation of wheat development

Phytochromes have been directly associated with mediating red light induced photomorphogenic responses in wheat. Chen et al (2014) identified loss of function mutations in the *PHYC* gene in tetraploid wheat (AABB). Plants deficient in *PHYC* function showed a significant increase in coleoptile length under red light, indicating an impairment of red light mediated photomorphogenesis (Chen et al, 2014). Flowering time was also affected in a *phyC* mutant in wheat. As demonstrated by the dramatic delay in time to flowering in the *phyC* mutant. A reduction in PHYC activity also resulted in altered spike and spikelet development, a reduced grain set and elongated rachises, glumes and glume awns (Chen et al, 2014). Furthermore, phytochrome has been associated with controlling the unrolling of wheat leaves (Beevers et al, 1970, Virgin, 1962). Red light was shown to stimulate the leaf unrolling response, which can be reversed by far-red light (Virgin, 1962).

1.3 bHLH transcription factors mediate phytochrome signalling

In Arabidopsis, the mechanism by which activated phytochrome alters gene expression is in part through interaction with a small subset of basic helix-loop-helix (bHLH) transcription factors known as the phytochrome-interacting factors (PIFs) (Duek and Fankhauser, 2005). The basic model of phytochrome-PIF signalling is described here (figure 1.3): Phytochrome is activated to its Pfr form by light which enables the translocation of Pfr into the nucleus. Once in the nucleus, phytochrome physically interacts with PIFs and promotes their degradation. Degradation of PIFs relieves the repression of genes involved in photomorphogenesis or inhibits the expression of genes involved in skotomorphogenesis (Pham et al, 2018).

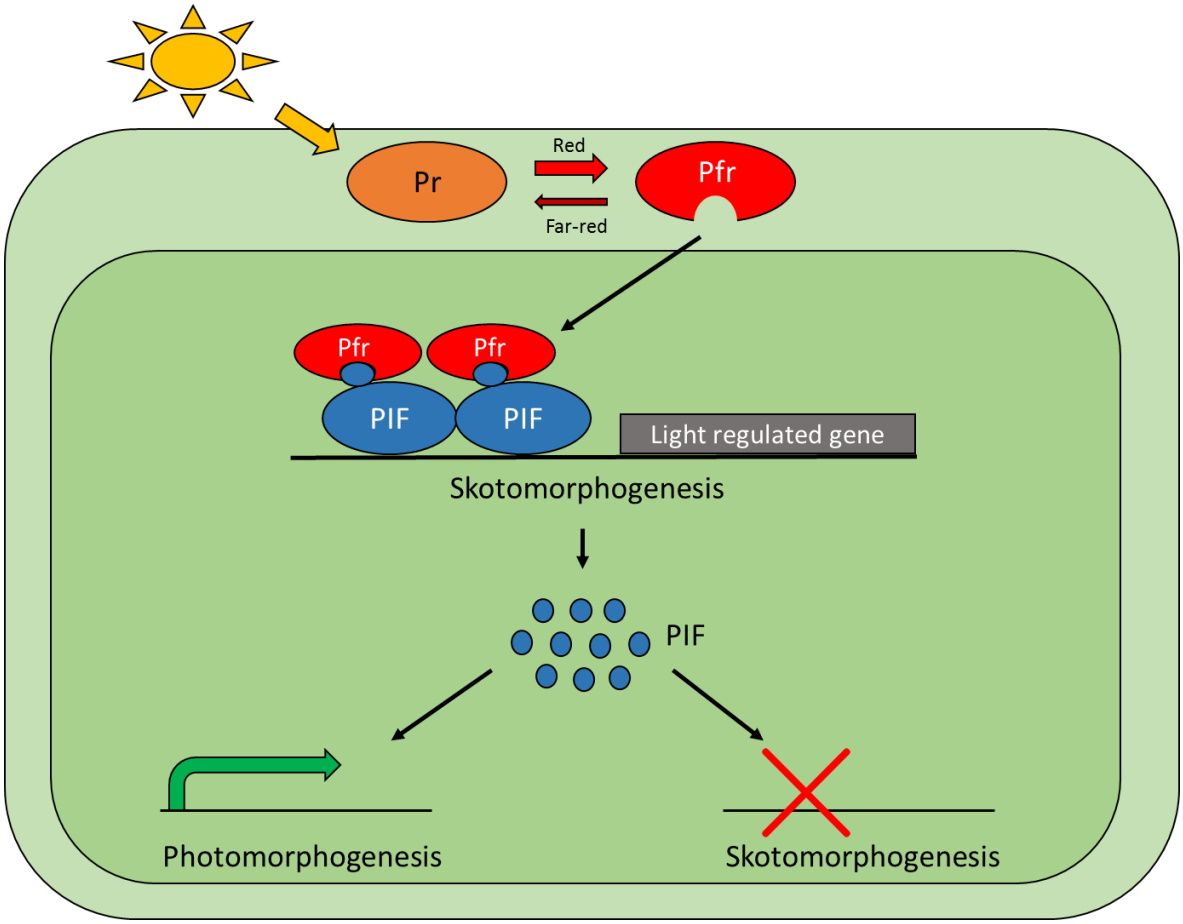


Figure 1.3 Phytochrome-PIF signalling. Phytochrome undergoes a conformational change from Pr to the biologically active Pfr form in response to light. Pfr is able to translocate to the nucleus and interact with PIFs causing their degradation. PIF degradation leads to the promotion of photomorphogenesis.

1.3.1 Identification of the PIF family in Arabidopsis

The PIFs belong to subset 15 of the Arabidopsis bHLH transcription factor family. The founding member of the Phytochrome Interacting Factor family, PIF3, was first isolated in a yeast two-hybrid screen of an Arabidopsis cDNA library using the C terminal domain of phytochrome B as bait (Ni et al, 1998). This protein family are defined by a signature bHLH domain consisting of ~60 amino acid residues. The bHLH domain itself is split into a DNA binding domain, which is located at the N terminal end of the domain and is termed the basic region. This region is often only ~ 15 amino acids in length and has a high proportion of basic residues (Toledo-Ortiz et al 2003). The HLH domain is responsible for the dimerization activity demonstrated by these proteins. The HLH domain itself is comprised of two amphipathic α-helices which are separated by a loop that varies in both length and amino acid composition (Nair and Burley, 2000).

In addition to the signature bHLH domain, all PIF members contain a conserved N-terminal region named the Active Phytochrome B-binding motif (APB). This conserved motif was identified by sequence alignments of the predicted full length Arabidopsis bHLH proteins in subfamily 15 (Khanna et al 2004). The APB motif was found to be present in 12 of the 15 subfamily 15 members. Further sequence alignments and searches in Arabidopsis sequence databases revealed that this conserved motif could not be located in any protein sequences apart from the sequences of the 12 subfamily 15 members. Within the APB motif four invariant residues have been identified (E31, L32, G37 and Q38) (figure 1.4). Positions refer to the positions in PIF3 and PIF5. Point mutations for each of these four residues were found to eliminate the detectable binding of PIF5 to phytochrome B whilst a single point mutation (G35A) prevented PIF4 interaction with phytochrome B and reduced PIF3 interaction by ~75% (Khanna et al 2004).

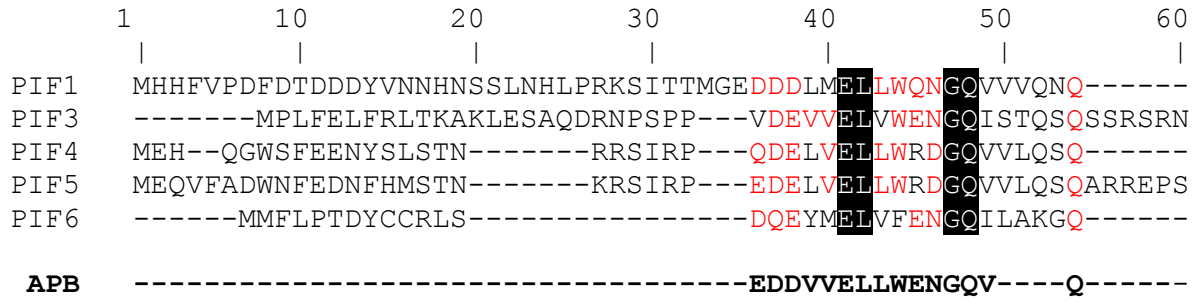


Figure 1.4 The APB motif of Phytochrome Interacting Factor proteins (PIFs). Alignment of the N-terminal amino acid sequences of Arabidopsis PIF proteins. Conserved amino acid residues are shown in red. Amino acid residues that are essential for APB function are highlighted in black. The consensus APB motif is given below.

PIF3 has been shown to bind to a palindromic hexanucleotide DNA sequence CACGTG known as a G-box motif using a random binding site selection procedure (Martinez-Garcia et al, 2000). This motif is a form of the more general E box motif CANNTG, which is considered to be the core nucleotide sequence to which bHLH proteins bind. The specificity of PIF3 binding to the G-box motif was confirmed with the use of electrophoretic mobility shift assays (EMSA) using a G-box containing labelled probe and recombinant PIF3 (Martinez-Garcia et al, 2000). EMSA was also performed with PIF3 and phyB together to determine if phyB can interact with DNA-bound PIF3. The chromophore conjugated Pfr form of phyB was seen to interact with DNA bound PIF3 due to the presence of a discrete lower mobility complex that presumably relates to a ternary complex of PIF3, phyB and the G-box containing probe (Martinez-Garcia et al, 2000).

The subcellular localisation of PIF3 is thought to be permanently in the nucleus. Observations from assays using a PIF3:GUS fusion protein have shown that under continuous white light the PIF3:GUS protein is strongly localised in the nucleus in onion epidermal cells (Ni et al, 1998). This localisation appears to be independent of phytochrome control as in a different experiment PIF3:GUS protein was still localised in the nucleus after far red irradiation and incubation in the dark. A latter study using Arabidopsis transgenic lines expressing PIF3:rsGFP (PIF3 fused to the red-shifted green fluorescent protein) also observed that PIF3 was detected as diffuse staining exclusively in the nucleus (Bauer et al, 2004). Expression was found in all cell types in examined etiolated seedlings. Accumulation

of the PIF3:rsGFP fusion protein was not detectable in the same seedlings that had been grown under continuous light.

1.3.2 Regulation of PIFs by light

The regulation of PIF3 by light is not thought to be at the level of transcription as neither the expression of *PIF3* nor the 35S:PIF3 rsGFP transgene was observed to be down regulated by light (Bauer et al, 2004). Instead FR irradiation was seen to cause rapid formation of nuclear speckles containing the PIF3:rsGFP protein and subsequently the loss of fluorescence. In a further study using Arabidopsis transgenic lines expressing YFP fusion proteins of PHYA, PHYB or PHYD as well as PIF3:CFP, it was observed that a short FR pulse caused translocation of PHYA:YFP from the cytoplasm to the nucleus. Overlay of microscopic images of the location of PHYA:YFP and PIF3:CFP after FR treatment indicated that the two fusion proteins colocalise in the nucleus. Observations made in this study strongly suggest FR causes the degradation of PIF3 in a manner that is reliant on phyA (Bauer et al, 2004). Similarly when seedlings were subject to R irradiation nuclear speckles containing PIF3:rsGFP formed rapidly and within 30 minutes fluorescence could no longer be observed. Microscopic images showed that a short illumination of R caused formation of PHYB:YFP and PHYD:YFP nuclear speckles and that overlay images indicated PHYB:YFP and PHYD:YFP colocalise with PIF3:CFP (Bauer et al, 2004). PIF3 was shown to only bind to the Pfr form of phyB by following a time course of R induced binding of phyB and PIF3 and observing the effect of a FR pulse (Ni et al, 1999). Binding of PIF3 and phyB occurred rapidly after a 5 minute R light pulse and continued more gradually up to 4 hours. Results highlighted phytochrome B Pfr can bind PIF3 in the dark without the need of continuous photoactivation. A pulse of FR at 2 hours caused rapid decrease in levels of bound phytochrome which returned to background levels.

1.3.3 PIF3 degradation is dependent on phytochrome and involves the activity of the 26S proteasome

In a study to examine if PIF3 degradation required the activity of the 26S proteasome, seedlings were transferred to red light from the dark in the presence of two 26S proteasome inhibitors (Park et al, 2004). Results indicated that PIF3 is degraded by the 26S proteasome as PIF3 degradation did not occur in the presence of the inhibitors. In support of PIF3 becoming polyubiquinated a higher molecular weight PIF3 band was observed. This band was apparent after light treatment and disappeared as the duration of light treatment increased as would be expected. The involvement of the proteasome system is supported by further observations that the proteasome inhibitor MG132 inhibits red light-induced PIF3 degradation (Al-Sady et al, 2006). In this more recent study, the electrophoretic mobility shift of PIF3 observed after R irradiation was proposed to be a product of an upstream event to proteasomal degradation. This mobility shift was seen to be abolished by in vitro treatment of PIF3 with calf-intestinal alkaline phosphatase (CIAP), strongly suggesting the mobility shift is the result of light-induced phosphorylation. Evidence that PIF3 phosphorylation is dependent on activated phyA or phyB binding was provided by mutating key amino acid residues on the PIF3 protein for phyA and phyB binding. The presence of a mobility shift was eliminated when phyA and phyB binding to PIF3 had been abolished (Al-Sady et al, 2006). Binding of phyA and phyB to PIF3 has been demonstrated to be necessary for PIF3 phosphorylation but the kinase responsible for this phosphorylation remained unclear.

Two decades ago, phytochrome was proposed to act as a light-regulated protein kinase. Sequencing of the cyanobacterium *Synechocystis* sp. PCC6803 revealed a putative histidine kinase of 748 amino acids that exhibited high sequence conservation surrounding the chromophore binding site of plant phytochromes (Wilde et al, 1997). The bacterial phytochrome was a light-regulated histidine kinase activating a bacterial two-component signalling pathway (Yeh et al, 1997). Plant phytochromes however are not histidine kinases like the phytochrome found in cyanobacteria, but are in fact serine-threonine kinases (Yeh et al, 1998).

In a recent study, Shin et al (2016) provided evidence that phytochrome directly phosphorylates PIFs *in vitro.* Purified recombinant oat phytochrome A (AsphyA) was shown to phosphorylate PIF3 to a greater extent than the control PKS1, the first reported phytochrome kinase substrate. In addition, AsphyA autophosphorylation was reduced in the presence of PIF3 but not PKS1 indicating PIF3 is a preferred substrate of phytochrome. Shin et al (2016) also observed that Arabidopsis phyB and phyD mediated PIF3 phospohorylation. Evidence that phytochrome directly phosphorylates PIFs *in vivo* was provided by analysis of transgenic Arabidopsis plants expressing AsphyA kinase impaired mutants. Transgenic plants expressing kinase-impaired AsphyA significantly inhibited PIF3 phosphorylation and degradation after far-red light (Shin et al, 2016).

1.3.4 The role of PIFs in light signalling

In Arabidopsis, members of the PIF family of transcription factors primarily act to repress photomorphogenesis in the dark. This is best exemplified in the *pif* quadruple mutant (*pif1 pif3 pif4 pif5*, termed *pifq*) that exhibits constitutive photomorphogenic phenotypes. Constitutive photomorphogenic phenotypes of *pifq* mutants include a short hypocotyl and complete cotyledon opening in the dark (Shin et al, 2009). Conversely, *PIF* overexpression lines display hyposensitive phenotypes in response to light (Pham et al, 2018). *PIFs* are known to act redundantly to regulate processes such as photomorphogenesis. However, some PIFs also function individually to mediate certain physiological responses. For instance, PIF1 plays a major role in repressing light-dependent seed germination. PIF1 is known to interact with two important regulatory proteins, LEUNIG_HOMOLOG and HFR1 (LONG HYPOCOTYL IN FAR-RED1) to coordinate seed germination by controlling expression of abscisic acid and gibberellin related genes (Pham et al, 2018). In addition, PIF1 also regulates chlorophyll biosynthesis and plastid development (Huq et al, 2004, Stephenson et al, 2009).

PIF3, the founding member of the Arabidopsis PIF family, functions predominately as a negative regulator of seedling deetiolation (Pham et al, 2018). Similarly to PIF1, PIF3 has a role in repressing chlorophyll synthesis and the development of active chloroplasts in dark-grown seedlings (Stephenson et al, 2009). In addition, a relatively new mode of action has

been described by which PIF3 can modulate seedling responsitivity to light by directly regulating the abundance of phyB levels (Leivar et al, 2008).

Hypocotyl elongation in response to light, shade, temperature and diurnal conditions is primarily regulated by PIF4 in Arabidopsis (Pham et al, 2018). PIF4 acts redundantly with PIF5 to regulate elongation in response to vegetative shade. However, PIF4 appears to uniquely regulate elongation growth in response to high temperature. PIF4 regulates the expression of *TAA1* and *CYP79B2*, two genes that encode enzymes in distinct auxin biosynthetic pathways. PIF4 dependent upregulation of these pathways is proposed to be necessary for an increased growth response (Franklin et al, 2011). Furthermore, PIF4 has been shown to promote flowering in response to high temperatures (Kumar et al, 2012). Chromatin immunoprecipitation (ChIP) demonstrated PIF4 is able to bind near to the transcription start site of the florigen FLOWERING LOCUS T (FT). Binding of PIF4 to the promoter of FT was found to be temperature dependent (Kumar et al, 2012).

PIF5 is known to regulate many of the same pathways as PIF4, often redundantly. Perhaps the most characterised role of PIF5 is its role in the shade avoidance response. PIF5 in collaboration with PIF4 is proposed to promote the expression of shade-induced genes such as *PIL1*, *ATHB2* and *HFR1* (Lorrain et al, 2008). Light abolishes the shade avoidance response by causing the phytochrome-mediated degradation of PIF4 and PIF5. In addition, PIF5 functions as a positive regulator of leaf senescence in the dark and as a negative regulator of red light-induced anthocyanin biosynthesis (Liu et al, 2015, Sakuraba et al, 2014).

1.3.5 PIFs repress chloroplast development

Members of the Arabidopsis PIF family have been shown to be important negative regulators of chloroplast development (Stephenson et al, 2009). *pif1* and *pif3* Arabidopsis single mutants accumulate a greater level of the chlorophyll intermediate protochlorophyllide (Pchlide) in the dark compared to WT with a *pif1pif3* mutant showing an additive response (Stephenson et al, 2009). In addition to increased Pchlide, *pif1pif3* seedlings show a partially developed chloroplast in the dark (Stephenson et al, 2009). Partially developed chloroplasts can be characterised by a reduced prolamellar body (PLB) and increased prothylakoid membrane development. Etioplasts of dark grown *pif1pif3*

seedlings have both increased prothylakoid membrane development and severely reduced PLBs (Stephenson et al, 2009). *pif1pif3* mutants exhibit additional aspects of a light-grown phenotype such as open cotyledons and lack of an apical hook when grown in the dark. This phenotype was also observed in single *pif1* and *pif3* mutants although less consistently (Stephenson et al, 2009). Observations that *pif3* mutants grown in the dark and transferred to R show a delay in greening rate and chlorophyll accumulation previously led to PIF3 being regarded as an activator to light signals (Monte et al, 2004). In contrast, reduced chlorophyll accumulation in *pif3* mutants is now thought to be due to photo-oxidative destruction rather than reduced synthesis (Stephenson et al, 2009). Increased Pchlide accumulation in *pif1* and *pif3* mutants is probably due to changes in *HEMA1* expression. *HEMA1* encodes the enzyme glutamyl tRNA reductase which catalyses the rate-limiting step to tetrapyrrole biosynthesis. *HEMA1* expression was seen to be strongly induced particularly in the *pif1pif3* double mutant at 2 and 3d after germination compared to WT (Stephenson et al, 2009). In a similar study *HEMA1* expression was seen to be 4 fold higher in *pif3* mutants compared to WT after 2 days grown in the dark (Shin et al, 2009). The chlorophyll biosynthesis gene *GUN5* encoding the ChlH subunit of Mg-chelatase was also observed to be expressed to higher levels in *pif3* mutants (Shin et al, 2009).

In addition to regulating tetrapyrrole biosynthesis, PIFs have been shown to have a regulatory role in carotenoid biosynthesis (Toledo-Ortiz et al, 2010). Carotenoid biosynthesis is vitally important to plants during deetiolation. Carotenoids have essential functions in photosynthesis but also importantly in photoprotection where they act by channelling excess energy from light away from chlorophylls and chlorophyll intermediates (Niyogi et al, 1999). Carotenoid biosynthesis is therefore strongly upregulated when an etiolated seedling perceives light for the first time. The rate determining enzyme for carotenogenesis in the dark is phytoene synthase encoded by the gene *PSY* (Rodriguez-Villalon et al, 2009). PIFs act to repress carotenoid accumulation in the dark as *pif1* mutants show increased levels of carotenoids (Toledo-Ortiz et al, 2010). PIF1 is not thought to act independently in regulating carotenoid biosynthesis as the double *pif1pif3* mutant and quadruple *pifq* mutant (*pif1 pif3 pif4* and *pif5*) accumulated carotenoids to a greater level than the single *pif1* mutant in the dark and during deetiolation. PIF1 has been shown to bind to a G-box element in the

promoter of the *PSY* gene in vitro and in vivo to repress expression (Toledo-Ortiz et al, 2010).

1.4 Characterization of Phytochrome Interacting Factor-Like (PIL) proteins in monocot species

1.4.1 PIF-like proteins in rice

In recent years, a lot of work has been undertaken with Arabidopsis to understand the role of PIFs in light signalling. It is my interest to elucidate how PIFs function in wheat. Little is known however about the functions of the PIF family in monocot species and to my knowledge no work has been published that characterises PIF activity in wheat. Phytochrome Interacting Factor-Like (PIL) genes, however, have been identified in the model monocot species rice (*Oryza sativa)*. In 2006, it was reported that rice contains 167 genes that encode a putative bHLH protein (Li et al, 2006). Further analysis of the rice genome has reported that there are six rice proteins that are highly homologous to members of the Arabidopsis PIF family (Nakamura et al, 2007). These six rice proteins named the OsPIL series (OsPIL11-OsPIL16) share not only a highly conserved bHLH domain with PIF3, but also a highly conserved APB motif. It was suggested that there are three pairs of OsPILs and that OsPIL13/14 are counterparts to Arabidopsis PIF4/PIF5 whilst the amino acid composition of OsPIL15/16 is highly similar to PIF3 (Nakamura et al, 2007).

OsPIL15, reported to be highly similar to Arabidopsis PIF3, was shown to localise in the nucleus in a subsequent study using an OsPIL15-GFP fusion protein (Zhou et al, 2014). Interestingly the OsPIL15-GFP protein caused the formation of dotted structures in the nucleus. These dotted structures show resemblance to the speckles formed after R irradiation of Arabidopsis plants expressing the PIF3:rsGFP fusion protein. These dotted structures were proposed to represent interaction between OsPIL15 and phytochrome causing phosphorylation and degradation of OsPIL15 (Zhou et al, 2014). *OsPIL15* overexpression lines grown in the dark for 7 days were seen to be poorly developed against the corresponding WT summarised by shorter aboveground parts and seminal roots (Zhou et al, 2014). When grown under continuous R and FR this phenotype is rescued. Evidence

suggests that light does not rescue the growth of OsPIL15-Ox seedlings by regulating *OsPIL15* transcript levels but likely by regulating protein levels of OsPIL15 in a similar way to regulation of PIF3 protein abundance in Arabidopsis.

In addition to OsPIL15, the function of OsPIL1 has been characterised (also known as OsPIL13) with an emphasis on its role in drought stress (Todaka et al, 2012). Arabidopsis PIF4 and PIF5, which have roles in both hypocotyl elongation and shade avoidance responses, are the closest orthologs based on sequence similarity to OsPIL1 (Nakamura et al, 2007). In accordance, it was shown that OsPIL1 also has a role in regulating plant height (Todaka et al, 2012). Firstly, it was shown that expression of OsPIL1 oscillated in a circadian manner with high expression levels during light periods. This expression pattern was disrupted in plants experiencing drought stress and drought stress was able to inhibit the upregulation of OsPIL1 expression in the light (Todaka et al, 2012). Similarly to OsPIL15, OsPIL1 was also shown to localise in the nucleus using a GFP fusion protein. Todaka et al (2012) demonstrated that OsPIL1 regulates plant height by generating OsPIL1 overexpression lines in rice. Plants that expressed OsPIL11 to a greater extent than control plants demonstrated a strikingly tall phenotype. It was concluded that the tall phenotype was conferred by enhanced cell elongation rather than increased cell division (Todaka et al, 2012). This conclusion was supported by alterations in internode cell size between OsPIL1 overexpressing plants and control plants. In addition, genes involved in cell wall biosynthesis and development were found to be enriched in genes that were upregulated in OsPIL1 overexpression lines (Todaka et al, 2012). However, OsPIL1 was not shown to interact with OsPhyB in a yeast two-hybrid assay and OsPhyB was shown to have no effect on the transcriptional activity of OsPIL1 (Todaka et al, 2012). In conclusion, the data provided by Todaka et al (2012) indicates that OsPIL1 is a key regulator of reduced plant height during drought conditions.

OsPIL1 has also been implicated in regulating chlorophyll biosynthesis, suggesting potential functional similarity to PIF1 and PIF3. Sakuraba et al (2017) identified a T-DNA insertion knockdown *ospil1* mutant to characterise additional functions of OsPIL1. *ospil1* mutants grown in paddy fields under natural long day conditions were shown to exhibit a pale-green phenotype. A pale-green phenotype corresponded to reduced chlorophyll and carotenoid levels in *ospil1* plants compared to WT (Sakuraba et al, 2017). In addition, leaves of 2-week-

old *ospil1* seedlings have significantly reduced levels of photosystem proteins. A reduction in the levels of light-harvesting complex of photosystem II (LHC II) subunits (Lhcb1, Lhcb2 and Lhcb4) , LHC I subunits (Lhca1 and Lhca2, and core subunits of photosystem II and photosystem I was observed. In accordance, transmission electron micrographs revealed that chloroplasts of *ospil1* mutants were reduced in size and had a looser grana structure compared to WT.

To examine the downstream signalling cascade of OsPIL1, Sakuraba et al (2017) performed genome wide microarray analysis to identify differentially expressed genes between WT and *ospil1*. In total 725 genes were found to be significantly upregulated and 840 genes significantly down-regulated in *ospil1.* Chlorophyll biosynthetic genes such as *OsHEMA*, *OsCHLH*, *OsPORA*, *OsPORB, OsDVR* and *OsCAO1* were downregulated in *ospil1* mutants. In particular, *OsHEMA, OsPORA, OsPORB* and *OsCAO1* are proposed to be direct targets of OsPIL1 due to them being upregulated in *OsPIL1* overexpression plants. Furthermore binding of OsPIL1 to the promoter regions of *OsPORB* and *OsCAO1* was confirmed in vivo using ChIP assays. However, OsPIL1 did not bind to the promoter of *OsPORA* in vivo. Of interest, is that microarray analysis revealed that two genes *OsGLK1* and *OsGLK2* encoding GOLDEN2-LIKE (GLK) transcription factors were down-regulated in *ospil1* (Sakuraba et al, 2012). In Arabidopsis two GLK transcription factors directly up-regulate the expression of genes encoding chlorophyll-binding photosystem subunits and chlorophyll biosynthetic enzymes. OsPIL1 was shown to act as a transcriptional activator of the *OsGLK* genes suggesting that down-regulation of *OsGLK* genes in *ospil1* contributes to the pale-green phenotype. In this study, Sakuraba et al (2017) provided evidence that suggests the roles of PIFs in regulating chlorophyll biosynthesis in Arabidopsis and rice might be different.

1.4.2 PIF-like proteins in maize

Phytochrome-interacting factors have also been identified in the crop species maize (*Zea mays*). As previously mentioned maize is a globally important crop. In addition, a fully sequenced and annotated genome is available allowing maize to be used as a model for other monocot species. Of importance is that, due to a recent genome duplication event, maize contains two homeologue pairs for each gene. Gao et al, (2015) identified a putative

PIF3 orthologue, named *ZmPIF3*, in a maize EST (Expressed Sequence Tag) database using the Arabidopsis *PIF3* gene sequence as a query probe. *ZmPIF3* encodes for an amino acid sequence of 638 amino acids, which was shown to include both a bHLH domain and an APB domain characteristic of PIF proteins. Firstly, ZmPIF3 was shown to localise in the nucleus in the same way as OsPIL1 and OsPIL15 (Gao et al, 2015). Of interest to this study was the function of ZmPIF3 in responses to abiotic stresses such as drought stress, salt stress and cold stress. To elucidate the role of ZmPIF3 under these stress conditions, transgenic rice lines were generated that expressed *ZmPIF3* under the control of an ubiquitin promoter. Rice plants that express *ZmPIF3* were grown in both a hydroponic solution and in soil and were shown to exhibit enhanced tolerance to drought and salt stress (Gao et al, 2015). Physiological markers associated with drought tolerance such as relative water content (RWC), chlorophyll content, photochemical efficiency and cell membrane stability were analysed in the transgenic plants. Interestingly, after drought stress all these parameters were enhanced in *ZmPIF3* transgenic plants compared to control suggesting a key role of ZmPIF3 in responding to drought stress (Gao et al, 2015).

Research to characterise phytochrome signalling in maize was also conducted by Kumar et al (2016). A maize bHLH containing protein (GRMZM2G387528) was identified as a PHYB interacting factor by using the C-terminal domain of maize phyB1 ('1' referring to subgenome 1) in a yeast-two hybrid screen. The identified gene subsequently termed *ZmPIF3.2* was found to have a homeologue that was designated *ZmPIF3.1* (Kumar et al, 2016). ZmPIF3.1 shares a very high sequence identity to ZmPIF3.2 with both proteins containing an APB motif and a bHLH domain similar to that of Arabidopsis PIF3. *ZmPIF3* characterised by Gao et al (2015) and *ZmPIF3.2* look to be the same gene as they encode for the same protein sequence. ZmPIF3.1 also interacted with the C-terminal domain of ZmPHYB.1 in a targeted yeast-two hybrid assay. Like most monocot species that have had their genomes characterised, maize contains three phytochrome genes, *PHYA*, *PHYB* and *PHYC* (Matthews and Sharrock, 1996). For each of these genes there are 2 copies due to the recent genome duplication event (Kumar et al, 2016). The results of the study indicated that maize phyB1 in the Pfr conformation has a strong binding affinity to ZmPIF3.1 and ZmPIF3.2 *in vitro*. In contrast, phyA2 showed little if any binding and phyC1 and phyB2 did not show any detectable binding to either of the PIF3 homeologs (Kumar et al, 2016).

In a subsequent study, Shi et al (2018) identified 15 putative PIF proteins in maize using Arabidopsis PIFs as query sequences in a BLASTP analysis. Seven of these putative PIF family members were shown to have a highly conserved APB domain and three were found to have a conserved APA domain. Two of these identified proteins, ZmbHLH16 and ZmbHLH27 were renamed ZmPIF4 and ZmPIF5 respectively due to sequence similarity with Arabidopsis PIF4 and PIF5 (Shi et al, 2018). RT-qPCR analysis demonstrated that *ZmPIF4* and *ZmPIF5* have temporal expression patterns. Expression was seen to increase during the night, peak at dawn and then decrease during the day. In addition, transcript levels of both *ZmPIF4* and *ZmPIF5* were rapidly induced by FR but repressed by R light. Similar to Arabidopsis PIF and rice PIL family members, ZmPIF4 and ZmPIF5 fused to GFP were shown to be exclusively localised in the nucleus.

To determine the ability of ZmPIF4 and ZmPIF5 to function as a PIF, Shi et al (2018) overexpressed *ZmPIF4* (*ZmPIF4-OE/pifq*) and *ZmPIF5* (*ZmPIF5-OE/pifq*) in Arabidopsis *pifq* mutants. As previously mentioned, *pifq* seedlings grown in the dark display a constitutive photomorphogenic phenotype. Overexpression of PIF4 and PIF5 resulted in partial complementation of this phenotype with significantly elongated hypocotyls in *ZmPIF4-OE/pifq* and *ZmPIF5-OE/pifq* seedlings relative to *pifq.* However, cotyledons were observed to be completely closed in *ZmPIF4-OE/pifq* seedlings and only some *ZmPIF5-OE/pifq* lines had open cotyledons. Furthermore, overexpression of *ZmPIF4* in Arabidopsis caused a constitutive shade avoidance phenotype (Shi et al, 2018). Constitutive shade avoidance phenotypes of *ZmPIF4* overexpression lines included earlier flowering times, elongated petioles, reduced leaf area, accelerated leaf senescence, slender inflorescences and plant lodging. Of interest is that Shi et al (2018) proposed that ZmPIF4 might interact with the gibberellin (GA) pathway to influence hypocotyl elongation. This is due to the finding that ZmPIF4 can interact with the DELLA protein REPRESSOR OF GA1-3 (RGA) (Shi et al, 2018).

1.5 Gibberellin signalling

As previously mentioned, this project aims to establish downstream signalling components of the wheat DELLA protein RHT-1. DELLA proteins are central regulators of gibberellin (GA) signalling that function antagonistically to the phytohormone to repress growth. GA was first identified in the fungus *Gibberella fujikuroi*, the pathogen responsible for causing 'foolish-seedling' disease of rice. Infection of plants caused excessive elongation of the stem (Yabuta and Sumiki, 1938). The gibberellins have now been classified as a large group of tetracyclic diterpenoid carboxylic acids that contain an ent-gibberellane (C_{20}) or a 20-nor-ent-gibberellane (C_{19}) carbon skeleton (Hedden & Thomas, 2012). Certain members of this group function as endogenous growth regulators in higher plants, acting to promote organ growth and to regulate switches between different stages of development growth (Hedden, 2012). Developmental processes that are regulated by GAs include seed germination, stem elongation, leaf expansion, trichome development, pollen maturation and the induction of flowering (Daviere and Achard, 2013). More than 130 different gibberellin structures have been identified to date. A nomenclature system is now widely used in which the compounds are assigned the names A_1 – A_{136} ordered according to time of discovery. These names are abbreviated to GA_1 for instance. Only a small number of these compounds are thought to act as bioactive hormones. GA_1 and GA_4 occur universally in plants as the main bioactive forms (Hedden and Thomas, 2012). Therefore, many non-bioactive gibberellins act as precursors to bioactive forms (Yamaguchi, 2008).

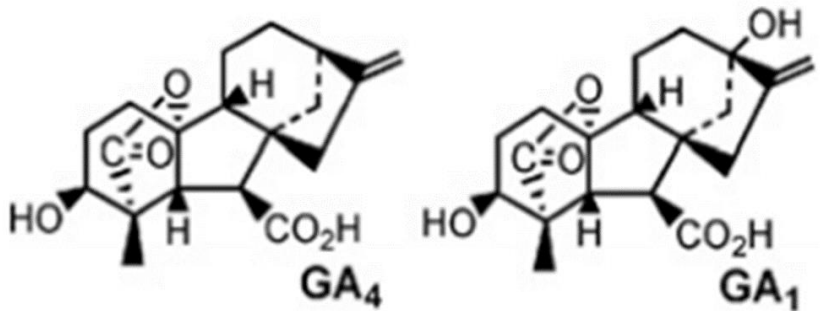


Figure 1.5 The structure of GA_1 and GA_4 (taken from Hedden and Thomas, 2012). The structure of the main biologically active GAs, GA_1 and GA_4, are shown.

1.5.1 Gibberellin biosynthesis

The biosynthesis of GAs in plants requires the activity of enzymes that can be classed into three groups: terpene synthases (TPSs), cytochrome P450 monooxygenases (P450s) and 2-oxoglutarate-dependent dioxygenases (2ODDs) (Yamaguchi, 2008). The first stage of GA biosynthesis occurs in plastids where trans-geranylgeranyl diphosphate (GGDP), a common C20 precursor for diterpenoids, is converted to ent-kaurene (figure 1.6). This conversion is catalysed by two TPSs (ent-copalyl diphosphate synthase and ent-kaurene synthase) both of which are localised in the plastid. ent-copalyl diphosphate synthase (CPS) firstly converts GGDP to ent-copalyl diphosphate. ent-copalyl diphosphate is subsequently converted to ent-kaurene by ent-kaurene synthase (KS). In Arabidopsis, both CPS and KS are encoded by single genes. Loss of function mutations in these genes cause a severe GA-deficient phenotype (Koorneef and van Der Veen, 1980). In other species, especially cereals, two or more copies of each gene have been identified. For instance wheat contains three CPS genes; *TaCPS1, TaCPS2* and *TaCPS3.* Toyomasu et al (2009) suggested that only *TaCPS3* is responsible for GA synthesis and that *TaCPS1* and *TaCPS2* are involved in phytoalexin biosynthesis.

In the next stage of GA biosynthesis, ent-kaurene is converted to GA_{12} within the endoplasmic reticulum (ER). This step involves the activity of two P450s, ent-kaurene oxidase (KO) and ent-kaurenoic acid oxidase (KAO). KO catalyses oxidation at the C-19 position to produce ent-kaurenoic acid, which is then converted to GA_{12} by KAO. GA_{12} represents the common precursor to all GAs in higher plants lying at the branch point of GA metabolism (Sponsel et al, 2004). 13-hydroxylation of GA_{12} to GA_{53} initiates the production of 13-hydroxyl GAs. Evidence has suggested the involvement of P450s and ODDs in this 13-hydroxylation step, however the exact enzymatic activity of this conversion remains unclear (Sponsel et al, 2004). On the opposite side of the branch, oxidation of GA_{12} at the C-20 position results in the formation of non-13-hydroxyl GAs. The formation of the active end products of each branch of the pathway requires the activity of 2ODDs called GA 20-oxidase (GA20ox) and GA 3-oxidase (GA3ox). GA20ox catalyses the sequential oxidation of C-20 on GA_{12} and GA_{53} causing the loss of C20 and CO_2, forming the products GA_9 and its 13-hydroxylated analogue $GA_{20.}$ GA20ox is therefore the enzyme responsible for the production of C_{19} GAs from C_{20} GAs (Yamaguchi, 2008). GA3ox then catalyses the introduction of a 3 β-

hydroxyl group to GA_9 and GA_{20} forming the bioactive compounds GA_4 and GA_1 respectively. In addition to GA_1 another bioactive GA can be formed from GA_{20}. A single GA3ox can catalyse the conversion of GA_{20} to the bioactive GA_3 via the intermediate GA_5.

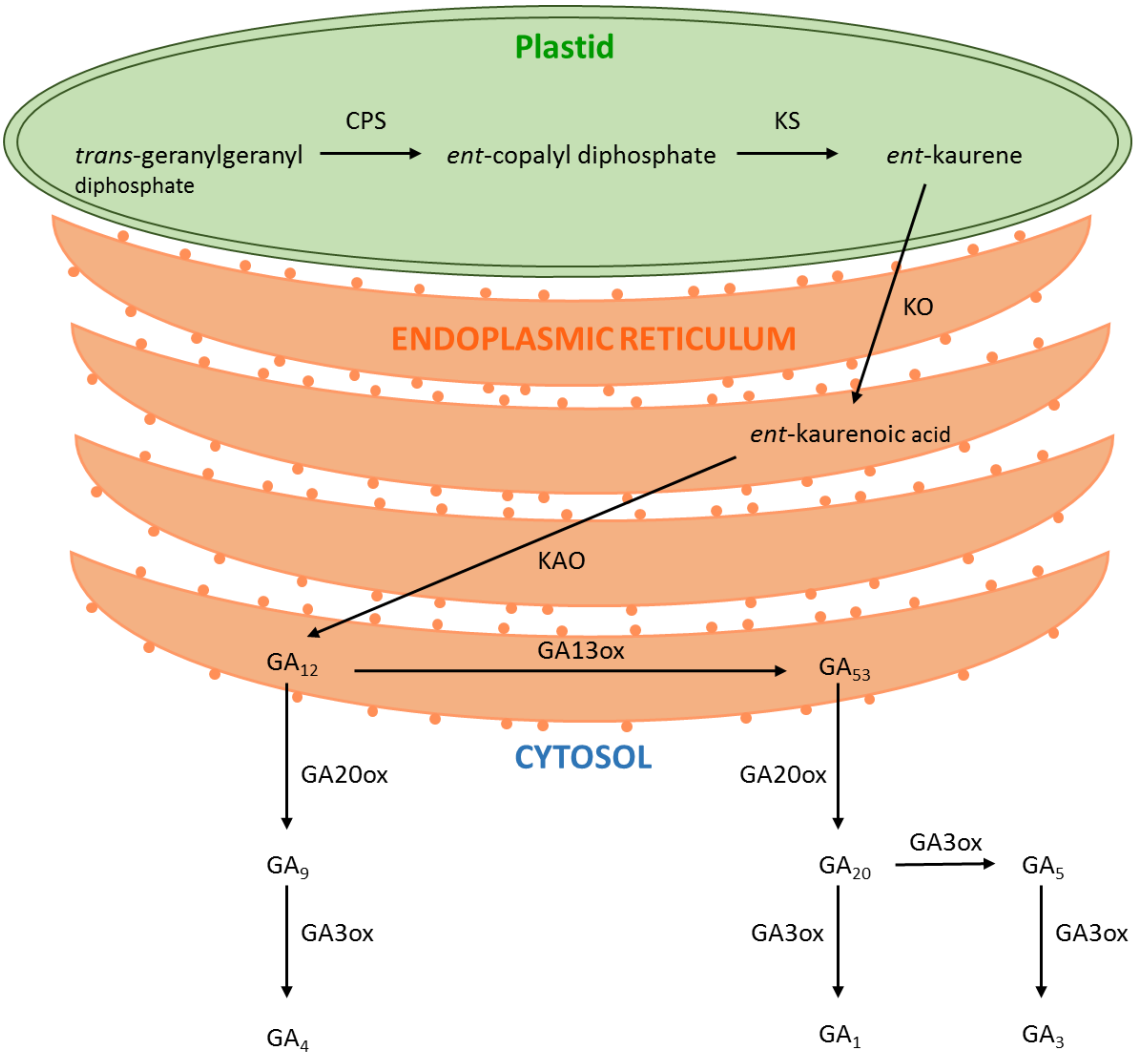


Figure 1.6 The GA biosynthetic pathway. The GA biosynthetic pathway from trans-geranylgeranyl diphosphate to bioactive GA_1, GA_3 and GA_4. The enzymes catalysing each step of the pathway are indicated and the subcellular compartmentalization is shown. CPS, copalyl diphosphate synthase; KS, *ent*-kaurene synthase; KO, *ent*-kaurene oxidase; KAO, *ent*-kaurenoic acid oxidase.

1.5.2 Sites of GA biosynthesis

The concentration of bioactive GAs at their target site will be subject to constant adjustment according to the developmental state of the organ or to changes in the environment. Plants therefore require the ability to precisely regulate GA levels and rapidly change them when needed. The highest concentration of bioactive GAs tend to be found in actively growing organs such as developing leaves rather than in more mature organs (Hedden, 2012). Kaneko et al (2003) used the expression of GA biosynthetic genes encoding GA 3-oxidase and GA 20-oxidase as molecular markers to investigate the sites of GA biosynthesis in rice. Using reporter gene expression assays, GA biosynthesis was shown to occur at the site of action in growing organs (Kaneko et al, 2003). GA biosynthesis was detected in the embryo of germinating seeds, young leaves of developing seedlings, the basal region of elongating internodes and in inflorescence and floral tissues (Kaneko et al, 2003).

A number of studies have suggested that different reactions of the GA biosynthetic pathway can occur in separate tissues, therefore requiring the transport of intermediates between cells. Yamaguchi et al (2001) showed that *CPS* expression was localised in the provasculature of the embryonic axis in germinating Arabidopsis seeds. In contrast, both *KO* and *GA3ox* genes were expressed in the cortex and endodermal cells during seed germination (Yamaguchi et al, 2001). This finding suggests transport of an intermediate, potentially *ent*-kaurene, from the provasculature. In addition to actively growing tissues, *KS* expression was shown to be active in the vasculature of mature leaves (Silverstone et al, 1997). The vasculature is the primary tissue in leaves without mature chloroplasts, therefore unlikely to be capable of *ent*-kaurene synthesis. In support of this finding, Ross et al, (2003) demonstrated that pea plants express *CPS* in mature tissues. These tissues however contained low levels of GA_1 and GA_{20} due to rapid inactivation. The localisation of GA biosynthesis in the vasculature raises the possibility that this tissue is a potential source of GA through export via the phloem (Hedden and Thomas, 2012).

There is evidence to suggest that GA is capable of moving between tissues within the plant to induce a response. In Arabidopsis, a dramatic increase in GA_4 levels can be detected in the shoot apices of plants just before flowering under short day conditions (Eriksson et al, 2006). An increase in GA_4 accumulation in these cells does not correlate with an increase in the expression of GA biosynthetic genes. Eriksson et al (2006) instead proposed GA_4 is

exported from the leaves to the shoot apex in order to induce flowering. Furthermore, in *Lolium temulentum,* exogenously applied GA_5 is transported from the leaf to the shoot apex to promote flowering under long day conditions (King et al, 2001). Results from experiments with pea concluded that GA_{20} is likely to be the mobile form of GA in this species (Proebsting et al, 1992).

Tissues with high levels of GA biosynthesis have been shown to act as sources for neighbouring tissues to support their development. For instance, aleurone cells in germinating cereal grains cannot actively produce GAs, but do have the ability to perceive GA (Kaneko et al, 2003). The aleurone layer instead relies on the import of GAs from the scutellum epithelium. Perception of GA triggers the aleurone to synthesise and secrete hydrolytic enzymes as well as undergo programmed cell death (PCD) (Kaneko et al, 2003). Within flowers, GA biosynthesis is prominent within the tapetum layer of the anthers (Hu et al, 2008). The anthers therefore act as a major source of GAs to organs of the flower that do not synthesise GA. Petals are an example of an organ that cannot synthesise GA and are dependent on transport of GAs from the anthers for normal growth (Hu et al, 2008). In addition to local transport of GA, studies have also supported long distance transport of GAs in plants. A series of micrografting experiments indicated that GA_{12} is the chemical form of GA undergoing long distance transport in Arabidopsis (Regnault et al, 2015). The study demonstrated that GA_{12} has the capacity to move from root to shoot as well as shoot to root. Furthermore, GA_{12} was detected in xylem and phloem exudates suggesting GA_{12} is transported from root to shoot by the xylem and from source to sink tissues in the phloem (Regnault et al, 2015). Recently, several GA transporters have been identified that belong to the NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER FAMILY (NPF) (Chiba et al, 2015, Saito et al, 2015). The NPFs are membrane proteins that have the ability to transport a wide array of compounds across the cell membrane including nitrate, peptides, glucosinolates and phytohormones (Regnault et al, 2016). Although not verified, the NPF multifunctional transporters may contribute to the translocation of GA_{12}.

1.5.3 Regulation of GA inactivation

Mechanisms exist in plants that deactivate GAs providing a means to regulate GA levels rapidly when required (Hedden & Thomas, 2012). Several deactivation mechanisms have been identified, but the most ubiquitous appears to be 2β-hydroxylation (Hedden, 2012; Yamaguchi, 2008). The presence of a 2β-hydroxyl group on GA disrupts the interaction between the GA and its receptor (Murase et al, 2008). 2β-hydroxylation is catalysed by GA2β-hydroxylases (GA2ox) that exist in two classes of 2ODDs. One class contains members that act on C_{19} GAs and the other class contains members that act on their C_{20} precursors. To avoid excessive accumulation of GAs in more mature organs there will be a reduced rate of GA biosynthesis coupled with high rates of GA deactivation. A high rate of GA deactivation is seen in late developing seeds where high levels of GA2ox prevents GAs from causing early germination (Hedden, 2012).

A second mechanism responsible for GA inactivation involves 16α, 17-epoxidation of non-13-hydroxylated GAs (Zhu et al, 2006). In rice, the *ELONGATED UPPERMOST INTERNODE* (*EUI*) gene was found to encode a previously uncharacterised cytochrome P450 monooxygenase. Recombinant EUI produced in yeast was able to catalyse 16α, 17-epoxidation of the precursor GAs, GA_9 and GA_{12} and bioactive GA_4 (Zhu et al, 2006). As a result, the *eui* mutant displays a tall phenotype due to a high accumulation of GA. In contrast, *EUI*-overexpressing transgenic plants have a dwarf phenotype (Zhu et al, 2006). A further mechanism of GA inactivation is mediated by the GA methyl transferases (GAMT). The Arabidopsis genes *GAMT1* and *GAMT2* encode enzymes that catalyse the formation of the methyl esters of GAs (Varbanova et al, 2007). Methylation of GA is suggested to deactivate GAs and initiate their degradation.

1.5.4 Regulation of GA biosynthesis

GAs act as mediators of environmental signals, therefore GA metabolism is sensitive to changes in ambient light conditions. Changes in light quantity, quality or duration may result in increased or decreased GA content. The photoreceptors, phytochrome and cryptochrome have both been implicated in regulating GA levels. Phytochrome regulates GA biosynthesis during seed germination through its interaction with PIF1 (Oh et al, 2006). PIF1 represses

the expression of GA biosynthetic genes *GA3ox1* and *GA3ox2* and activates the expression of the GA catabolic gene *GA2ox2* (Oh et al, 2006). Light stimulated degradation of PIF1 therefore promotes germination by increasing GA levels. During early seedling development, GA signalling suppresses photomorphogenesis in dark-grown seedlings. Illumination overcomes this suppression by causing a rapid reduction in GA content. This reduction involves the activity of phytochrome and cryptochrome and causes the up-regulation of *GA2ox* expression and the down-regulation of *GA20ox* and *GA3ox* genes (Zhao et al, 2007). This activity has been demonstrated by Reid et al (2002) in pea. In this study, the expression of *PsGA3ox1,* the gene encoding the enzyme that catalyses the conversion of GA_{20} to GA_1, was clearly downregulated in response to 0.5 h of exposure to red, far-red or blue light. In addition, expression of *PsGA2ox2* encoding a GA 2-oxidase that converts GA_1 to inactive GA_8 was upregulated by 1 or 2 h of red light and to a lesser extent with blue light. Whilst phyA appears to be responsible for changes in GA levels in response to red and far-red light, Cry1 and phyA act redundantly to reduce GA_1 accumulation in pea seedlings in response to blue light (Foo et al, 2006).

GA metabolism is also regulated through interactions with other hormone signalling pathways. Auxin signalling has been observed in several species to influence the regulation of GA concentrations (Frigerio et al, 2006; O'Neill and Ross, 2002). In Arabidopsis auxins act on the GA biosynthetic pathway by regulating the expression of two genes that encode enzymes that catalyse rate-limiting steps, *AtGA20ox1* and *AtGA20ox2* (Frigerio et al, 2006). This supports the theory that auxin generally promotes GA biosynthesis. In other cases, auxin signalling has also been seen to cause down-regulation of *GA2ox* genes therefore reducing the deactivation of GAs (Hedden & Thomas, 2012). For instance in barley, IAA from the developing inflorescence stimulates stem elongation by upregulating GA 3-oxidation and down regulating GA 2-oxidation (Wolbang et al, 2004). In contrast to the stimulatory effects on GA biosynthesis by auxin, ABA has been observed to suppress GA biosynthesis in imbibed and developing seeds (Seo et al, 2006). It is well documented that GA and ABA have antagonistic roles during seed germination. What has remained elusive is whether ABA regulates endogenous GA levels or vice versa. In a study by Seo and his colleagues it was observed that GA_4 levels in dark imbibed seeds are elevated in an ABA-deficient mutant (Seo et al, 2006). This increase in GA_4 levels linked to an increase in AtGA3ox1 and AtGA3ox2

expression in the ABA-deficient mutant suggests ABA does have a role in regulating endogenous GA levels.

1.6 Gibberellin signalling pathway

The plant hormone GA regulates many processes associated with plant growth and development including seed germination, hypocotyl elongation and leaf expansion (Xu et al, 2014). Plants deficient in GA have a dwarfed phenotype which can be rescued with treatment of GA (Richards et al, 2001). GA promotes growth by targeting the growth repressing DELLA proteins for degradation (Gao et al, 2011). The nuclear localised DELLA proteins belong to the GRAS family of putative transcriptional regulators that are defined by two leucine heptad repeats (LHRI and LHRII) and three domains with highly conserved residues, VHIID, PYFRE and SAW (figure 1.7). In addition, DELLA proteins contain a highly conserved N-terminal DELLA motif (Gao et al, 2011). Various GA signalling mutants were used to identify DELLA proteins.

1.6.1 DELLA proteins are negative regulators of GA signalling

The Arabidopsis genome encodes five DELLA proteins: GAI, RGA, RGL1, RGL2 and RGL3 (Jiang and Fu, 2007). The first DELLA mutant identified in Arabidopsis was the *gai* (gibberellic acid insensitive) mutant. *gai* was shown to be a semidominant mutation that confers a dwarf phenotype with dark green leaves and reduced apical dominance (Koornneef et al, 1985). Therefore, the phenotype of *gai* resembles that of GA biosynthetic mutants. However, *gai* mutants are unresponsive to the application of exogenous GA and accumulate intracellular GA to higher levels than WT, indicating that GA biosynthesis is not impaired. (Koornneef et al, 1985, Talon et al, 1990). These findings therefore demonstrated that the *GAI* gene is likely to be involved in modulating the response of plant cells to GA.

Comparison of *GAI* and *gai* DNA sequences revealed that the *gai* mutant contains a 51-bp deletion within the ORF of *GAI* (Peng et al, 1997). The 51-bp deletion results in a 17 amino acid deletion situated close to the amino terminus of the prediction protein. Analysis of the GAI DNA and protein sequence revealed that GAI is closely related to the Arabidopsis gene

SCARECROW (*SCR*). GAI shares high sequence homology to the carboxyl terminus of *SCR*, particularly to the VHIID domain that characterises this family of regulatory proteins (Peng et al, 1997). In addition, the GAI protein contains two basic regions that are characteristic of nuclear localisation signals (NLS) and two motifs, VHALL and LHKLL, that are proposed to mediate binding of transcriptional coactivators to nuclear receptors. These findings suggested that GAI functions in the nucleus potentially as a negative regulator of GA responsive genes, but whose activity is opposed by GA. Furthermore, Peng et al (1997) proposed that *gai* is a mutant repressor and due to the structural alteration of the protein sequence maintains repression irrespective of the presence of GA.

The *RGA* (*repressor of ga1-3*) gene was subsequently identified as being involved in GA responses (Silverstone et al, 1997b). Mutant alleles at this locus were found to suppress the GA deficient *ga1-3* mutant phenotype (Silverstone et al, 1997). These results therefore suggested that the RGA protein functions as a negative regulator of GA responses in a similar manner as GAI. Alignments of the amino acid sequences of RGA and GAI show that they share high sequence homology. Further examination of the RGA protein sequence revealed that RGA contains a consensus motif for binding of transcriptional coactivators to nuclear receptors, a putative NLS and a VHIID domain (Silverstone et al, 1998). In addition, RGA was shown to contain an N terminal DELLA domain which is also present in GAI (figure 1.7). These findings suggested that RGA might function in a similar or overlapping manner to GAI in controlling the GA response pathway (Silverstone et al, 1998). Furthermore, the 17 amino acids deleted in *gai* are identical between GAI and RGA. Of interest is that the DELLA domain previously mentioned is located within the 17 amino acid deletion. To support the hypothesis that RGA functions similarly to GAI as a repressor of GA, Dill et al (2001) reproduced the same 17 amino acid deletion in the RGA sequence (*rga-Δ17*). Expression of the *rga-Δ17* sequence in Arabidopsis caused a sterile and severe dwarf phenotype and the *rga-Δ17* mutant did not respond to GA treatment to rescue leaf expansion or stem growth, similar to the *gai* mutant. In addition, whilst RGA levels are rapidly reduced in response to GA, the rga-Δ17 protein remained relatively stable. Therefore, Dill et al, 2001 proposed that the DELLA domain is necessary for GA induced degradation and the *gai* and *rga-Δ17* mutants are constitutive repressors of GA signalling.

Sequencing of the Arabidopsis genome identified three additional genes with high homology to *RGA* and *GAI: RGL1, RGL2* and *RGL3 (RGA-LIKE 1, 2* and *3*) (Dill and Sun, 2001, Sanchez-Fernandez et al, 1998). All three are DELLA domain-containing members of the GRAS family and share 56-60% amino acid sequence identity to RGA and GAI (Dill and Sun, 2001). Due to the conserved DELLA domain, the *RGL* genes were also hypothesized to encode negative regulators of the GA signalling pathway. Isolation of *rgl* knockout mutants, revealed that RGL2 plays a primary role in regulating seed germination (Tyler et al, 2004). This result supports previous findings that only *rgl2* affects germination of *ga1-3* seeds (Lee et al, 2002). Further mutant analysis indicates that Arabidopsis DELLAs act redundantly to control flower development. *rgl1* and *rgl2*, in combination with *rga*, increased stamen filament growth, anther development and fertility of *ga1-3* flowers (Tyler et al, 2004). It remains to be seen whether GAI and RGL3 also contribute to flower development.

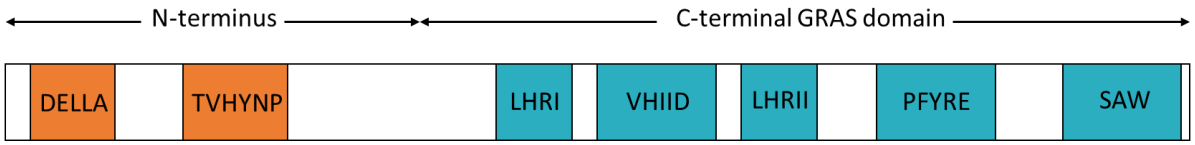


Figure 1.7 Schematic overview of DELLA proteins. A diagram representing the consensus domains that constitute DELLA proteins. The C-terminal GRAS domains Leucine heptad repeat I (LHRI), VHIID, Leucine heptad repeat II (LHRII), PFYRE and SAW are highlighted in blue. Highlighted in orange are the N-terminal conserved domains, DELLA and TVHYNP.

Orthologs of the Arabidopsis *GAI/RGA* genes have been identified in monocot species such as rice, maize, barley and wheat. The rice *slender* gene (*SLR1*) was first identified by characterising the slender-type mutant in rice (*slr1-1*) (Ikeda et al, 2001). Shoots of the *slender* mutant grow more than two-fold taller than WT and resemble the phenotype of WT plants saturated with GA. In addition, application of the GA biosynthesis inhibitor uniconazole had no effect on the increased height phenotype of *slender* (Ikeda et al, 2001). Cloning of the *SLR1* gene revealed that the SLR protein sequence contains the conserved DELLA domain, a putative NLS, and the VHIID domain. The *slr1-1* mutant was found to

contain a single base deletion that alter the N terminal region of SLR (Ikeda et al, 2001). The *slr1-1* mutation was proposed to be a loss-of-function mutation suggesting that the SLR1 protein is a negative regulator of GA signalling. To confirm that SLR1 has a similar role to GAI, 17 amino acids that contain the DELLA domain were deleted from the SLR1 sequence. In a similar manner to *gai,* deletion of the 17 amino acids caused a dwarfed phenotype (Ikeda et al, 2001).

The *slender* (*sln1*) mutant in barley was first reported by Foster (1977). In a similar manner to the *slender* mutant in rice, plants that are homozygous for the *sln-1* allele can be characterized by a tall elongated stem and leaf phenotype. Treatment of the GA biosynthesis inhibitor ancymidol to WT barley causes a severe decrease in plant height. However, *sln1* plants are unaffected by ancymidol demonstrating that the mutants are not responsive to endogenous GA levels (Lanahan and Ho, 1988). *SLN1* was therefore postulated to be an ortholog of the *GAI/RGA/SLR1* genes. In addition, sequence analysis of the SLN1 protein sequence revealed SLN1 is closely related to GAI and RGA and contains an N terminal DELLA domain (Chandler et al, 2002, Fu et al, 2002).

Maize contains two DELLA proteins, encoded by two paralogues *DWARF8* (*D8*) and *DWARF9* (*D9*). A number of mutant alleles of *D8* have been identified and characterised, all of which confer a mild to severe dwarf phenotype (Winkler and Freeling, 1994). The similarities between the *d8* mutant alleles and *gai* first suggested that these genes could be orthologous. Peng et al (1999) established the molecular basis of three d8 mutant alleles: *d8-1, d8-2023* and *d8-Mpl.* Each allele was found to contain a mutation that alters the N-terminal region of the encoded protein. The most severely dwarfing allele *D8-1*, contains a deletion that results in four missing amino acids within the conserved DELLA domain. This allele is therefore similar to the *gai* mutant further indicating that D8 acts as a DELLA protein in maize. Although the *D9* gene had previously been described as a likely paralogue to *D8*, *D9* has only recently been characterised and confirmed as a DELLA gene. The D9 protein sequence shares 92.6% sequence identity to D8 (Lawit et al, 2010). Only one GA-insensitive allele of *D9* has so far been identified. This allele, *d9-1*, encodes a protein that contains an indel and a number of amino acid changes throughout the protein. Lawit et al (2010) proposed one of these mutations, an E-K substitution at the C terminal, could affect

the binding properties of the protein. Furthermore, the E-K mutation was shown to be both necessary and sufficient to cause dwarfing and early flowering in transgenic Arabidopsis.

In hexaploid wheat, mutations of the *Reduced height-1* (*Rht-1*) gene were found to cause a range of dwarf phenotypes (reviewed by Borner et al, 1996). The reduction in plant height was attributed to an abnormal response to the plant hormone GA. In total, seven mutant alleles of *Rht-1* have been identified, four located on chromosome 4B and three on chromosome 4D (Borner et al, 1996). The range of dwarfism conferred by these mutations can be seen in figure 1.8. Two alleles of *Rht-1* in particular, *Rht-B1b* and *Rht-D1b* (formerly *Rht1* and *Rht2* respectively), have been extensively used in agriculture in order to decrease plant height of high yielding varieties. It was shown that *Rht-1* is orthologous to the maize *D8* and Arabidopsis *GAI* genes (Peng et al, 1999). Sequence analysis revealed that GAI/RGA/Rht-D1a/D8 proteins contain two regions of closely related sequence at the N terminus. One of these regions contains the DELLA domain previously shown to be important for GA responses.

Rht-B1b and *Rht-D1b* semi-dwarf alleles only cause about a 20% reduction in height (Flintham et al, 1997). In contrast, alleles *Rht-B1c* and *Rht-D1c* produce a more severe dwarf phenotype and are about half the height of the control plants. Peng et al (1999) showed that both *Rht-B1b* and *Rht-D1b* contain nucleotide substitutions that create premature stop codons. A T-C substitution converts the Q64 codon to a translational stop codon in the *Rht-B1b* mutant. Similarly, in *Rht-D1b*, a T-G substitution converts the E61 codon into a stop codon. Peng et al (1999) proposed that translational reinitiation occurs at one of several methionines that closely follow the premature stop codons. This would lead to a N-terminally truncated product that confers dwarfism through increased repression of GA signalling. Dwarfism in the *Rht-B1d* and *Rht-B1e* alleles is also caused by premature stop codons within the coding sequence of the respective genes. Surprisingly, *Rht-B1d* contains the same point mutation as *Rht-B1b* with no further mutations detected (Pearce et al, 2011).However, *Rht-B1d* plants clearly display a less severely dwarfed phenotype compared to *Rht-B1b* indicating a mutation outside of the Rht-B1 coding sequence is affecting plant height. Characterising the *Rht-B1e* allele revealed a novel mutation that causes a translational stop mutation only three codons upstream of the *Rht-B1b* mutation (Pearce et al, 2011). The extreme dwarf phenotypes of *Rht-B1c* and *Rht-D1c* are not caused by single

point mutations. Pearce et al (2011) identified a 90bp in-frame insertion within the region encoding the N terminal DELLA domain in the *Rht-B1c* mutant. Of interest, a much larger insertion was detected in the genomic DNA sequence. This insertion contains potential splice sites 5' of the 90 bp insertion detected in the transcript. This suggests that the majority of the insertion is subsequently excised during post-transcriptional processing. The extreme dwarf allele *Rht-D1c* is caused by multiple copies of the *Rht-D1b* allele. It is proposed that this results in elevated transcript levels and potential higher amounts of the constitutive growth repressing DELLA.

The semi-dwarfing alleles *Rht-B1b* and *Rht-D1b* were hugely important to the vast increases in wheat yields during the green revolution. These two alleles can now be found in the majority of wheat varieties grown worldwide (Pearce et al, 2011). However, Flintham et al (1997) described the need to tailor the use of different dwarfing alleles to the specific genetic and ecological context. For example, when grown at high temperatures, the yield advantages of *Rht-B1b* and *Rht-D1b* are eroded. Therefore, these two alleles may be too extreme for use during the hot, dry summers of Southern Europe. Alleles that confer an even shorter phenotype such as *Rht-B1c* and *Rht-D1c* have not been exploited for commercial use. Whilst some studies have suggested that *Rht-B1c* has no negative effect on grain yield (Flintham et al, 1983), Li et al (2006) reported a negative effect of *Rht-B1c* on grain yield due to a significant reduction in biomass and a greater susceptibility to adverse growing conditions thus making these plants less reliable.

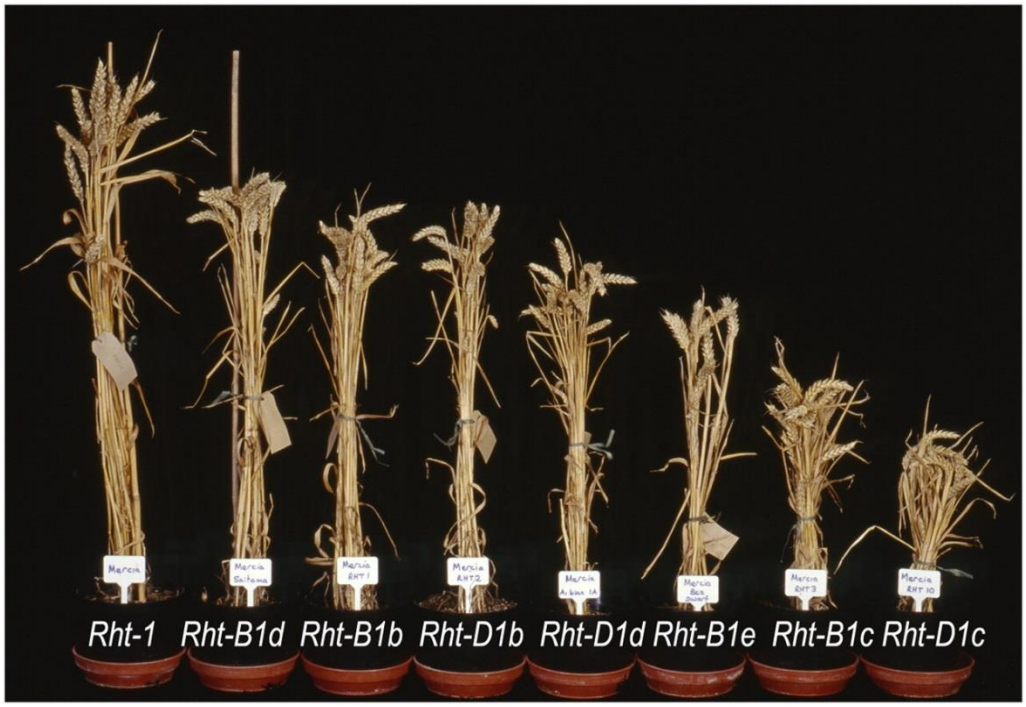


Figure 1.8 Range of *Rht-1* dwarf phenotypes. The height of different *Rht-1* dwarfing alleles are shown compared to wild-type *Rht-1*. Figure taken from Pearce et al. (2011).

1.6.2 The GA receptor: GID1

Studies in Arabidopsis and rice have demonstrated that GA binding to its receptor GA INSENSITIVE DWARF1 (GID1) initiates targeted degradation of DELLAs. Ueguchi-Tanaka et al (2005) first identified GID1 in a screen for GA-insensitive dwarf mutants in rice. One mutant, *gid1-1*, showed a severe dwarf phenotype and lacked GA-responsive phenotypes such as elongation of the second leaf sheath. Analysis of the *slr1* single mutant and the *gid1slr1* double mutant indicated that GID1 functions in the same pathway as the rice DELLA SLR1. Furthermore, GID1 was shown to be essential for degradation of SLR1. GFP tagged SLR1 could be detected in the nuclei of *gid1-1* cells after GA treatment, but not in WT cells (Ueguchi-Tanaka et al, 2005). Database searches could not identify a gene homologous to GID1 in rice whereas there are three homologues genes in Arabidopsis.

Three gibberellin receptor genes, *AtGID1a*, *AtGID1b* and *AtGID1c* were identified and cloned from Arabidopsis (Nakajima et al, 2006). The encoded proteins of all three of these genes showed reversible GA binding activity. Furthermore, all three AtGID1 proteins were able to

bind to each of the five Arabidopsis DELLA proteins in a GA dependent manner. Heterologous expression studies further implicated AtGID1s as GA receptors by the finding that all three genes can rescue the rice *gid1-1* phenotype (Nakajima et al, 2006). Phylogenetic analysis of the deduced amino acid sequences revealed that AtGID1a and AtGID1c are very similar. However, AtGID1b is grouped into a different subgroup indicating a possible difference in biological function.

Crystal structures of GA bound GID1 and the N-terminal GAI DELLA domain have been instrumental in our understanding of the GA-GID1-DELLA complex (Murase et al, 2008). The structure shows a globular GID1 that is bound on one side by the DELLA domain of GAI. GID1A is a monomeric protein, comprising an α/β core domain and an N-terminal extension. The GID1 core domain contains a central deep pocket that accommodates GA. Once GA has bound to the pocket, the N-terminal extension forms a lid that covers GA (figure 1.9). Closing of the N-terminal lid of GID1 allows DELLAs to bind through their conserved N-terminal DELLA/TVHYNP domains. The N-terminal extension therefore lies between between the DELLA domain of GAI and the GID1 core domain. In this way, there is no direct contact between GA and the DELLA domain (Murase et al, 2008).

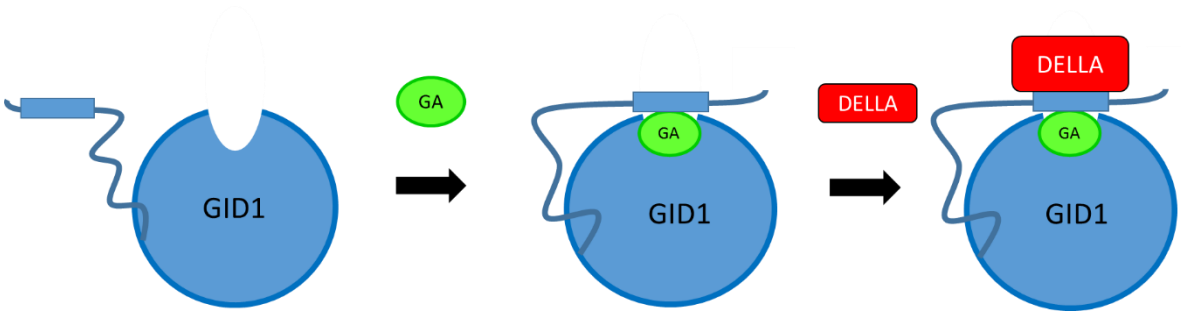


Figure 1.9 The GA-GID1-DELLA complex. A model of gibberellin (GA) regulated GID1-DELLA interaction. GA binds to the pocket of GID1, which causes the closing of the N terminal extension to form a lid. DELLA proteins are then able to bind to GID1 via their conserved DELLA/TVHYNP domains.

1.6.3 Molecular mechanism of DELLA degradation

The formation of a tripartite GID1-GA-DELLA complex initiates the degradation of the DELLA protein (figure 1.10). Specifically, DELLAs are recognised by F-box components of an SCF E3 ubiquitin ligase. This leads to ubiquitination of the DELLA protein and its subsequent degradation via the 26S proteasome (Dill et al, 2004). The SCF complex is named after the three core components, Skp1, cdc53/cullin and F-box proteins. The F-box protein binds to Skp1 and recruits the target via its C-terminal protein-protein interaction domain. In this way, the F-box protein provides the substrate specifically to the SCF E3 complex (Kipreos and Pagano, 2000). The SLY1 (SLEEPY) protein from Arabidopsis and GID2 from rice have been shown to be highly homologous F-box proteins. Loss of function *sly1* and *gid2* mutants display a GA insensitive dwarf phenotype therefore demonstrating the role of SLY1 and GID2 as positive regulators of GA signalling (Sasaki et al, 2003, Steber et al, 1998). In addition, the *sly1* and *gid2* mutants accumulate increased levels of the DELLA proteins RGA and SLR1 respectively (McGinnis et al, 2003, Sasaki et al, 2003). As mentioned the *sly1* loss of function mutant exhibits a dwarf phenotype. This phenotype was rescued in a *sly1rga* double mutant demonstrating that RGA acts downstream of SLY1 in the GA signalling pathway (McGinnis et al, 2003). In a similar manner, Sasaki et al (2003) revealed that the *gid2* dwarf phenotype is rescued in the *gid2slr1* double mutant. This shows that SLR1 also acts downstream of GID2. These results strongly indicate that SLY1 and GID2 are F-box components of the SCF E3 ubiquitin ligase complex that target DELLAs for degradation. The finding that GID2 interacts with a SCF Skp1 component OsSkp2 in a yeast two-hybrid assay further supports this hypothesis (Sasaki et al, 2003).

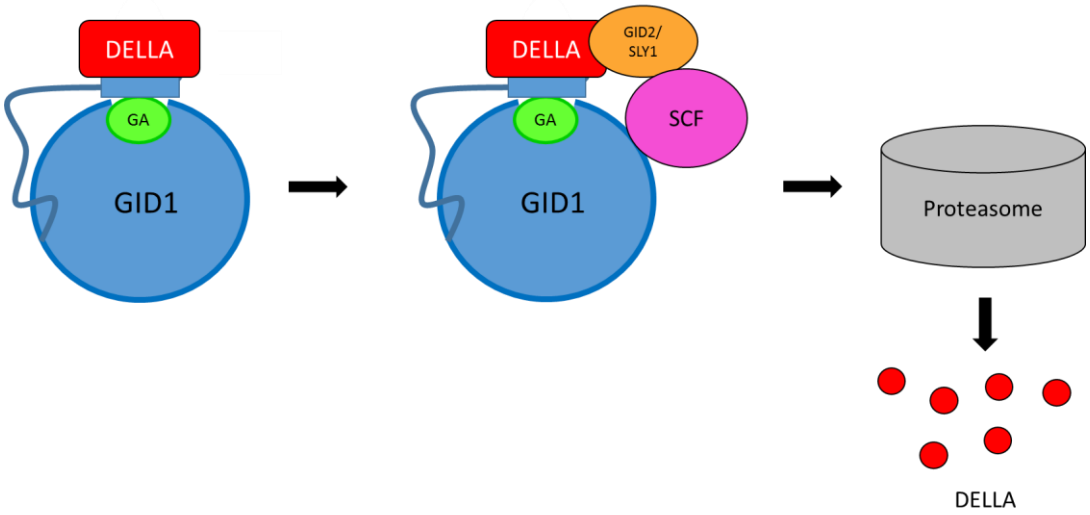


Figure 1.10 GA mediated DELLA degradation. The pathway leading to DELLA degradation by the 26S proteasome. GA binds to the pocket of GID1 causing the N-terminal lid to close. DELLAs bind to the N-terminal lid and are recognised by F-box proteins GID2 or SLY1. This leads to ubiquitination of the DELLA by the SCF complex and subsequent degradation by the proteasome.

1.7 Role of DELLAs during plant development

1.7.1 DELLAs restrict growth and protect against oxidative damage under stress conditions

It has been shown that Arabidopsis plants grown in high salinity conditions contain reduced levels of bioactive GAs (Achard et al, 2006). As would be expected, reduced accumulation of GA corresponds to increased accumulation of DELLAs. The increase in DELLAs both slowed the rate of growth and promoted plant survival in such adverse conditions (Achard et al, 2006). To establish if restriction of growth and greater tolerance to salt stress was correlated, plant height and stress tolerance were measured in Arabidopsis mutants that were deficient in single, double, triple or quadruple DELLAs (Achard et al, 2008). A strong correlation was seen between the relative growth of the plant and the degree of salt tolerance witnessed suggesting a common regulatory mechanism. In order to identify factors involved in the regulatory mechanism, transcript levels of WT, *ga1-3* (GA deficient) and a *ga1-3* quadruple-DELLA mutant were analysed after short exposure to high salt conditions (Achard et al, 2008). Analysis of the changes in mRNA levels revealed 2.2% of the

24,576 Arabidopsis genes represented in the study showed DELLA dependent changes (Achard et al, 2008). Comparisons with additional public available microarray data led to the identification of many DELLA-regulated genes that respond to oxidative stress. Among these genes that showed DELLA dependent upregulation were the genes that encode the antioxidant enzymes Cu/Zn superoxide dismutases, catalases, peroxidases and glutathione S-transferases (Achard et al, 2008). To confirm that upregulation of these antioxidant systems conferred protection against oxidative damage, the levels of reactive oxygen species (ROS) were determined in WT, *ga1-3* and *ga1-3* quadruple DELLA mutant roots with the use of a ROS sensitive dye. In the absence of salt stress the levels of ROS in the WT and *ga1-3* quadruple DELLA roots were greater than in the roots of the GA-deficient plant. Upon application of 50 mM NaCl, ROS levels increased in the WT roots and increased to a greater extent in the roots of the *ga1-3* quadruple DELLA mutant. In contrast, there was no detectable accumulation of ROS in the roots of salt-treated *ga1-3* mutants (Achard et al, 2008). In conclusion, under stress conditions it appears that DELLAs accumulate at least in part due to a reduction in bioactive GA. The accumulation of DELLAs causes restriction in plant growth and upregulation of antioxidant systems that reduces accumulation of ROS.

1.7.2 Regulation of gene expression by DELLAs

Evidence suggests that DELLA proteins are nuclear localised and have a role in regulating gene expression (Cao et al, 2006). In a study using chromatin immunoprecipitation, the direct interaction between DELLA proteins and promoters of GA-responsive genes could not be observed (Feng et al, 2008). This proposes that gene regulation mediated by DELLAs must occur indirectly by influencing the activity of transcription factors. In Arabidopsis, DELLAs have been shown to interact with the PIFs (Feng et al, 2008; de Lucas et al, 2008). The interaction between DELLAs and PIFs could therefore represent a convergence point between light and hormonal signalling.

Evidence that DELLAs physically interact with PIF3 and PIF4 originates from yeast two-hybrid assays, *in vitro* pull-down assays and bimolecular fluorescence complementation analysis (Feng et al, 2008; de Lucas et al, 2008). Further analysis using an immunoprecipitation approach revealed the interaction between PIF3 and the Arabidopsis DELLA protein RGA in

dark-grown seedlings. This interaction is dependent on the relative abundance of GA such that treatment with paclobutrazol enhances the PIF3-RGA interaction and high levels of GA causes PIF3 release. Deletion studies have revealed that the interaction between PIF4 and RGA is dependent on both the conserved bHLH DNA recognition domain in PIF4 and the first conserved heptad leucine repeat in the RGA protein (de Lucas et al, 2008).

These two independent studies have demonstrated that DELLAs can physically interact with PIF3 and PIF4 and can sequester them from binding to their promoter targets (Feng et al, 2008; de Lucas et al, 2008). In these studies, sequestration of PIFs by DELLAs was observed to inhibit hypocotyl elongation. More recently, DELLAs have also been shown to regulate PIF3 activity by a second, distinct mechanism (Li et al, 2016). Li et al (2016) demonstrated that in cases where DELLA proteins accumulate, such as in a constitutively active DELLA mutant, PIF3 protein abundance is reduced. A decrease in PIF3 protein abundance corresponded to a shorter hypocotyl length. DELLA-mediated PIF3 degradation was shown to be independent of light-mediated degradation of PIF3, which requires phytochrome. The mechanism by which DELLAs promote PIF3 degradation is unknown. However, the E3 ubiquitin ligase COP1 and DET1 have been shown to be essential for the DELLA-dependent decrease of PIF3 protein abundance (Li et al, 2016). Li et al (2016) provided evidence that both mechanisms, sequestration and degradation of PIF3, contribute to the inhibition of PIF3 activity (figure 1.11). Briefly, the binding of PIF3 to five target promoters was analysed under different conditions. In PAC-treated seedlings, the enrichment of these promoters was significantly decreased. On the other hand, treatment with PAC and a proteasome inhibitor (MG132) increased the enrichment of the five promoters compared to PAC-treated seedlings. This indicates that sequestration by itself can inhibit PIF3 activity, whilst degradation provides an additional level of inhibition. In addition to PIF3, protein abundance of PIF1, PIF4 and PIF5 was also shown to be negatively regulated by DELLAs (Li et al, 2016).

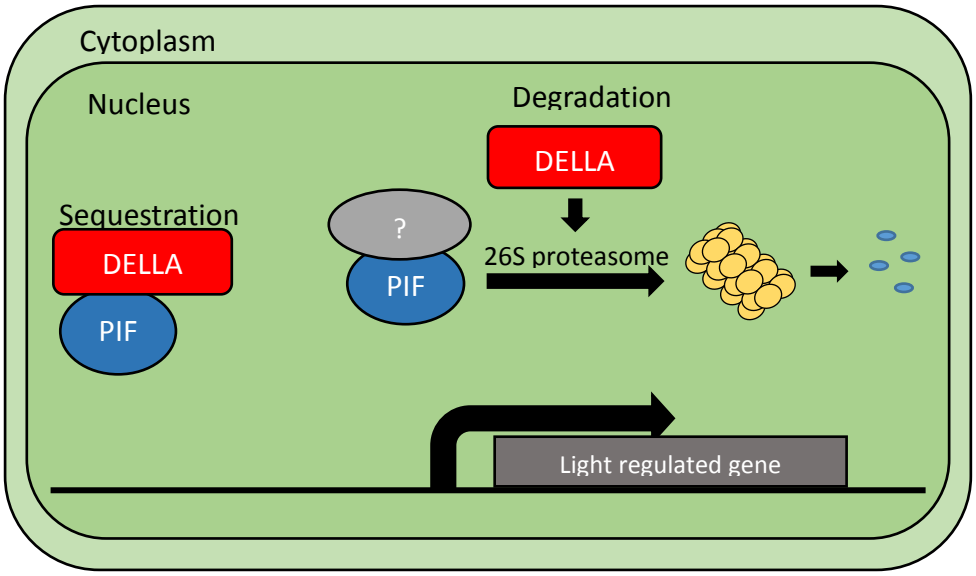


Figure 1.11 DELLA regulation of PIF activity. DELLAs negatively regulate PIFs by sequestering them from target promoters or by promoting PIF degradation via the 26S proteasome.

1.7.3 DELLA-PIF complex regulates expression of chlorophyll and carotenoid biosynthetic genes in the dark

The tight regulation of chlorophyll biosynthetic genes and photosynthetic genes by the DELLA-PIF complex was indicated by comparing the expression patterns of genes under the control of both DELLAs and PIFs (Cheminant et al, 2011). Included in the at least 40 chlorophyll biosynthetic and photosynthetic genes that overlapped were *LHB1B1*, *LHCB2.2* and *LHCB1.1* encoding chlorophyll binding proteins. Reduced GA levels due to treatment with PAC caused increased expression of all three genes in a WT background and in single and double *pif* mutant backgrounds. Enhanced expression due to reduced GA was not seen however in *pifq* seedlings indicating that DELLAs prevent PIF1, PIF3, PIF4 and PIF5 from repressing the three genes. DELLAs inhibit PIFs from repressing light-regulated genes by preventing PIFs from binding to the G-box element on the promoters of these genes. PIF1 was observed to bind to the promoter of *LHCB2.2*. This interaction was reduced when DELLAs were able to accumulate due to the presence of PAC. The same results were observed with the chlorophyll biosynthetic genes *PORC, CHLOROPHYLL A OXYGENASE (CAO)* and *CONDITIONAL CHLORINA (CHLH)* (Cheminant et al, 2011). The notion that DELLAs

prevent PIFs interacting with their promoter elements is supported by Feng et al (2008) who selected five putative PIF3 target genes and confirmed PIF3 interaction with the five promoters. In all cases, DELLA accumulation mediated by PAC reduced PIF-promoter binding. PIF4 interaction with its target promoters has also been seen to be reduced by PAC treatment (de Lucas et al, 2008). Treatment with GA reversed the effect of PAC by destabilising the DELLAs. Interestingly PIF1 has been reported to regulate the transcription of two DELLA proteins GAI and RGA in germinating seeds. PIF1 was observed to bind to G-box elements on the promoters of *GAI* and *RGA* (Oh et al, 2007). This finding suggests a feedback mechanism between DELLAs and PIFs.

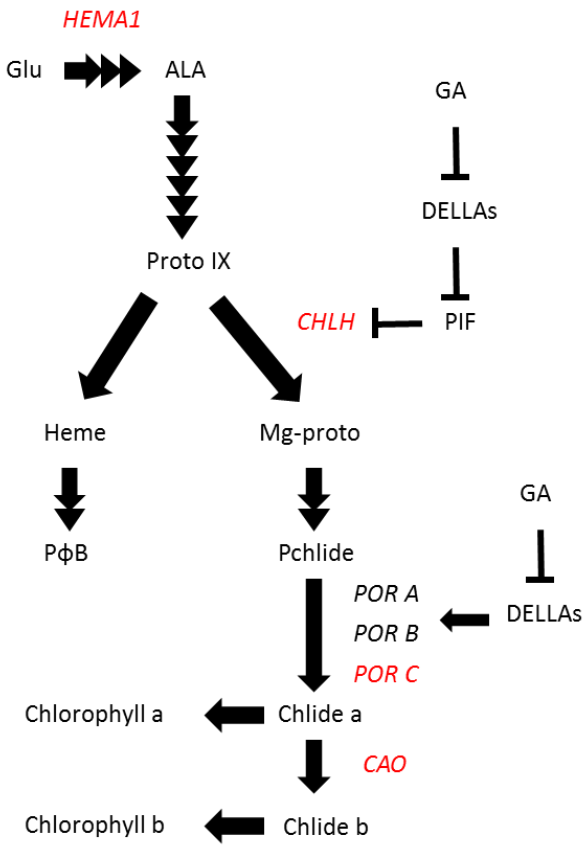


Figure 1.12 Regulation of the chlorophyll biosynthetic pathway. DELLAs regulate expression of key chlorophyll biosynthetic genes by preventing PIFs from binding to the promoters of light-regulated genes. Genes highlighted in red are proposed to be under DELLA-PIF regulation. ALA, 5-aminolevulinic acid; HEMA, glutamyl-tRNA reductase; Proto IX, Protoporphyrin IX; PϕB, Phytochromobilin; CHLH, Mg-chelatase H subunit; Mg-proto, Mg-Protoporphyrin IX; Pchlide, Protochlorophyllide; POR, NADPH:protochlorophyllide oxidoreductase A, B and C; Chlide, Chlorophyllide; CAO, Chlorophyllide a oxygenase. Figure adapted from Cheminant et al (2011).

Carotenoid biosynthesis is also known to be repressed in the dark by the PIFs (Toledo-Ortiz et al, 2010). This repression can be lifted in the dark with treatment of PAC (Rodriguez-Villalon et al, 2009). This is supported by findings that Arabidopsis seedlings unable to synthesise GA accumulate higher levels of the four main species of carotenoid (lutein, β-carotene, violaxanthin and neoxanthin) (Cheminant et al, 2011). This derepression of carotenoid accumulation is DELLA-dependent. This can be seen in a DELLA global-deficient mutant where the increase in carotenoid levels is abolished. Similar to the DELLA-PIF control of the previously mentioned chlorophyll biosynthesis genes, it was observed that PAC treatment reduces the binding between PIF1 and the G-box motifs on the *PSY* promoter (Cheminant et al, 2011). In the dark DELLA levels are low due to being targeted for degradation by GA. During the transition of the plant moving from the dark to the light, GA levels will decline allowing DELLAs to accumulate. The requirement for carotenoid biosynthesis due to the new light conditions will be achieved by DELLAs sequestering PIFs from binding to their target promoters.

1.7.4 DELLAs confer protection against photodamage

A vital feature distinguishing etioplasts from other plastid types is the prolamellar body (PLB). The PLB is a highly regular curved lattice of tubular membranes consisting predominantly of the lipid monogalactosyldiacylglycerol, the enzyme NADPH:protochlorophyllide oxidoreductase (POR) and carotenoids, typically lutein and violaxanthin (Franck et al, 2000; Park et al, 2002). POR catalyses the reduction of the chlorophyll precursor Pchlide, the only reaction in the chlorophyll biosynthetic pathway that directly requires light (Frick et al, 2003). In the dark chlorophyll synthesis halts at Pchlide accumulation. POR binds to its substrate Pchlide poised such that a photon of light leads to immediate reduction of Pchlide to Chlide (Frick et al, 2003). Levels of POR and Pchlide need to be stoichiometrically linked as unbound Pchlide acts as a potentially lethal photosentizer (op den Camp et al, 2003).

Etiolated Arabidopsis seedlings that accumulate DELLAs accumulate greater levels of Pchlide in the dark (Cheminant et al, 2011). Etiolated GA-deficient seedlings therefore would be expected to be severely bleached when transferred from the dark to the light. In fact, *ga1-3*

(GA deficient mutant) and *gai* mutants were more resistant to photobleaching than the WT control despite having greater levels of Pchlide with the potential to cause photo-oxidative damage. This is the opposite of *pif* mutants which have been shown to be a lot more susceptible to photobleaching than WT seedlings (Shin et al, 2009, Stephenson et al, 2009). Resistance conferred by GA deficiency was partially supressed by the absence of DELLAs in a DELLA knockout mutant indicating DELLAs must have a role in protecting against photo-oxidative damage (Cheminant et al, 2011).

DELLA-mediated resistance to photooxidative damage was shown to be partly due to upregulation of the photoprotective enzyme POR (Cheminant et al, 2011). Arabidopsis contains three POR genes, *PORA*, *PORB* and *PORC* which encode structurally-related but differentially-regulated PORs (Frick et al, 2003). All three *POR* genes have been seen to be induced in cotyledons of dark grown GA-deficient Arabidopsis seedlings (Cheminant et al, 2011). The involvement of DELLAs in regulating the increase in *POR* expression in *ga1-3* mutants was indicated due to *ga1-3* DELLA global mutants supressing this upregulation. The mechanism by which DELLAs regulate *POR* expression appears to be independent of PIFs since POR expression is upregulated in *ga1-3 pif3-1* mutants (Cheminant et al, 2011). This upregulation in the absence of PIF3 suggests that DELLAs cannot be inducing expression of *POR* genes by relieving repression mediated by the PIFs. The regulatory mechanism of POR expression exerted by DELLAs remains unclear, however, it is likely to be indirect due to failure to obtain evidence of direct binding of GAI to the promoter of *PORA* and *PORB* (Cheminant et al, 2011).

1.8 Project aims

This project aims to identify the signalling pathways required to protect against photo-oxidative stress during chloroplast development in wheat. Chloroplasts are essential organelles in the plant cell that are responsible for harvesting the energy obtained from the sun. The development of chloroplasts, however, has to be carefully orchestrated to prevent the excitation of chlorophyll intermediates and the accumulation of reactive oxygen species

(ROS). Previous work with Arabidopsis has indicated that the plant hormone gibberellin (GA) has a major role in oxidative stress responses. GA signalling mediates a wide array of developmental processes and plant responses by causing the degradation of the growth repressing DELLA proteins. The wheat DELLA protein, RHT-1, was central to the green revolution. It was elucidated that gain-of-function mutant alleles of the *Rht-1* gene were responsible for the varying degrees of dwarfism witnessed during the green revolution. The mechanism by which RHT-1 functions in wheat however is largely unknown. This project aims to determine downstream signalling components of RHT-1 in hexaploid wheat.

In Arabidopsis, DELLAs have been shown to function through an interaction with members of the Phytochrome Interacting Family (PIF). PIFs are light regulated transcription factors that regulate different aspects of plant development including chloroplast biogenesis. Therefore, DELLA-PIF interactions could represent a convergence between hormonal signalling and light signalling in plant development. How DELLAs and PIFs interact in wheat however is unknown. Understanding these signalling pathways in wheat will provide the tools to improve seedling establishment and yield, factors central to the issue of food security.

This project will test the hypothesis that the DELLA protein RHT-1 interacts with specific PIFs to regulate chloroplast development in wheat.

Chapter 2. Materials and Methods

2.1 Plant growth conditions

Triticum aestivum cv. Cadenza plants were grown in individual 13-cm pots in the glasshouse at the University of Southampton. Prior to sowing, wheat seeds were imbibed on moist filter paper at 4°C for 3 days. Seeds were sown to a depth of approximately 2 cm in compost containing equal volumes of Levingtons F+S modular compost, John Innes No 2 and Vermiculite. Compost was supplemented with 5 g/L of slow release fertiliser. Plants were grown under a 16 hour photoperiod provided by natural light and supplemented with banks of SON-T 400W sodium lamps giving 400-1000 µmol m^{-2} sec^{-1} total light.

2.1.2 Growth of Arabidopsis seedling material

For each genotype, approximately 200 *Arabidopsis thaliana* seeds were surface sterilised using 10% (v/v) bleach for 10 minutes. Seeds were subsequently washed three times using sterile water. Seeds were then sown onto autoclaved ½ MS (Murashige & Skoog, 1962) media (1% agar) in 90 mm diameter plates and sealed with parafilm. Plates were wrapped in aluminium foil and placed in a cold room (4°C) for 48 hours, for stratification. Plates were routinely exposed to 2 hours white light to stimulate uniform germination. Plates were subsequently re-wrapped in foil and placed in a dark cabinet (23°C) for 5 days for protochlorophyllide extraction and measurement of hypocotyl length.

2.2 Protochlorophyllide extraction

To extract protochlorophyllide (Pchlide), 20 *Arabidopsis thaliana* seedlings were homogensised using a micropestle in 0.5 mL cold extraction solvent (acetone:0.1 M NH_4OH, 9:1, v/v). Samples were then centrifuged at 13,000 g in a bench-top microcentrifuge. The supernatant was kept and the pellet re-extracted with 0.3 mL fresh extraction solvent. Extracts were subsequently combined for Pchlide quantification.

Pchlide was measured by relative fluorescence at 628 nm following excitation at 440 nm using a Hitachi F-3010 fluorescence spectrophotometer (Hitachi, Tokyo, Japan).

2.3 Identification of wheat TILLING lines and wheat crossing

Mutant *Triticum aestivum* lines for *TaPIF3* were identified from the Cadenza TILLING population (www.wheat-tilling.com). This TILLING population was generated by ethyl methanesulfonate (EMS) mutagenesis of wheat cv Cadenza. The exome of 1200 mutagenized M3 lines was sequenced and mapped to the IWGSC reference sequence. Sequence reads generated from the exome capture were mapped to a subset of gene targets that included the three homoeologues of *TaPIF3*. TILLING lines for *TaPIF3* were selected that contain premature stop codons or disruption of splice sites.

Plants containing mutations in each homoeologue of *TaPIF3* were crossed to produce a *TaPIF3* triple mutant. Buds from the female donor plant were emasculated by removing three anthers per floret. Anthers were removed prior to flower opening and pollen release to prevent self-fertilisation. Glumes of the ear were cut back to facilitate pollen transfer from a male donor plant. After emasculation, the resulting ear was left for 2-3 days to mature. The emasculated ear was subsequently pollinated with an ear from the male donor plant. Grain resulting from successful crosses was allowed to mature prior to harvesting. The genotype of resulting offspring was determined by sequencing (2.5.14).

2.4 Phenotypic characterisation of wheat lines

A mature plant phenotype experiment was set up using the *TaPIF3* triple TILLING mutant, the *TaPIF3* TILLING WT segregant, three independent overexpression lines (B3602 R6P1 and B3602 R7P1) including a selection marker only control (B3603 R2P3) and a wild-type Cadenza control. Eight individuals were planted for both the *TaPIF3* triple TILLING mutant and the *TaPIF3* TILLING WT segregant. Six plants were planted of each of the overexpression lines and the wild-type control. Plants were grown on a bench in the glasshouse measuring 145x324 cm. The growing bench was divided into 5 columns and 8 rows. Plants were set up

in a completely randomised design with latinisation to ensure even distribution of the six genotypes across the rows and columns. A completely randomised design was used to account for environmental variation within the glasshouse.

2.4.1 Phenotype characterisation

Once the plants had reached maturity and were dried down, the total number of tillers were counted and the following length measurements were taken for the three tallest tillers in cm; total tiller, flag leaf, ear, emergence above flag leaf, peduncle, internode 2, internode 3 and internode 4. Days to anthesis was recorded as when the anthers first extruded from the floret. Spikelet number of each ear from the three tallest tillers per plant was recorded.

2.4.2 Statistical analysis

The mean for each measurement across the three tillers and the individuals of a single genotype was calculated and used in an Analysis of Variance (ANOVA). ANOVA was applied to the data to consider the overall significance of differences between the lines using the F-test. Following the F-test, the standard error of the difference (SED) on the residual degrees of freedom (df) was used to compare pairs of lines of interest invoking the least significant difference (LSD) at the 5% level of significance. No transformation of data was required upon inspection of residual plots. The GenStat statistical package (17th edition, 2014, ©VSN International, Hemel Hempstead, UK) was used for the analysis.

2.5 Molecular Biology

Unless stated otherwise, all chemicals used in the creation of solutions, buffers, media etc. were purchased from Sigma Aldrich (Sigma-Aldrich Company Ltd., Dorset, UK).

2.5.1 DNA extraction

Genomic DNA (gDNA) was extracted from wheat leaf tissue using the DNAMITE^{TM} Plant Kit (Microzone, West Sussex, UK) according to the supplied protocol. Frozen plant tissue was

homogenised using a pestle and mortar. Pelleted DNA was resuspended in 30 µL sterile dH_2O. DNA concentration was quantified using a NanoDrop 1000 spectrophotometer (ThermoFisher Scientific, DE, USA). Extracted DNA was used immediately or stored at -20°C.

2.5.2 RNA extraction

Extraction of RNA was performed using the QIAGEN RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) according to manufacturer's instructions. A maximum of 100 mg plant tissue powder was decanted into liquid nitrogen-cooled 2 mL microcentrifuge tubes. Plant material was not allowed to thaw. RNA preparations were vortexed vigorously and kept at room temperature after the addition of 450 µL of lysis buffer RLT or RLC. 10 µL β-mercaptoethanol was added to 1 mL buffer RLT or RLC prior to use. The resulting lysate was transferred to a Qiagen QIAshredder spin column that had been placed in a 2 mL collection tube before centrifugation for 2 minutes at 15,700 x g. The supernatant was transferred to a new microcentrifuge tube before adding 0.5 volume of ethanol (96-100%). Samples were mixed well by pipetting before being transferred to an RNeasy Mini spin column. Spin columns were centrifuged at ≥ 8000 x g for 15 seconds. The flow through was discarded. 700 µL of wash buffer RW1 was subsequently added to the spin column to wash membrane-bound RNA. Buffer RW1 was passed through the spin column by spinning at ≥ 8000 x g for 15 seconds. The flow through was discarded. 500 µL of a mild washing buffer RPE was then added to the column which was spun at ≥ 8000 x g for 15 seconds. The flow through was discarded before a second 500 µL of buffer RPE was added. The centrifugation step was repeated and the flow through discarded. Residual ethanol was removed from the membrane of the column by spinning for a further 1 minute at ≥ 8000 x g. The collection tube of the spin column was replaced with a new, sterile 1.5 mL microcentrifuge tube. RNA was eluted by adding 30-50 µL of RNase-free water directly onto the membrane of the spin column and centrifuging for 1 minute at ≥ 8000 x g. The RNA concentration was quantified using a Nanodrop (2.5.3).

2.5.3 Quantification of DNA/RNA

The concentration and quality of nucleic acid samples was determined using a NanoDrop 1000 spectrophotometer (ThermoFisher Scientific, DE, USA) using the program ND-1000 V3.7.1. The instrument is first initialised by loading 2 µL of sterile 18 MΩ H_2O onto the sample pedestal. After initialisation, 2 µL of eluent was loaded onto the sample pedestal to use as blank reading. The concentration of nucleic acid in samples is determined by loading 2 µL of the sample onto the sample pedestal. The absorbance of the sample was measured over the range of 200-300 nm. Absorbance at 260 nm determines DNA/RNA quantity whilst the ratio of absorbance at 260 nm and 280 nm was used to assess the purity of DNA and RNA.

2.5.4 Reverse transcription

Total RNA was used to synthesize first strand complementary DNA (cDNA) using the Superscript III First-Strand Synthesis System (Invitrogen™ California, USA). Prior to cDNA synthesis, RNA was subject to a DNase treatment using RQ1 RNase-Free DNase (Promega, Wisconsin, USA) to degrade contaminating DNA in the RNA sample. A 2 µg aliquot of RNA was resuspended in 16 µL of RNase free water, 2 µL DNase and 2 µL DNase buffer. Samples were then incubated at 37°C for 1 hour. After the incubation period the reaction was terminated by the addition of 2 µL DNase STOP buffer and a second incubation at 65°C for 10 minutes. 15 µL of DNase treated RNA was transferred to a sterile microcentrifuge tube. 1 µL of 10mM dNTPs (consisting 10mM of each dATP, dTTP, dCTP and dGTP) and 1 µL 50 µM oligo dT primer were added to the RNA and mixed well by pipetting. The sample mix was subsequently incubated at 65°C for 5 minutes and then cooled to 4°C. The contents of the tube were collected by centrifugation before the addition of 4 µL 5X first-strand buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM $MgCl_2$), 1 µL 0.1M DTT, 1µL Superscript III reverse transcriptase and 1 µL RNaseOUT. Reaction components were mixed by pipetting and incubated for 60 minutes at 50 °C. To inactivate the reaction, samples were heated to 70 °C for 15 minutes. After 15 minutes newly synthesised cDNA was allowed to cool before storing at -20 °C.

2.5.5 Polymerase Chain Reaction (PCR) amplification of DNA

PCR reactions were used to amplify DNA from cDNA or genomic DNA templates. Primers used for PCR were designed using the bioinformatics platform Geneious version (8) (http://www.geneious.com, Kearse et al., 2012). Primer sequences are listed in tables S.1, S.2 and S.3. PCR reactions were performed in sterile PCR tubes and cycled using a PCR thermocycler. BioMix (Bioline, London, UK) and HotShot Diamond (Clent Life Science, Stourbridge, UK) are ready to use 2X reaction mastermixes that include a DNA polymerase and dNTPs. The volumes of PCR reagents used are given in table 2.1. The cycling conditions used for the different PCR reactions are given in table 2.2.

Table 2.1 Volumes of PCR reagents used for each PCR reaction.

Reagent	GoTaq polymerase	Phusion polymerase	Biomix	HotShot Diamond
5X Green GoTaq buffer	4µL	-	-	-
5X Phusion HF or GC buffer	-	4 µL	-	-
2X Biomix	-	-	10 µL	-
2X HotShot Diamond				10 µL
10 µM forward primer	0.5 µL	0.25 - 1 µL	0.4 µL	0.4 µL
10 µM reverse primer	0.5 µL	0.25 – 1 µL	0.4 µL	0.4 µL
10 mM dNTPs	0.4 µL	0.4 µL	-	-
$MgCl_2$ (25mM)	1.2 µL	-	-	-
DMSO	-	0.6 µL	-	-
GoTaq polymerase (5u/µL)	0.15 µL	-	-	-
Phusion polymerase	-	0.2 µL	-	-
Template DNA	variable	Variable	variable	variable
Sterile 18Ω H_2O	Up to 20 µL	Up to 20 µL	Up to 10 µL	Up to 20 µL

Table 2.2 Cycling conditions used for each PCR reaction

Step	GoTaq polymerase	Phusion polymerase	Biomix	HotShot Diamond
Initial denaturation	95°C, 2 minutes	98°C, 30 seconds	95°C, 4 minutes	95°C, 5 minutes
Denaturation	98°C, 10 seconds	98°C, 10 seconds	94°C, 25 seconds	95°C, 30 seconds
Annealing	Variable, 30 seconds	Variable, 30 seconds	Variable, 30 seconds	Variable, 30 seconds
Extension	72°C, 1min/kb	72°C, 15-30 sec/kb	72°C, 15-30 sec/kb	72°C, 15-30 sec/kb
Repeat cycle	35-40 cycles	35-40 cycles	35-40 cycles	35-40 cycles
Final extension	72°C, 5 minutes	72°C, 5 minutes	72°C, 5 minutes	72°C, 7 minutes
Storage	4°C, indefinite	4°C, indefinite	4°C, indefinite	4°C, indefinite

2.5.6 Agarose gel electrophoresis

PCR products and restriction enzyme digest products were assessed and separated using agarose gel electrophoresis. Samples were loaded onto a 1% (w/v) agarose TAE or TBE gel. Ethidium bromide or Gel Red (Biotium Cambridge Biosciences, UK) was used for DNA detection. 5 µL of GeneRuler 1 kb DNA Ladder (ThermoFisher Scientific, DE, USA) or 5 µL of 2-Log DNA Ladder (New England BioLabs) was loaded to assess product size. Loading buffer was added to samples requiring buffer before gel electrophoresis. Gels were run at 120 volts for 40-50 minutes in 1X TAE/TBE buffer before visualising under UV light.

2.5.7 Purification of DNA from agarose gel

Following gel electrophoresis, target DNA fragments were extracted from a 1% (w/v) agarose gel using the Qiagen Gel Extraction Kit according to the manufacturer's protocol. The DNA fragment was extracted from the gel using a clean, sharp scalpel whilst under UV light. Buffer QG was added to the gel piece at a 3:1 ratio before the sample was incubated at 50°C to dissolve the gel. One volume isopropanol was subsequently added before the sample was applied to a QIAquick spin column. To bind the DNA to the column the sample was centrifuged for 1 minute at 15,700 x g and the flow through discarded. A further 0.5 mL buffer QG was added to the spin column which was spun at 15,700 x g for 1 minute to remove traces of agarose. The flow through was discarded. To wash the column 0.75 mL of buffer PE was added to the column before the column was again centrifuged for 1 minute at 15,700 x g. Flow through was discarded. Re-spinning the column for a further 1 minute was carried out to remove ethanol remnants. The column was placed in a new, sterile microcentrifuge tube. 30-50 µL elution buffer was added directly to the column. To elute DNA the column was centrifuged at 15,700 x g for 1 minute.

2.5.8 Cloning

2.5.8.1 Cloning of PCR-amplified *TaPIF3* gene fragments for sequencing

PCR products amplified using GoTaq polymerase or BioMix were cloned into the pGEM-T Easy vector (Promega Corporation, Wisconsin, U.S.A.) following the manufacturer's

instructions. Concentrations of PCR products were determined prior to cloning (2.5.3). For ligation into pGEM-T Easy vector a 3:1 molar ratio of insert:vector was used. The calculation for the 3:1 ratio is as follows where A refers to the amount of vector DNA, B refers to the length of the insert in kb and C refers to the length of the vector in kb:

((A ng vector DNA x B kb insert DNA)/ C kb of vector DNA) x 3 = ng of insert DNA

Ligation reaction components are detailed in table 2.3. The ligation reaction was mixed by pipetting before incubating at 4°C overnight. After incubation, the reaction was stored at -20°C until transformation.

Table 2.3 Reaction components of pGEM-T Easy ligation reaction.

Reaction component	Standard reaction	Positive control reaction
2X Rapid Ligation Buffer	5 µL	5 µL
pGEM-T Easy vector (50ng)	1 µL	1 µL
PCR product	Variable	-
Control Insert DNA	-	2 µL
T4 DNA Ligase (3 Weiss units/µL)	1 µL	1 µL
Nuclease-free water	to final volume of 10 µL	to final volume of 10 µL

Blunt-ended PCR products were ligated into the pSC-B vector using the StrataClone Blunt PCR Cloning Kit (Agilent Technologies, California, USA). pSC-B supports blue/white screening and confers resistance to ampicillin and kanamycin. 3 µL of StrataClone Blunt Cloning Buffer was combined in order with 2 µL of PCR product (5-50 ng) and 1 µL StrataClone Blunt Vector Mix in a 1.5mL sterile microcentrifuge tube as detailed by the manufacturer's instructions. The reaction was incubated for 5 minutes at room temperature before being placed on ice and stored at -20°C until use.

2.5.8.2 Cloning of *TaPIF3* expression vectors

Cloning into Gateway entry vectors and the *pRRES125* vector was carried out by a restriction digest and ligation protocol. All restriction endonucleases used in this study were sourced from New England Biolabs UK (Hitchin, U.K.) or Promega (Wisconsin, U.S.A.). Restriction digest reactions were performed using the supplied buffers and following the recommended digest conditions. In all cases, restriction digests were performed using at least one sticky end cutter to ensure ligation of DNA fragments in the correct orientation. Digests of vector and insert DNA were typically incubated at 37°C for 2-3 hours. Products of each digest reaction were separated using gel electrophoresis (2.5.6) and extracted by the method detailed in section 2.5.7.

Ligation of the insert into the linearized vector was carried out using a T4 DNA ligase reaction. Typically, a 20 µL ligation reaction was performed using a 3:1 molar ratio of insert to vector DNA. Approximately 50 ng of vector was used for each ligase reaction. Ligation reactions were incubated at room temperature for approximately 2 hours. Ligations were subsequently transformed into DH5α competent cells (2.5.9) and positive transformants were selected based on antibiotic resistance. The success of each cloning step was analysed by a diagnostic digest of miniprep DNA and subsequent sequencing.

Recombination of *TaPIF3* from Gateway entry vectors into destination vectors was carried out using the Gateway™ LR Clonase™ Enzyme mix (ThermoFisher Scientific, DE, USA) following manufacturer's instructions. 150 ng of entry vector was combined with 150 ng of the destination vector and TE buffer (pH 8) to a final volume of 8 µL. 2 µL of LR Clonase II enzyme (previously thawed on ice) was added to the reaction mix. The reaction mix was vortexed briefly, centrifuged briefly to collect the contents and incubated at 25°C overnight. To terminate the reaction, 1 µL of proteinase K enzyme was added to the reaction. The reaction mix was then vortexed briefly and incubated at 37°C for 10 minutes. The LR reaction was subsequently transformed into DH5α *E.coli* cells as described in section 2.5.9.

2.5.9 Transformation of chemically competent *Escherichia coli*

2 µL of plasmid DNA was added to a 50 µL aliquot of chemically-competent *Escherichia coli* (*E.coli*) cells. The cells were mixed by gently flicking the transformation tube before incubating on ice for 20 minutes. The competent cells were subsequently heat-shocked by placing the tube in a 42°C waterbath for 45 seconds. After heat-shock, the cells were placed immediately on ice. 250 µL LB medium (1% (w/v) tryptone; 0.5% (w/v) yeast extract and 1% (w/v) NaCl) was added to the cells before incubation at 37 °C and vigorous agitation. After the incubation period, cells were spread onto freshly poured LB agar (LB medium in 15% (w/v) agar) containing antibiotic selection and, when stated, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) for blue/white screening. Selection plates were incubated at 37°C overnight. When transforming pSC-B vector, Cre recombinase expressing competent cells were used that are supplied with the StrataClone Blunt PCR Cloning Kit.

2.5.10 Selection of colonies for overnight culture

Single well-isolated colonies that had grown on LB selection plates were selected for plasmid DNA analysis. Selected colonies were swabbed using a sterile pipette tip or toothpick and inoculated in 5 mL LB broth containing antibiotic selection. Cultures were subsequently incubated at 37 °C overnight with agitation.

2.5.11 Extraction of Plasmid DNA

Plasmid DNA from an overnight culture was extracted using the Qiagen Mini-prep kit or the GeneJET Plasmid Mini-prep kit (ThermoFisher Scientific, DE, USA) following manufacturer's instructions. DNA was eluted using 50 µL elution buffer and the concentration of plasmid DNA was determined using the NanoDrop (2.5.3).

2.5.12 Diagnostic restriction enzyme digest

Restriction enzyme digestion was conducted to verify the correct insertion of DNA into a vector. 0.5 – 1 µg of plasmid DNA was added to 1 µL restriction enzyme 10X buffer. The reaction mixture was made up to a volume of 9 µL with 18 MΩ H_2O. 1 µL of restriction enzyme(s) (single or double digest) was added to reach a final volume of 10 µL. The manufacturer's instructions were followed to select the appropriate restriction enzyme 10X buffer for the restriction enzyme in use. The reaction was incubated at the temperature recommended for optimal enzyme activity for 1-4 hours. Products of the restriction enzyme digest were analysed on a 1% (w/v) agarose gel following gel electrophoresis.

2.5.13 Sequence analysis

Plasmid DNA was sent at a concentration of 100 ng/µL in a 1.5 mL microcentrifuge tube to either Eurofins Genomics (Luxembourg) or Source Bioscience (Nottingham, UK) for sequencing. Samples were diluted to the required concentration and sent in a sterile 1.5 mL microcentrifuge tube. Sequencing primers were selected according to the particular plasmid. Sequencing data was analysed using the bioinformatics platform Geneious (Aukland, New Zealand). The quality of the chromatogram files was inspected in all cases.

2.5.14 Genotyping

TILLING lines were genotyped using genotyping by sequencing. Genomic DNA was extracted from approximately 1g leaf tissue and used as template in a BioMix PCR reaction. PCR primers designed to amplify the homoeologue specific region of interest were used for the amplification. PCR products were purified using the QIAGEN QIAquick PCR Purification kit according to the supplied protocol and then sent for sequencing (2.5.13). Sequencing results were mapped to the WT *TaPIF3* sequence using the map to reference function in Geneious to identify lines containing the desired mutant allele.

2. 6 5′ Rapid amplification of cDNA ends (RACE)

2.6.1 Generating RACE-Ready cDNA

Synthesis of RACE-Ready cDNA was performed using the SMARTer RACE cDNA Amplification Kit (Clontech, California, USA). RACE-Ready cDNA synthesis was performed using RNA extracted from wheat seedlings (cv. Mercia) and wheat floral tissue (cv. Cadenza). In addition, RACE-Ready cDNA synthesis was performed using the control mouse heart total RNA (supplied by the manufacturer). A 1 µg aliquot of RNA was combined with 1 µL 12 µM oligo dT primer and made up to 3.75 µL with sterile H_2O in a 1.5 mL microcentrifuge tube. The sample was incubated at 72°C for 3 minutes then cooled to 42°C for 2 minutes. Tubes were spun for 10 seconds at 14,000 x g. 1µL of the SMARTer IIA oligo was added to each cDNA synthesis reaction. 5.25 µL of mastermix containing 2.0 µL 5X first-strand buffer, 1.0 µL 20 Mm DTT, 1.0 µL 10 mM dNTP mix, 0.25 µL RNase Inhibitor and 1.0 µL SMARTScribe reverse transcriptase was added to each cDNA reaction. Reactions were incubated at 42°C for 90 minutes and subsequently heated to 70°C for 10 minutes. The RACE-Ready cDNA was diluted in 100 µL Tricine-EDTA buffer and stored at -20°C.

2.6.2 5′ RACE PCR

5′ RACE PCR was performed using the SMARTer RACE cDNA Amplification Kit to amplify 5′ *TaPIF3* cDNA fragments. Gene specific primers for RACE PCR were designed using Geneious (see table S.2 for primer sequences). Primers were designed to be 23-28 nucleotides in length and have a $T_m \geq 65$°C. A control PCR reaction using mouse heart cDNA and the control 5′-RACE TFR primer was performed following manufacturer's instructions. RACE PCR reactions were performed in sterile PCR tubes and cycled using a PCR thermocycler. The volumes of PCR reagents used are shown in table 2.4 and the cycling conditions detailed in table 2.5.

Table 2.4 Volumes of reagents used for RACE PCR

Reagent	5’ RACE sample reaction	5’ RACE control reaction
10X Advantage 2 PCR buffer	5.0 µL	5.0 µL
dNTP (10 mM)	1.0 µL	1.0 µL
5’ RACE-Ready cDNA	2.5 µL	-
RACE-Ready mouse heart cDNA	-	2.5 µL
Universal Primer Mix	5.0 µL	5.0 µL
Gene specific primer (10 µM) (antisense)	1.0 µL	-
Control 5’-RACE TFR primer	-	1.0 µL
50X Advantage 2 polymerase mix	1.0 µL	1.0 µL
PCR-Grade water	34.5 µL	34.5 µL

Table 2.5 PCR cycling conditions for 5’ RACE reactions.

5 cycles:	
94°C	30 seconds
72°C	2 minutes
5 cycles:	
94°C	30 seconds
70°C	30 seconds
72°C	2 minutes
32 cycles:	
94°C	30 seconds
68°C	30 seconds
72°C	2 minutes

2.6.3 Nested PCR

Nested PCR was performed to improve specificity of the amplified 5' RACE product. 5 µL of primary PCR product was diluted with 245 µL of Tricine-EDTA buffer. Nested PCR was performed using 5 µL diluted primary PCR product as template in place of RACE-Ready cDNA. 1 µL of a nested universal primer (supplied by the manufacturer) and 1 µL of a nested gene specific primer were used for amplification. Volumes of 10X Advantage PCR buffer, dNTPs and 50X Advantage 2 polymerase mix were kept constant (table 2.4). Cycling conditions used for nested PCR are detailed in table 2.6.

Table 2.6 Cycling conditions used for nested PCR.

20 cycles:	
94°C	30 seconds
68°C	30 seconds
72°C	2 minutes

2.7 *Agrobacterium*-mediated transformation of Arabidopsis

2.7.1 *Agrobacterium* transformation

DNA was extracted from transformed *E.coli* using a Mini-prep kit according to the manufacturer's instructions and quantified by using a NanoDrop 1000 spectrophotometer (ThermoFisher Scientific, DE, USA). 2 µL of plasmid was then transferred to a 1.5 mL Eppendorf tube containing 50 µL GV3101 *Agrobacterium* cells. Following gentle mixing, cells were rapidly frozen by submersion in liquid nitrogen. Cells were then incubated at 37°C for 5 minutes before being shaken (200 rpm) for 2 hours in 1 mL LB broth. After the two hour incubation period, cells were centrifuged at 15,700 x g for 30 seconds and resuspended in 100 µL LB broth. Cells were subsequently spread on LB agar plates containing 25 µg/mL rifampicin, 30 µg/mL gentamycin and 100 µg/mL spectinomycin. They were then cultured at 28°C for 2 days. Two colonies that grew on selection plates were used as template in a

BioMix PCR reaction (2.5.5) using primers pGWB_F and TaPIF3_5'_R to confirm the presence of the plasmid (see table S.1 for primer sequences).

2.7.2 Arabidopsis transformation

Following confirmation of a positive PCR result, colonies were shaken in 10 mL LB broth containing 25 µg/mL rifampicin, 30 µg/mL gentamycin and 100 µg/mL spectinomycin at 28°C for 20 hours (200 rpm). 5 mL of the Agrobacterium culture was used to inoculate 200 mL LB broth containing 25 µg/mL rifampicin, 30 µg/mL gentamycin and 100 µg/mL spectinomycin. Cultures were incubated at 28°C for 16 hours with agitation (200 rpm) before centrifugation at 4,750 x *g* for 15 minutes. The supernatant was removed and the cells were resuspended in 400 mL dH_2O containing 5% (w/v) sucrose. 360 µL Silwet L77 per litre of culture was added to the Agrobacterium immediately prior to dipping the plant material. The aerial parts of approximately 6-week-old Arabidopsis plants were dipped in the Agrobacterium culture for 20 seconds with gentle agitation. Transformed plants were returned to the growth room in an area of low light for 24 hours. Seeds were harvested when dry.

2.7.3 Selection of transformants

T1 seeds were plated on ½ MS 1% (w/v) agar media containing 40 µg/mL hygromycin antibiotic. Prior to plating, seeds were surface sterilised in 70% (v/v) ethanol for 5 minutes followed by 15% (v/v) bleach for 10 minutes. Seeds were washed four times using sterile water. Plates were initially incubated at 4°C for 3 days, then treated with 110 µmol m^{-2} sec^{-1} white light (WL) at 23°C for 6 hours to induce germination. Subsequent to the WL treatment, plates were placed in the dark for four days. After 4 days, plates were observed for seedling growth to identify T1 transformants. Plates were returned to continuous WL 110 µmol m^{-2} sec^{-1} for 3 days to allow seedlings to mature. Seedlings that survived were transferred to soil and allowed to self-fertilise. T2 seed was plated on ½ MS 1% (w/v) agar media containing 40 µg/mL hygromycin. The survival ratio was noted to identify single insertion transformants. Lines identified as single insertion events were transferred to soil

and allowed to self fertilise. Homozygous T3 seedlings were identified by a 100% survival rate in the next generation.

2.8 Yeast two-hybrid assays

2.8.1 Yeast transformation

Plasmids were transformed into yeast competent cells using a heat shock protocol. 150 µL of *Mav203* competent yeast cells were combined with 1 µg of each plasmid, 2 µL of 10 mg/mL sheared salmon sperm DNA (ThermoFisher Scientific) and 350 µL 50% polyethylene glycol (PEG) and incubated at 30°C for 1 hour. The reaction was subsequently chilled on ice for 2-3 minutes and then centrifuged at 14,500 rpm for 1 minute in a table-top centrifuge. Cells were resuspended in 100 µL sterile distilled water. 100 µL of cells were spread onto SC-leu/-trp + 2% bacto-agar plates and incubated at 30°C for 48-72 hours.

2.8.2 Specific yeast two-hybrid assay

To test for a specific interaction, yeast transformed with the appropriate bait and prey plasmids were grown on SC media lacking leucine, tryptophan and histidine with increasing concentration of 3-amino-1,2,4-triazole (3-AT). Appropriate volumes of a 1M solution of 3-AT were added to SC-Leu/Trp/His media to produce final concentrations of 10 mM, 25 mM, 50 mM and 100 mM. Typically, one colony of yeast was inoculated into 200 µL sterile distilled water. 5 µL of this inoculation was then spotted onto plates in a grid format. Three technical replicates were spotted for each yeast strain. Once the culture had dried onto the plate, plates were incubated at 30°C for 48 hours. Assay plates were then photographed and scored for visible growth of each yeast strain.

Chapter 3. Identification of a *PIF3* orthologue in hexaploid wheat

3.1 Introduction

3.1.1 Sequencing of the wheat genome

The first objective of this research project was to identify a putative *PIF3* orthologue in the wheat genome. In recent years, a lot of work has been dedicated to providing the wheat research community with a high quality reference genome. The nature of the bread wheat genome however has impeded progress due to many factors. Firstly the wheat genome is very large, approximately 17 Gb, more than five times the size of the human genome. Secondly, the bread wheat genome contains a high number of repetitive regions making assembly of contiguous sequences (contigs) from sequenced strands of DNA very difficult (The International Wheat Genome Sequencing Consortium, 2014). Bread wheat also has an allohexaploid genome due to bread wheat deriving from three separate grasses that hybridised in two independent polyploidization events (Petersen et al, 2006). The three 'sub genomes' are termed A, B and D with each subgenome comprised of seven chromosomes. The consequence of having three separate sub genomes is that for each gene there are expected to be three homoeologues.

In 2014, significant progress in delivering a quality reference genome was made with the publication of the Chromosome Survey Sequence (CSS) for *Triticum aestivum* cv. Chinese Spring. This assembly was produced by the International Wheat Genome Sequencing Consortium (IWGSC) (The International Wheat Genome Sequencing Consortium, 2014). Due to the complexity of the bread wheat genome, the IWGSC project sequenced the three subgenomes one chromosome arm at a time. An exception to this was chromosome 3B which was isolated and sequenced as a complete chromosome. This made it possible to assign an assembled chromosome fragment to the correct subgenome. Each chromosome arm was sequenced to a depth of between 30x and 241x using Illumina sequencing technology (The International Wheat Genome Sequencing Consortium, 2014). The IWGSC estimated that the CCS assembly represented 61% of the bread wheat genome. Annotation of protein-coding gene sequences to the CSS assembly was based on comparisons to

annotated genes in related species such as rice and barley. In addition, mapping publically available wheat full-length cDNAs and RNA-seq data to the assembly facilitated the annotation of gene sequences. Based on the level of completion of the genome assembly the IWGSC estimated that the wheat genome contains 106,000 functional protein-coding genes. Each diploid subgenome is therefore estimated to contain between 32,000 and 38,000 genes. Gene distribution between the three subgenomes was found to be slightly uneven, with a greater number of gene loci on the B genome (35%) compared to the A (33%) and D (32%) genomes. Despite these differences, these numbers suggest that polyploidization events in wheat did not cause rapid genome changes or functional dominance of one subgenome over the others. (The International Wheat Genome Sequencing Consortium, 2014).

Despite an apparent equal distribution of gene loci between the ancestral diploid genomes, it has remained unclear whether the three subgenomes contribute equally or differentially to total wheat gene expression. The relative contribution of the three potential homoeoloci (A, B and D) to total gene expression was dissected in a comprehensive study using high coverage RNA sequencing (Leach et al, 2014). The study assessed the expression of 2,314 genes present on chromosomes 1 and 5. Of this set of genes, Leach et al (2014) estimated that at least 45% are expressed as three distinct homoeoloci in shoot and root tissues. However, the majority of genes expressed as three homoeologues demonstrate a bias whereby the expression of a single homoeologue dominates total gene expression. Leach et al (2014) also reported that no global bias towards expression of homoeoloci from a particular subgenome could be detected.

Shortly after the release of the IWGSC CSS, a new wheat genome assembly was released that claimed better coverage than the IWGSC CSS assembly (Chapman et al, 2015). This genome assembly is from the synthetic wheat variety W7984, a reconstitution of hexaploid wheat formed by hybridizing tetraploid wheat *Triticum turgidum* (AABB) with the diploid grass *Aegilops tauschii* (DD). Using a whole-genome shotgun (WGS) approach a *de novo* sequence assembly was constructed that represented 9.1 Gb of the ~ 17Gb hexaploid wheat genome. Of the 9.1 Gb assembly, 7.1 Gb is anchored to chromosomal locations. The coverage of the W7984 genome assembly was compared to the IWGSC CSS assembly by searching for six known sequences. Briefly, the three homoeologues of *Rht-1* and the

gibberellin biosynthetic enzyme *KAO* were searched for in the two assemblies. These six gene sequences can be found on six scaffolds of at least 20 kbp in the W7984 assembly. In contrast, in the IWGSC CSS assembly one gene could not be found and the other five are captured across multiple scaffolds (Chapman et al, 2015). Further analysis concluded that the protein coding regions captured by each assembly do not completely overlap. Therefore, the databases can be used complementarily to each other.

More recently, the TGACv1 sequence assembly was released (Clavijo et al, 2017). This genome assembly represented a new whole-genome shotgun sequence assembly of *Triticum aestivum* cv. Chinese Spring and claimed to represent ~80% of the 17 Gb genome. Nearly 3 million contigs greater than 500 bp were generated using the newly developed 'w2rap' assembly software that is designed to facilitate large and complex genomes. Long-range scaffolding was then completed using SOAPdenovo. This methodology attained an average scaffold length of 88.8 kb (N50 value of 88.8 kbp). This is 3.7 fold greater than that of the W7984 assembly and 30 fold greater than the IWGSC CSS assembly (Clavijo et al, 2017). When published, the TGACv1 sequence assembly therefore represented the most complete assembly available. In total, 104,091 high confidence genes were annotated to the assembly. The ability to perform BLAST searches against this new assembly is an invaluable resource to scientists working with wheat genetics.

In January 2017, the IWGSC made available the IWGSC RefSeq v1.0 wheat genome assembly. Prior to this, the IWGSC had released the IWGSC Whole Genome Assembly (WGA) v0.4 in June 2016. The IWGSC WGA v0.4 was comprised of Illumina short sequence reads that were assembled using the NRGene's DeNovoMAGIC™ software. This resulted in scaffolds that captured 14.5 Gb of the wheat genome with an N50 value of 7.0 Mb. The RefSeq v1.0 assembly is an integration of the IWGSC WGA v0.4 combined with additional resources developed by IWGSC members. Additional resources included physical maps for all chromosomes and sequenced bacterial artificial chromosomes (BACs). This integration has substantially increased the quality of the assembly and has increased the scaffold N50 to 22.8 Mb. In July 2018, the IWGSC RefSeq v1.0 was uploaded to Ensembl Plants, replacing the TGACv1 assembly. The progress made in sequencing the wheat genome has vastly increased the ability to identify and characterise genes in wheat.

The described wheat genome sequence assemblies were released throughout the duration of this project. At the time this project started, only the IWGSC CSS sequence assembly was available. Initial identification of a putative *PIF3* orthologue in the wheat genome was therefore completed using this database only. The W7984 assembly was released during the first year of this project. This database was used in parallel with the IWGSC CSS assembly to identify three homoeologues of a putative *PIF3* orthologue in wheat and to annotate gene models to the homoeologous sequences. The identification of a *PIF3* orthologue in wheat was completed prior to the release of the TGACv1 and the IWGSC RefSeq v1.0 wheat genome assemblies. Once the TGACv1 and the IWGSC RefSeq v1.0 databases became available, they were searched to confirm the identified sequences and to identify homoeologous sequences not present in the IWGSC CSS or W7984 assemblies.

3.1.2 Analysis of bHLH transcription factors in rice

There has been no published research that has characterised PIF-like bHLH transcription factors in wheat, although PIF proteins have been identified in other monocot species and this information can be used to help obtain candidate gene sequences for a wheat *PIF3* orthologue (*TaPIF3*). Rice, in addition to being an important food crop, is often used as a model species for the monocots. Rice has a considerably smaller genome than other cereals, approximately 420-450 Mb in size (Goff, 1999) for which a high quality genome reference is available (Kawahara et al, 2013). Co-linearity of gene order is thought to exist between rice and other major crop species such as wheat and maize after comparisons of genetic maps (Ahn et al, 1993). In addition to the conservation of gene location, studies have demonstrated considerable homology between gene families in cereal species (Goff, 1999). Information can therefore be elucidated about the functions of orthologous genes/gene families in other cereal species from the analysis of the rice genome.

In 2006, a genome-wide analysis of basic helix-loop-helix (bHLH) transcription factors in rice and Arabidopsis was completed (Li et al, 2006). In this study, 167 rice bHLH genes (*OsbHLH*) were identified using a criteria to define a bHLH protein developed by Atchley et al (1999). Principally, the bHLH motif contains 19 conserved residues, five of which are in the basic region, five in the first helix, one in the connecting loop and eight in the second helix

(Atchley et al, 1999). Using TBLASTN against the rice genome database, Li et al (2006) extracted all putative bHLH proteins that had 11 or more of the 19 conserved residues. Phylogenetic analysis revealed clustering of *AtbHLH* and *OsbHLH* genes into 25 subfamilies. Of particular interest to this project was subfamily E that grouped *OsbHLH* genes with Arabidopsis sequences such as *PIF3, PIF4,* and *PIF6*. In total, 10 *OsbHLH* genes were clustered into subfamily E, each of which has an OsbHLH number according to its position in the multiple sequence alignment. In addition to an OsbHLH number, the gene name derived from the Rice Genome Annotation Project located at Michigan State University (MSU) was provided. Genes annotated by MSU have a locus identifier in the form LOC_OsXXgYYYYY where LOC_Os refers to *the Oryza sativa* locus, XX refers to the chromosome, g refers to gene and YYYYY refers to the gene order on the chromosome. The phylogenetic analysis identified two rice genes, LOC_Os01g18290 and LOC_Os05g04740, as putative orthologues of Arabidopsis *PIF3*.

In a separate study by Nakamura et al (2007) a set of highly conserved *OsPIL* (*Phytochrome Interacting Factor 3-Like*) genes were identified and characterised. In total six rice genes were identified that encoded proteins with high sequence homology to Arabidopsis PIFs and contained both a bHLH domain and, critically, an APB motif located near the N terminus. These six genes were termed *OsPIL11* to *OsPIL16*. OsPIL15 and OsPIL16 were inferred to be highly similar to PIF3. The set of *OsPIL* genes identified by Nakamura et al (2007) have annotated gene ID codes that have been adopted from the Rice Annotation Project (RAP) (table3.1).

3.2 Results

3.2.1 Identification of putative *PIF3* orthologues in wheat

3.2.1.1 Sequence analysis of PHTOCHROME INTERACTING FACTOR-LIKE genes in rice.

To identify a *PIF3* orthologue in the wheat genome, rice *PIL* sequences were used to search the wheat database. To obtain *PIL* sequences in rice, the genomic and CDS of the six *OsPIL* members were downloaded and imported into Geneious version (7) (http://www.geneious.com, Kearse et al, 2012) from Ensembl plants using the RAP gene ID. The genomic and CDS of the 10 genes from the Li et al (2006) study were obtained from the Rice Genome Annotation Project (MSU) (http://rice.plantbiology.msu.edu/) by entering the locus identifier into the locus search tool. Comparisons between the two sets of rice genes, the ten MSU annotated genes and the six rice *PIL* genes, were made by producing multiple and pairwise sequence alignments. As anticipated, the sequence alignments revealed in many cases a gene sequence from one study/annotation matched identically to a gene sequence from the other study/annotation (Table 3.1). The CDS for *OsPIL16*, *OsPIL11* and *OsPIL13* share 100% sequence identity to LOC_Os05g04740, LOC_Os12g41650 and LOC_Os03g56950 respectively. *OsPIL12* and *OsPIL14* represent the same genes as LOC_Os03g43810 and LOC_Os07g05010 respectively despite differences in annotation at the 3' end. The CDS of *OsPIL15* extends 281 bp upstream of LOC_Os01g18290, however the two sequences subsequently share 1914 consecutive bases. A common set of rice genes have therefore been independently identified as being closely related to Arabidopsis *PIF* sequences. From now on, this project will adopt the nomenclature for rice PIL genes designated by Nakamura et al (2007).

Table 3.1 Nomenclature of *PIL* genes in rice. A set of common rice genes were identified as being closely related to Arabidopsis *PIFs*. *OsPIL* genes identified by Nakamura et al (2007) have annotated gene IDs indicated in parentheses adopted from the Rice Annotation Project (RAP). Genes identified by Li et al (2006) have gene names derived from the Rice Genome Annotation Project located at Michigan State University (MSU). Genes highlighted in red were grouped with *PIF3* in each study.

RAP annotation	MSU annotation
OsPIL11 (Os12g0610200)	LOC_Os12g41650
OsPIL12 (Os03g0639300)	LOC_Os03g43810
OsPIL13 (Os03g0782500)	LOC_Os03g56950
OsPIL14 (Os07g0143200)	LOC_Os07g05010
OsPIL15 (Os01g0286100)	LOC_Os01g18290
OsPIL16 (Os05g0139100)	LOC_Os05g04740

3.2.1.2 Wheat orthologues of *OsPIL15* and *OsPIL16*

Initial identification of a *PIF3* orthologue was completed using the IWGSC CSS assembly. Conducting a search for orthologues of the six *OsPIL* genes against the IWGSC genome assembly in Ensembl plants revealed a number of candidate wheat genes. In particular, genes Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 became promising candidates for a *PIF3* orthologue in wheat. Both Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 were identified in a BLASTN search of the IWGSC database using the CDS of Arabidopsis *PIF3* and *OsPIL16*. The CDS of Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 and corresponding genomic scaffolds IWGSC_CSS_1BS_scaff_3473017 and IWGSC_CSS_1DS_scaff_1896224 were imported into Geneious. Traes_1BS_D1FCBFBE8 is comprised of four exons and is predicted to encode a protein of 506 amino acids in length whilst Traes_1DS_55E91A134 consists of 5 exons and encodes a protein of 655 amino acids (for gene architecture see Figure 3.1). The predicted proteins that the two wheat genes encode contain a bHLH domain with high sequence homology to Arabidopsis PIF1, PIF3, PIF4 and PIF5 (figure 3.1). The CDS of Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 will be checked by RNA-Seq mapping due to signs of incorrect annotation. For instance, the start site of Traes_1DS_55E91A134 is annotated as a CCA codon encoding a proline residue.

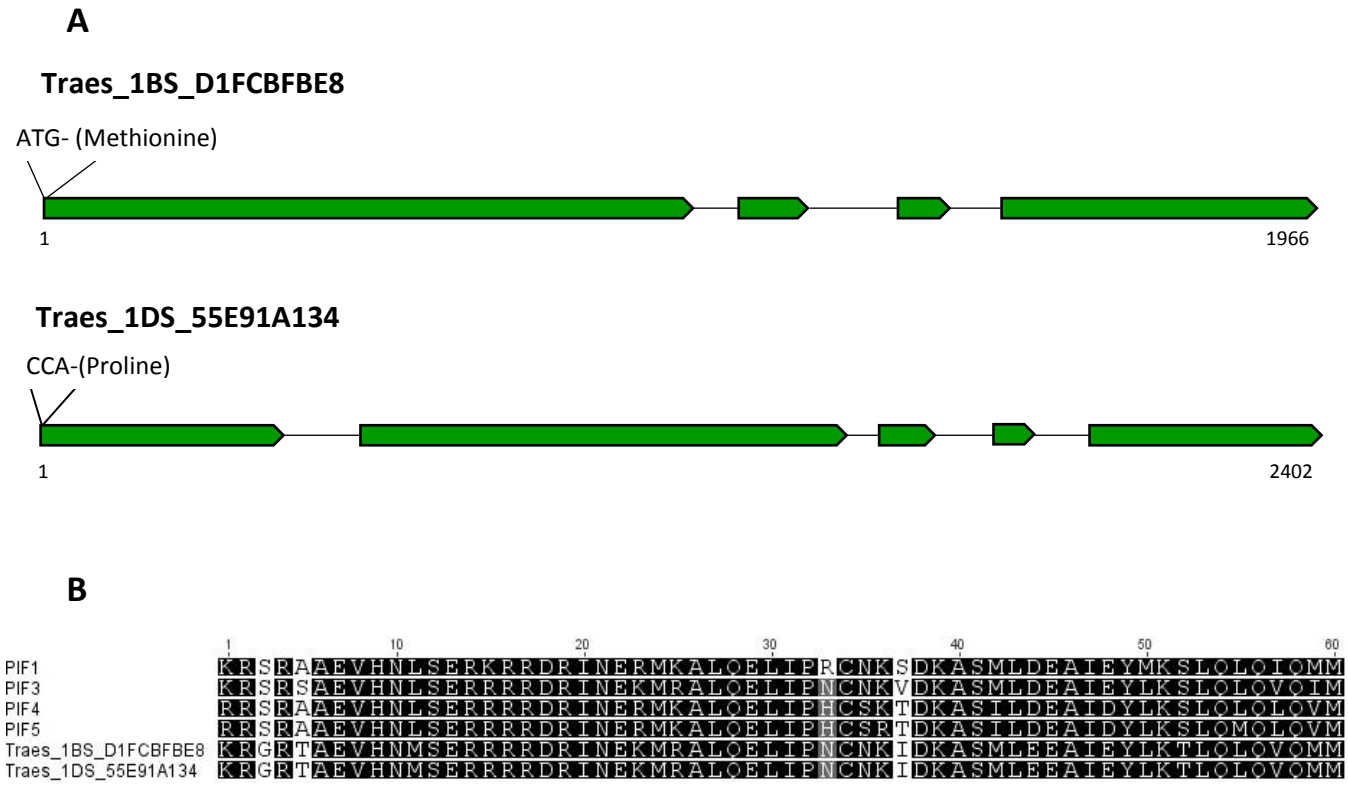


Figure 3.1 Schematic representation of the Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 genes. (A) Solid green arrows indicate exon coding sequences. A solid line represents non-coding introns. The putative translational start site is highlighted for each gene model and sequence length is indicated in bp. (B) A multiple sequence alignment of the bHLH domains of Arabidopsis PIF members, Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134.

Functional analysis of the Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 protein sequences was performed by entering the amino acid sequences into InterPro (Mitchell et al, 2015). InterPro allows the user to enter a protein sequence of interest and provides a functional analysis of that sequence by predicting protein domains and important sites of the sequence. The InterPro profiles of OsPIL16, Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 matched almost exactly, grouping all three genes in the gene family 'Basic helix-loop-helix (bHLH) transcription factors ALC-like, plant (IPR031066)'. In addition, all three proteins contained a Myc-type, basic helix-loop-helix domain (IPR011598) and were further classified into a STEROL REGULATORY ELEMENT-BINDING PROTEIN family (PTHR12565). In total, the IWGSC database contained 355 genes that encode a protein

containing a Myc-type, basic helix-loop-helix domain (IPR011598). Only a small number of these 355 proteins are likely to represent PIFs such as they do in Arabidopsis. Searching the same database revealed 90 wheat genes that encode proteins with the PTHR12565 signature. In addition to OsPIL16, Traes_1BS_D1FCBFBE8, Traes_1DS_55E91A134 and Arabidopsis PIF1, PIF3, PIF4, PIF5 and PIF6 were all grouped as PTHR12565 proteins. Sterol regulatory element-binding proteins (SREBPs) are recognised as transcription factors that activate lipid synthesis (Shimano, 2001). Of interest is that, SREBPs contain a bHLH domain and have the ability to bind to and activate palindromic E-box containing promoters. As reviewed previously, the PIFs bind to a palindromic hexanucleotide sequence called the G-box, which is representative of the more general E-box motif. The PTHR12565 signature therefore became a promising target to search for additional PIF3-like genes in the wheat genome.

3.2.1.3 Identification of wheat scaffolds

Wheat gene sequences that encoded for a protein with the PTHR12565 signature were searched for and imported into Geneious using BioMart through Ensembl plants. The database was selected to 'Plant Mart' and the dataset 'Triticum aestivum genes (IWGSC1.0 + popseq(2.2))' was chosen. PTHR12565 was entered into protein domains and families in the filter tool to restrict the output to only genes associated with this signature. In total 90 genes were identified and imported into Geneious with the corresponding IWGSC scaffolds.

To identify the regions of the 90 scaffolds that are expressed and to confirm intron/exon boundaries, RNA-Seq reads were mapped to the wheat scaffolds. Leaf and stem RNA-Seq data from hexaploid wheat cv. Chinese Spring was used for RNA-Seq mapping. RNA-Seq data was generated by INRA GDEC and is publicly available to download at URGI. The libraries are described in Pingault et al (2015). A single folder containing all 90 wheat scaffolds in fasta format was imported into the bioinformatics platform Galaxy**.** The RNA sequence reads were then mapped to the imported reference genome using the read mapping algorithm TopHat (Kim et al, 2013). Settings were kept to default with the exception that the data type was changed from single-end to paired-end data. Once the mapping was completed the data was merged and imported into Geneious. Mapping of RNA-Seq reads to the 90

scaffolds gave a visualisation of coding-sequences on the genomic scaffolds. This allowed the annotation of the identified wheat genes to be confirmed or edited.

The mapping of RNA-Seq data to scaffold IWGSC_CSS_1BS_scaff_3473017 suggested the annotation of Traes_1BS_D1FCBFBE8 was incorrect at the 5' end of the gene. RNA-Seq reads did not map to the ATG translational start codon that had been annotated for Traes_1BS_D1FCBFBE8. Exon 1 of Traes_1BS_D1FCBFBE8 was therefore reduced by 14 bases to align with the RNA-Seq reads leaving the gene without a start site (figure 3.2). RNA reads continued to map to the genomic scaffold approximately 130 bp upstream of the edited exon 1. Interestingly, the first two bases supported by RNA-Seq reads upstream of exon 1 were AT and the first base of the edited exon 1 was now a G therefore forming an ATG codon split by an intron (figure 3.3). The translation of this novel gene annotation yields a predicted protein of 502 amino acids. This annotation of Traes_1BS_D1FCBFBE8 will subsequently be referred to as *TaPIF3-B1*.

The annotation of Traes_1DS_55E91A134 also produced a gene model lacking a conventional ATG start site. RNA-Seq mapping to scaffold IWGSC_CSS_1DS_scaff_1896224 confirmed the gene annotation of Traes_1DS_55E91A134 except for the 5' end. RNA reads did not map to the 5' end of the first exon. The first exon coding sequence was therefore restricted leaving just two bases (AT) at the start of the gene (figure 3.2). The start codon is completed with a G after the first intron. This new gene model will be referred to as *TaPIF3-D1*.

To exclude the possibility of alternative upstream ATG translational start sites for *TaPIF3-B1* and *TaPIF3-D1*, the presence of alternative ATG codons was investigated in the regions upstream from the annotated translational start sites. Scaffold IWGSC_CSS_1BS_scaff_3473017 extends 4040 bp upstream of the annotated start site of *TaPIF3-B1*. RNA-Seq reads cease mapping to the genomic scaffold approximately 250-300 bp upstream of the annotated ATG start codon. A single, alternative ATG codon is located in this 250-300 bp region supported by RNA-Seq reads. The translation of this putative gene model however revealed that this ATG codon is out of frame and introduces multiple stop codons located throughout the CDS. Further potential upstream ATG start codons could not be found.

Scaffold IWGSC_CSS_1DS_1896224 extends 2874 bp upstream of the annotated start site for *TaPIF3-D1*. Similarly to *TaPIF3-B1*, RNA reads cease mapping to the genomic scaffold approximately 250-300bp upstream of the current start site. A search for alternative ATG codons within this 250-300bp region identified one ATG codon located 275bp upstream. Annotating the described ATG codon as the start site for *TaPIF3-D1* however alters the open reading frame introducing a premature stop codon.

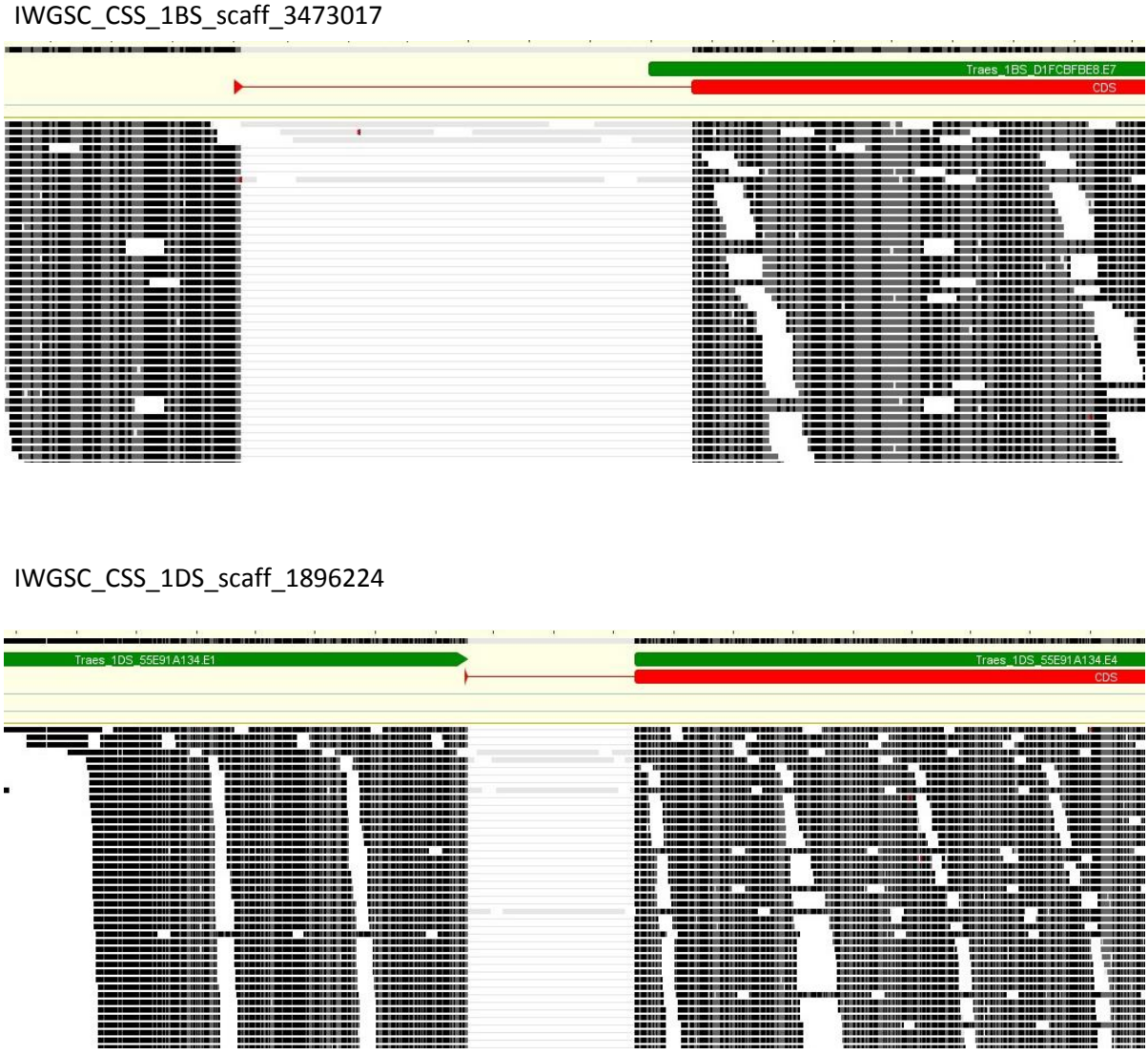


Figure 3.2 RNA-Seq reads mapped to scaffolds IWGSC_CSS_1BS_scaff_3473017 and IWGSC_CSS_1DS_scaff_1896224. RNA reads were mapped to a set of 90 IWGSC CSS scaffolds. RNA reads (in black) map to the coding sequences of the scaffold. Annotation of Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 can be seen in green. New gene annotation of *TaPIF3-B1* and *TaPIF3-D1* is represented in red.

A third homoeologue, *TaPIF3-A1*, could not be identified by searching the IWGSC wheat database or by analysing the wheat scaffold sequences extracted from biomart. During the analysis, Chapman et al (2015) released the W7984 sequence assembly. PIF-like sequences were searched for in this new genome assembly by performing a BLASTN search using the *TaPIF3-B1* and *TaPIF3-D1* gene sequences. Three genomic scaffolds were obtained that contained full-length sequences, two that matched the IWGSC scaffold sequences of the B and D homoeologue and a third, scaffold 2481858, that presumably represented the A homoeologue. Mapping of IWGSC sequencing reads from chromosome 1 of genome A to the third scaffold provided evidence that the A homoeologue of *TaPIF3* had been obtained. To correctly annotate a gene model to the A scaffold, RNA-Seq reads were mapped to all three scaffolds obtained from the new genome assembly. A coding sequence of 1540 bp was annotated to the A scaffold. The architecture of the gene looks to be conserved with the B and D homoeologue with an ATG start codon split by an intron (see figure 3.3). This gene model will subsequently be referred to as *TaPIF3-A1*.

Alternative translational start codons for *TaPIF3-A1* were searched for upstream of the annotated CDS using the same methodology for *TaPIF3-B1* and *TaPIF3-D1*. Scaffold 2481858 extends only 640bp upstream of the start site of *TaPIF3-A1*. Mapping of RNA-Seq reads to this scaffold shows that RNA reads map to the region 250-300bp directly upstream of the ATG start site . Within this 250-300bp region, one alternative ATG codon is present. However, this codon is not in the same reading frame as *TaPIF3*.

Aligning the cDNA sequences of the three homoeologues confirmed high sequence homology between the three homoeologues (table 3.2). The most significant difference between the three sequences was an 18bp insertion of repeated sequence in exon 2 of the A homoeologue.

Table 3.2 Sequence identities between *TaPIF3* homoeologues. The sequence identities of the CDS between each pair of *TaPIF3* homoeologues is shown.

	TaPIF3-A1	*TaPIF3-B1*	*TaPIF3-D1*
TaPIF3-A1	-	94.96%	94.89%
TaPIF3-B1	94.96%	-	97.42%
TaPIF3-D1	94.89%	97.42%	-

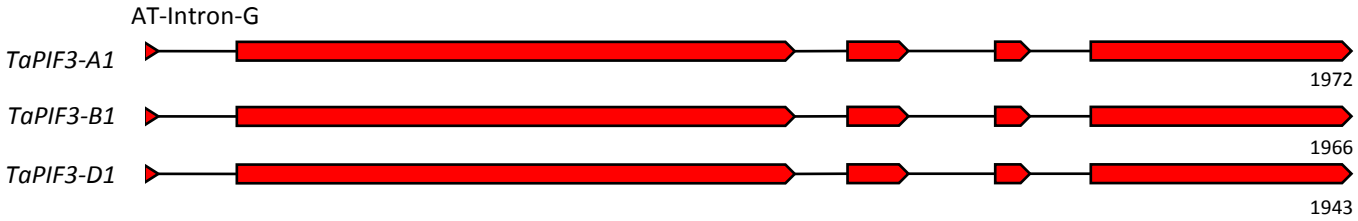


Figure 3.3 Intron/exon architecture of three *TaPIF3* homoeologues. Solid red arrows indicate exon coding sequences. Introns are annotated by a solid line. Length of the genomic sequences for each homoeologue are given in bp. The novel translational start codon split by an intron is indicated.

3.2.1.4 Evidence to support 5′ annotation of *TaPIF3*

The mapping of RNA-Seq reads to IWGSC and W7984 genomic scaffolds led to the current annotation of the three homoeologues of *TaPIF3* (figure 3.3). All three genes were annotated with an ATG start codon that was split by an intron between the T at position +2 and the G at position +3. The described IWGSC and W7984 scaffolds however only extend a short distance upstream of the three *TaPIF3* homoeologues. To eliminate the possibility of additional coding sequence upstream of the current *TaPIF3* annotation it was necessary to obtain a set of longer, more complete genomic scaffolds. Following the release of the TGAC v1 assembly the *TaPIF3* genomic sequences were used to BLAST against the newly released TGAC v1 assembly. The results of the BLAST search identified three genomic scaffolds: TGACv1_scaffold_019045_1AS, TGACv1_scaffold_049601_1BS and

TGACv1_scaffold_080393_1DS. Pairwise sequence alignments between the TGAC v1 scaffolds and IWGSC/W7984 scaffolds confirmed that the TGACv1 scaffolds corresponded to the current set of scaffolds for *TaPIF3*. TGACv1_scaffold_019045_1AS extends 9,245 bp upstream of the ATG start codon of *TaPIF3-A1* therefore greatly extending the current W7984 scaffold. In contrast, TGACv1_scaffold_049601_1BS only extends 46 bp upstream of scaffold IWGSC_CSS_1BS_scaff_3473017. This results in a genomic scaffold for the B fhomoeologue that extends 4,085 bp upstream of the annotated start site. An alignment between TGACv1_scaffold_080393_1DS and IWGSC_CSS_1DS_scaff_1896224 failed to generate an extended scaffold for the D homeologue of *TaPIF3*. TGACv1_scaffold_080393_1DS contains a region of unknown sequence approximately 380 bp upstream of the annotated CDS. IWGSC_CSS_1DS_scaff_1896224 therefore remained the most complete scaffold for *TaPIF3-D1.*

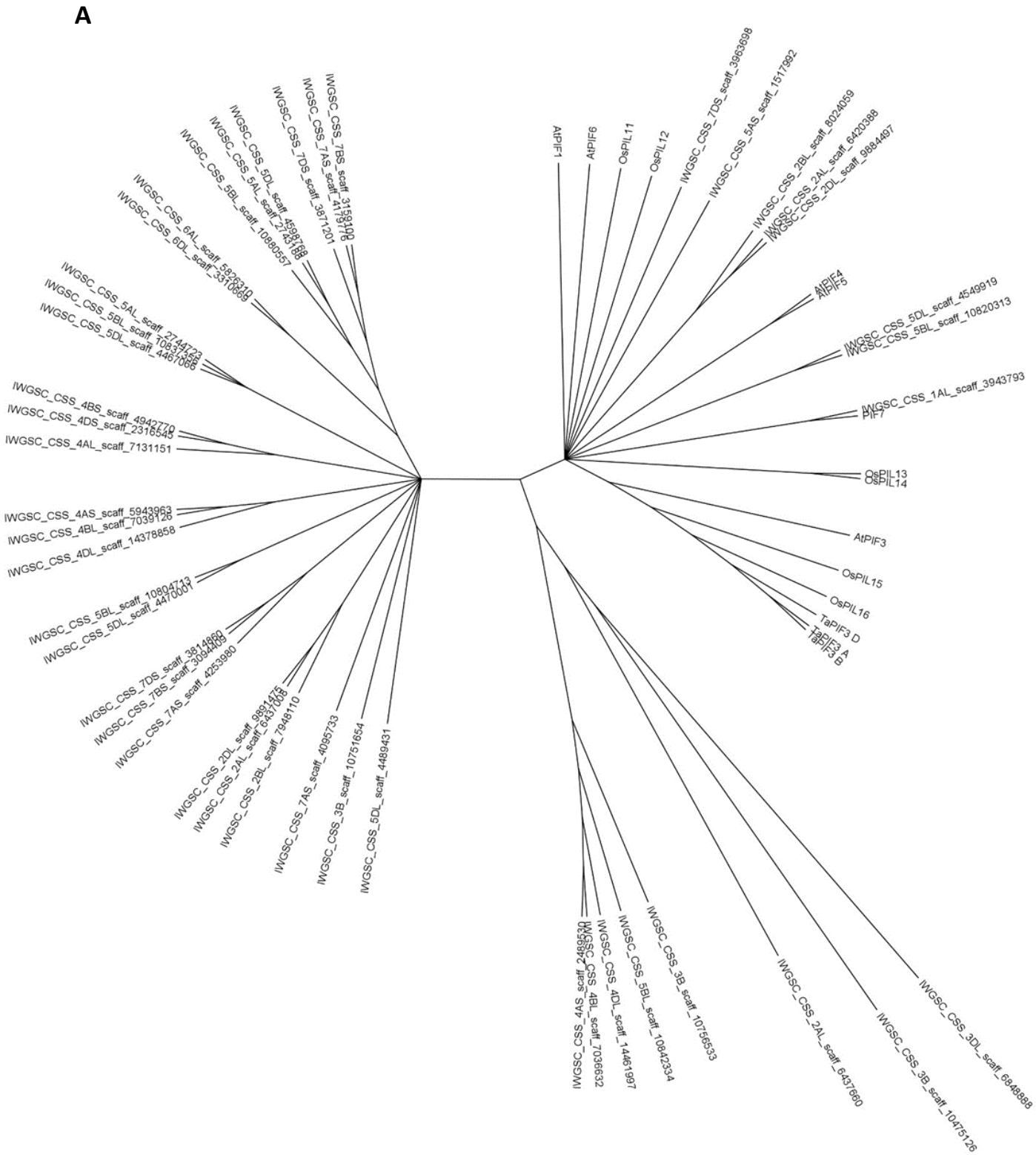
To map RNA-Seq reads to the extended A homoeologue, scaffold TGACv1_scaffold_019045_1AS, TGACv1_scaffold_049601_1BS and scaffold IWGSC_CSS_1DS_scaff_1896224 were imported into Galaxy. RNA-Seq reads described in section 3.2.1.3 were mapped to the three scaffolds using the read-mapping algorithm TopHat (Kim et al, 2013). RNA reads mapped to the TGACv1_scaffold_019045_1AS approximately 250-300bp upstream of the annotated *TaPIF3* start site. RNA-Seq reads did not map to the genomic scaffold further upstream of this 250-300 bp region indicating that the current annotation of *TaPIF3-A1* represents the full length CDS. No further information could be elucidated about the annotation of the B or D homoeologues of *TaPIF3*.

3.2.1.5 Phylogenetic analysis of PIF-like sequences in wheat

Editing of wheat gene annotation using RNA-Seq mapping provided a final set of candidate PIF-like sequences in wheat. Phylogenetic analysis was subsequently used to determine which sequences are most closely related to Arabidopsis PIF3. In addition, performing this analysis could identify putative orthologues of other Arabidopsis *PIFs*. Failure to find a translational start codon for four of the wheat genes meant that these sequences were deleted from the set. The CDS of the identified wheat genes were grouped with the six rice *OsPIL* sequences listed in table 3.1 and the gene sequences of Arabidopsis *PIF1*, *PIF3*, *PIF4*,

PIF5, *PIF6* and *PIF7*. The final list of coding sequences was translated in Geneious, and a multiple sequence alignment of the protein sequences was generated. Prior to the alignment, the list of protein sequences was refined by deleting any sequence less than 200 amino acids in length. The smallest PIF protein sequence in Arabidopsis is comprised of 364 amino acids and therefore 200 amino acids was chosen as the threshold.

Phylogenetic analysis should only be performed on residues with positional homology. It is therefore essential that protein sequences are correctly aligned relative to one another. Insertions, deletions and substitutions in the amino acid sequences complicate the production of a good multiple sequence alignment. Due to the high degree of sequence conservation observed in the bHLH domain of these sequences it was decided that the alignment should be restricted to the bHLH domain. Sequence conservation outside the bHLH domain was highlighted to be minimal through observation of full-length multiple sequence alignments. Initially a multiple sequence alignment was generated using the full-length PIF-like protein sequences using the MUSCLE alignment tool in Geneious. The alignment was manually edited to remove sequences with low sequence similarity to the consensus. The alignment was then constrained to the bHLH domain defined by Arabidopsis PIF3 (figure S.2). A phylogenetic tree was constructed based on the alignment of the bHLH domains of the Arabidopsis, rice and wheat protein sequences. The phylogenetic tree was constructed using the Neighbour-Joining (NJ) distance based algorithm and the Jukes-Cantor genetic distance model with 500 bootstrap replicates (figure 3.4). The majority of wheat sequences formed separate groups, often grouping with sequences assumed to represent homoeologues. PIF3 however groups with OsPIL15, OsPIL16 and the three homoeologues of TaPIF3. This indicates that the TaPIF3 protein is closely related to PIF3 and that *TaPIF3* is likely to represent the orthologous gene to *PIF3*.


A
IWGSC_CSS_7BS_scaff_3159100
IWGSC_CSS_7AS_scaff_4179178
IWGSC_CSS_7DS_scaff_3871201
IWGSC_CSS_5DL_scaff_4598788
IWGSC_CSS_5AL_scaff_2743788
IWGSC_CSS_5BL_scaff_10880557
IWGSC_CSS_6AL_scaff_5823319
IWGSC_CSS_6DL_scaff_3310669
IWGSC_CSS_5AL_scaff_2744723
IWGSC_CSS_5BL_scaff_10831353
IWGSC_CSS_5DL_scaff_4467066
IWGSC_CSS_4BS_scaff_4942770
IWGSC_CSS_4DS_scaff_2316545
IWGSC_CSS_4AL_scaff_7131151
IWGSC_CSS_4AS_scaff_5943963
IWGSC_CSS_4BL_scaff_7039126
IWGSC_CSS_4DL_scaff_14378858
IWGSC_CSS_5BL_scaff_10804713
IWGSC_CSS_5DL_scaff_4470007
IWGSC_CSS_7DS_scaff_3814860
IWGSC_CSS_7BS_scaff_3094409
IWGSC_CSS_7AS_scaff_4253980
IWGSC_CSS_2DL_scaff_9891475
IWGSC_CSS_2AL_scaff_6437008
IWGSC_CSS_2BL_scaff_7948110
IWGSC_CSS_7AS_scaff_4095733
IWGSC_CSS_3B_scaff_10751654
IWGSC_CSS_5DL_scaff_4489431
AtPIF1
AtPIF6
OsPIL11
OsPIL12
IWGSC_CSS_7DS_scaff_3963698
IWGSC_CSS_5AS_scaff_1517992
IWGSC_CSS_2BL_scaff_8024059
IWGSC_CSS_2AL_scaff_6420388
IWGSC_CSS_2DL_scaff_9884497
AtPIF4
AtPIF5
IWGSC_CSS_5DL_scaff_4549919
IWGSC_CSS_5BL_scaff_10820313
IWGSC_CSS_1AL_scaff_3943793
PIF7
OsPIL13
OsPIL14
AtPIF3
OsPIL15
OsPIL16
TaPIF3.D
TaPIF3
TaPIF3.B
IWGSC_CSS_4AS_scaff_2489530
IWGSC_CSS_4BL_scaff_7036632
IWGSC_CSS_4DL_scaff_14461997
IWGSC_CSS_5BL_scaff_10842334
IWGSC_CSS_3B_scaff_10756533
IWGSC_CSS_2AL_scaff_6437660
IWGSC_CSS_3B_scaff_10475126
IWGSC_CSS_3DL_scaff_6848889

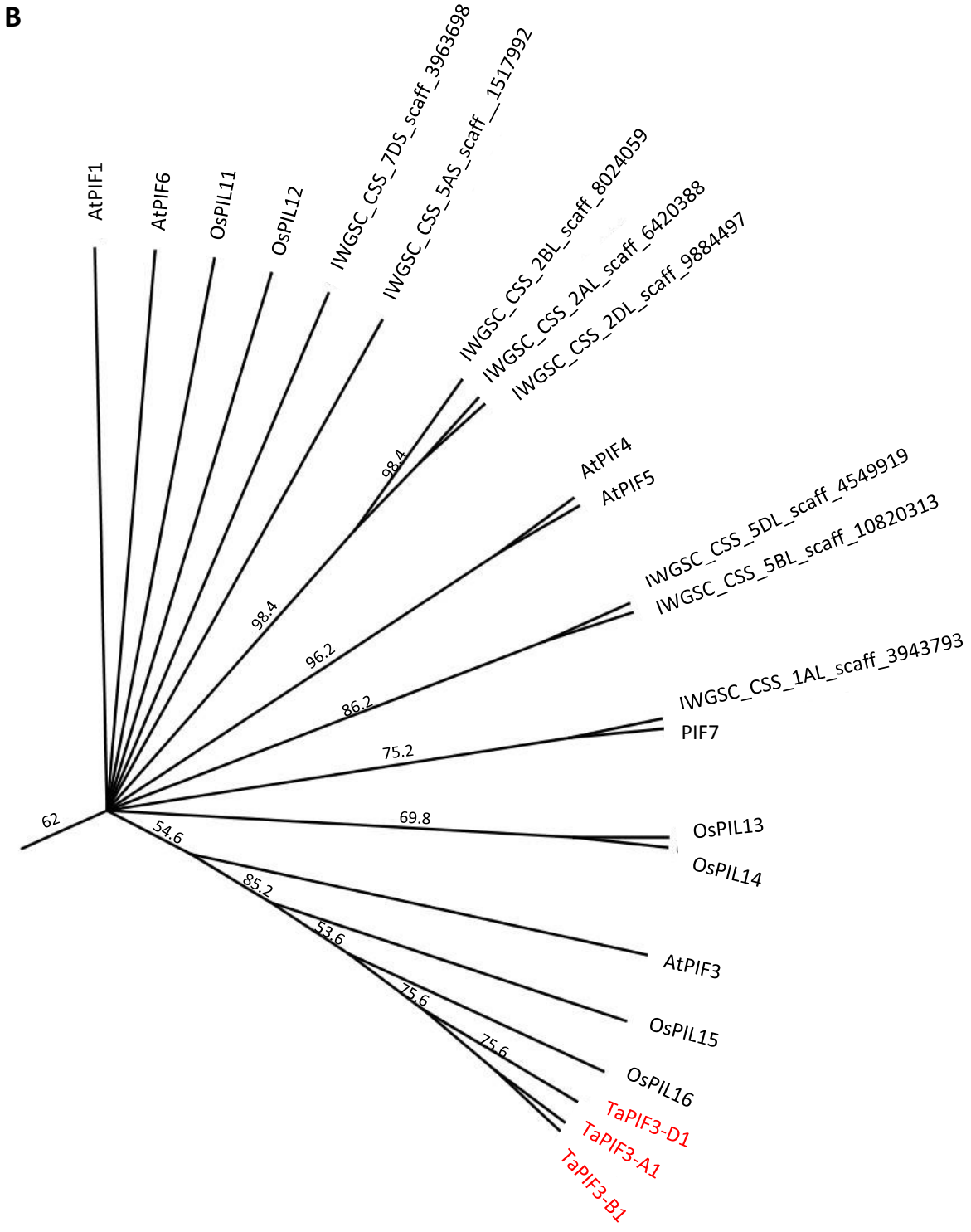


Figure 3.4 Phylogenetic analysis of identified wheat sequences. (A) A neighbour joining tree was produced using the amino acid sequence of the bHLH domain of identified wheat sequences, Arabidopsis PIFs and rice PILs. Wheat sequences are named by the genomic scaffold that gene is located on. Tree was constructed using 500 replicates. (B) Subtree of phylogenetic tree shown in figure 3.4 (A). TaPIF3 homoeologues are coloured in red. Consensus support (%) values are shown on the branches.

3.2.1.6 Identification of APB domains in TaPIF3 proteins

A defining component of Arabidopsis PIF proteins is a conserved Active Phytochrome B-binding motif (APB) identified by Khanna et al (2004), which is located at the N-terminus. Residues that define the APB domain were shown to be conserved in the OsPIL proteins identified by Nakamura et al (2007). In addition, the APB domain is conserved in ZmPIF3.1 and ZmPIF3.2 as illustrated in a multiple sequence alignment with PIF3 (Kumar et al, 2016). To identify a similar motif in the TaPIF3 proteins, a sequence alignment was generated using the amino acid sequences of Arabidopsis PIFs, OsPIL15, OsPIL16, ZmPIF3.1, ZmPIF3.2 and the three homoeologues of TaPIF3. A sequence alignment of the N-terminus of these proteins would also provide additional evidence to whether the 5' annotation of *TaPIF3* is correct. In Arabidopsis, residues E41, L42, G47 and Q48 marked by an asterisk in figure 3.5 were shown to be invariant for PHYB binding (Khanna et al, 2004). Of these four residues, the alignment illustrates that the first three amino acids are conserved in all of the Arabidopsis, rice, maize and wheat sequences. However, the glutamine residue at position 48 is not conserved in OsPIL16 or TaPIF3. TaPIF3 instead contains a proline residue at this position. However, it is hypothesised that additional amino acids to the four invariant residues are important for interaction with PHYB. For instance, two proteins in Arabidopsis contain all four invariant residues but show no detectable binding to PHYB. The multiple sequence alignment also illustrates that in the case of the Arabidopsis, rice and maize PIFs, the APB motif is located in very close proximity to the N-terminus of the protein. This observation indicates that *TaPIF3* does not have additional coding sequence upstream of the annotated ATG start codon, therefore supporting the current gene model of *TaPIF3*.

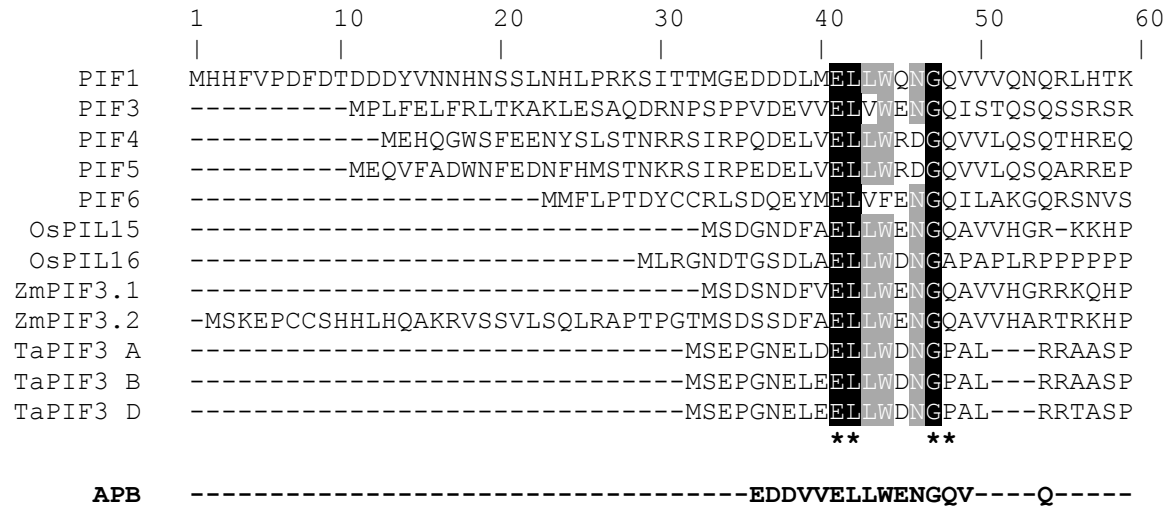


Figure 3.5 TaPIF3 contains a conserved APB motif. A multiple sequence alignment of the APB motif of Arabidopsis, rice, maize and putative wheat PIF proteins. The consensus APB motif is indicated in bold. Residues shown in reverse are highly conserved among the sequences. Residues shown in grey are partially conserved. Asterisks indicate residues shown to be invariant for PHYB interaction by Khanna et al (2004).

3.2.1.7 Searching the IWGSC RefSeq v1.0 assembly

The identification of a *PIF3* orthologue in the wheat genome was completed using the IWGSC CSS and W7984 assemblies that were available at the time. As previously mentioned, the IWGSC made available the IWGSC RefSeq v1.0 wheat genome assembly in 2017. The IWGSC RefSeq v1.0 represents the first chromosome-scale assembly of hexaploid wheat (IWGSC, 2018). To locate the *TaPIF3* gene in the new wheat assembly and to identify potential paralogues of *TaPIF3*, the *TaPIF3* sequence was used in a BLASTN search against the IWGSC RefSeq v1.0 assembly. Results of the BLASTN search identified three genes, TraesCS1A01G083000, TraesCS1B01G100400 and TraesCS1D01G084200, located on chromosomes 1A, 1B and 1D. These three genes made up the top fifteen hits from the BLAST search. Pairwise sequence alignments revealed that the CDS of TraesCS1A01G083000, TraesCS1B01G100400 and TraesCS1D01G084200 share 100% sequence identity to the CDS of *TaPIF3-A1*, *TaPIF3-B1* and *TaPIF3-D1* respectively. Genes

that represent potential duplications of *TaPIF3* were not identified by performing BLAST searches against the IWGSC RefSeq v1.0 assembly.

The release of the first chromosome-scale assembly also allows researchers to determine genes located upstream and downstream of a gene of interest. The IWGSC RefSeq v1.0 assembly on Ensembl plants was used to determine the genes located upstream and downstream of the three homoeologues of *TaPIF3*. Elucidating whether the genes upstream and downstream of the three *TaPIF3* sequences are conserved would provide further evidence that *TaPIF3-A1*, *TaPIF3-B1* and *TaPIF3-D1* represent true homoeologues. Genes TraesCS1A02G082900, TraesCS1B02G100300 and TraesCS1D02G084100 were found to be located directly upstream of *TaPIF3-A1*, *TaPIF3-B1* and *TaPIF3-D1* respectively. Sequence alignments of these three gene sequences strongly suggest that genes TraesCS1B02G100300 and TraesCS1D02G084100 are homoeologous due to a high degree of sequence similarity. In addition, genes downstream of *TaPIF3* also appear to be conserved between the three sub-genomes. Genes TraesCS1A02G083100, TraesCS1B02G100600 and TraesCS1D02G084400 located downstream of *TaPIF3-A1*, *TaPIF3-B1* and *TaPIF3-D1* respectively share pairwise sequence identity scores of 97.6%, 97.9% and 99.1%. These findings show a degree of conservation of gene order between the A, B and D genomes and further suggests the three homoeologues of *TaPIF3* are equivalent.

3.2.2 Cloning of the *TaPIF3* cDNA sequence

3.2.2.1 Amplification of *TaPIF3* cDNA sequence

Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 had initially been identified as promising candidates for a wheat orthologue of *PIF3*. To confirm the sequences of these two genes and to ensure they are expressed in the Cadenza background, a PCR approach was used to amplify these two genes from a cDNA template. PCR was performed prior to the bioinformatics that identified a third homoeologue of *TaPIF3* and that led to the annotation of *TaPIF3* described in section 3.2.1.3. Due to the early uncertainty of the 5' architecture of the two wheat genes, primers were designed to amplify a 3' fragment of Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134.

A common forward primer was designed, TaPIF3_3'_F, that was positioned at the beginning of the last exon of both genes (for primer positions see figure 3.7). Reverse primers TaPIF3_R1 (B) and TaPIF3_R1 (D) were positioned at the end of the last exon. The target amplicon was expected to be approximately 400 bp. PCRs were performed using cDNA synthesised from total RNA isolated from leaf tissue of wheat (cv Cadenza). PCR performed with primers TaPIF3_3'_F and TaPIF3_R1 (D) successfully amplified a product of the expected size (figure 3.6). The resulting PCR product was cloned and sequenced. Seven independent clones were sequenced and the resulting sequencing reads were aligned to the Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 sequences. Of the seven sequencing reads, one read aligned to each Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 supporting the cDNA sequence in each case. The five remaining sequencing reads were subsequently shown to correspond to the A homoeologue of *TaPIF3*.

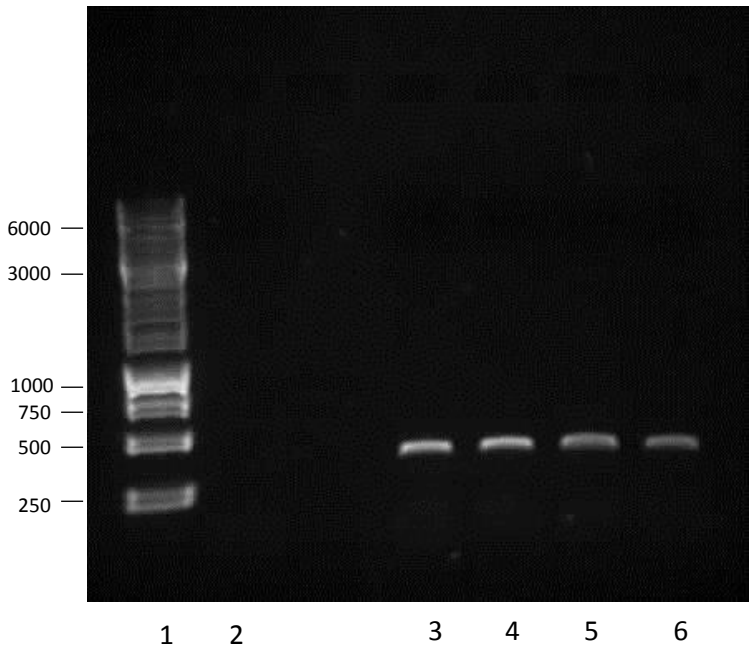


Figure 3.6 Amplification of a 3' fragment of Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134 from cDNA. Agarose gel electrophoresis of PCR performed to amplify 3' Traes_1BS_D1FCBFBE8 and Traes_1DS_55E91A134. PCR amplifications are from cDNA synthesised from leaf total RNA (var. Cadenza). The expected product size is 429 bp. PCR reaction involved four repeats (lanes 3, 4, 5, 6) using the same primer sets. 1kb gene ruler loaded into lane 1 (band sizes indicated in bp). Lane 2 contained a no template negative control.

Subsequent to completing the bioinformatics analysis to provide a full-length sequence for *TaPIF3*, primers were designed to amplify the full-length cDNA sequence for all three homoeologues. The amplification of a small fragment of *TaPIF3* offered promise that the full-length putative gene sequence could be amplified. A common forward primer that anneals to the first 18 bp of *TaPIF3* was designed for the B and D homoeologues. In addition, a specific forward primer was designed for the A homoeologue. Reverse primers, that had been shown to amplify the B and D homoeologue sequences from cDNA were already available, but a specific reverse primer for the A homoeologue was also designed. An initial PCR was performed with primers TaPIF3_F1 (A) and TaPIF3_R1 (A). Attempts to amplify *TaPIF3* focused on the A homoeologue as the A homoeologue was enriched in the sequencing results of the amplified 3' fragment, therefore, indicating that this homoeologue might be expressed to a greater level than the B and D homoeologues. PCRs to amplify *TaPIF3* from a leaf cDNA sample, however, did not yield a band of expected size when run on a gel. Efforts to optimise the PCR conditions included using a temperature gradient for

the annealing stage and adjusting the extension times. Attempts to amplify the full-length sequence with forward primer TaPIF3_F1 (B,D) and reverse primers TaPIF3_R1 (B) and TaPIF3_R1 (D) also did not result in the amplification of a PCR product with the expected size of 1.5 kb. PCR amplifications were repeated using different cDNA samples synthesised from RNA isolated from leaf and immature floral tissue, but with no success. Difficulties in amplifying the full-length coding region of this gene could be attributable to a number of different factors. For example, the full length CDS of *TaPIF3* was calculated to have a 68.4%, 67.8% and 68% GC content for the A, B and D homoeologues, respectively. Sequences with high GC content are widely recognised to be difficult to amplify due to high melting temperatures and the formation of secondary structures.

3.2.2.2 Confirmation of *TaPIF3* CDS by amplification from cDNA

To confirm the gene sequence of T*aPIF3* in the Cadenza background, multiple sets of primers were designed with the purpose of amplifying shorter, overlapping fragments of the three homoeologues (see figure 3.7). Using different combinations of forward and reverse primers to amplify overlapping fragments would also allow confirmation of intron/exon boundaries. Primer sequences are listed in table S.1. Due to the sequence similarity between the three homoeologues, common primers that anneal to more than one homoeologue were often designed.

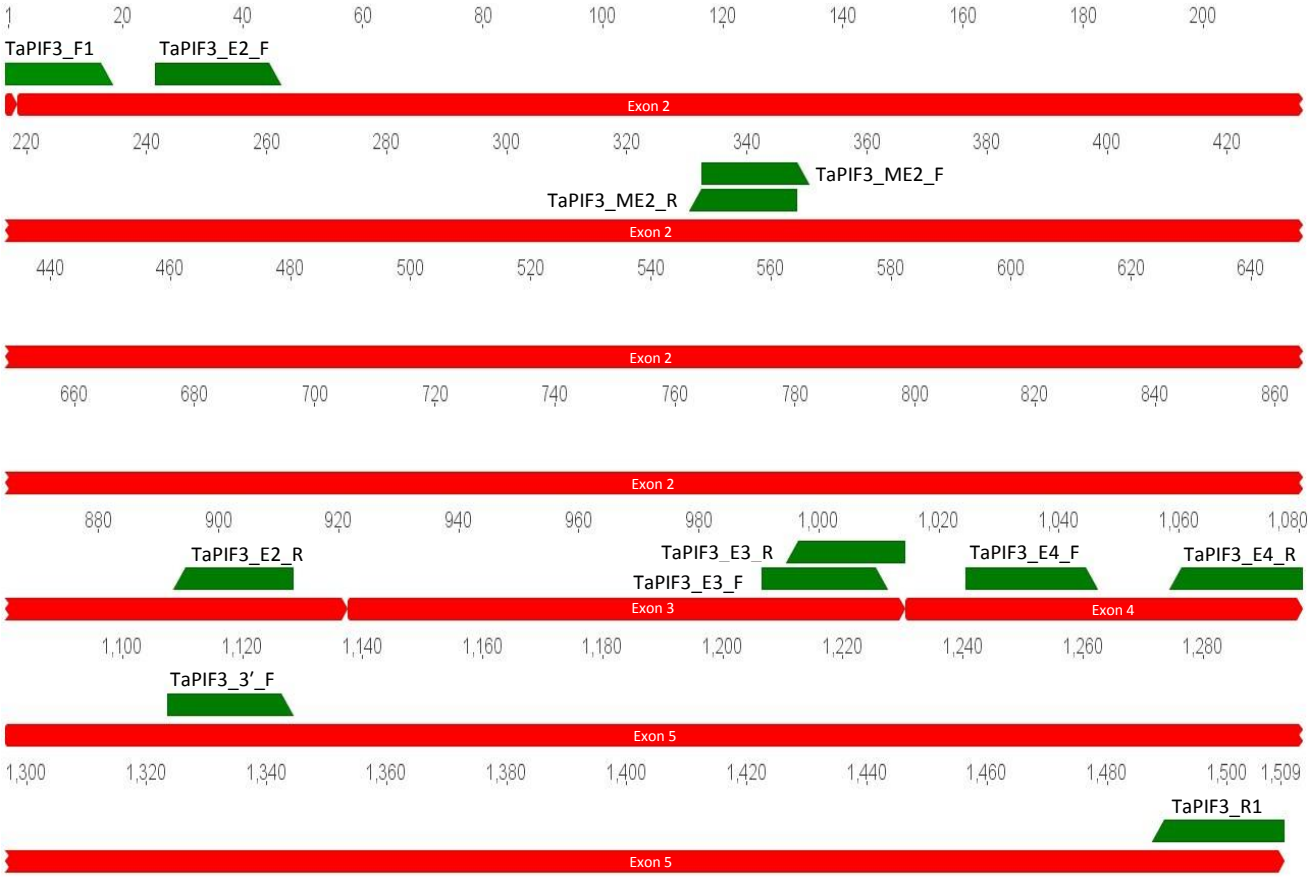


Figure 3.7 Primer annealing positions within the *TaPIF3* sequence. Primer positions are shown relative to the predicted *TaPIF3* CDS. Primers are labelled and annotated in green. Primer orientation is represented by direction of annotation. The *TaPIF3* CDS is annotated in red. Exon numbers are labelled. Sequence length in relation to the predicted initiating ATG is indicated by the nucleotide label above the sequences.

Performing PCR amplifications to test a variety of primer combinations revealed a solitary clear band of expected size when using primers TaPIF3_E2_F and TaPIF3_E2_R with a leaf cDNA template. Primers had been designed to amplify exon 2 of *TaPIF3*, that comprises 920 bp of the total ~1500 bp CDS. This PCR product was therefore an important candidate for sequencing. The PCR was repeated using the proof-reading Phusion DNA polymerase to minimise potential errors introduced during amplification (figure 3.8). The PCR products were cloned and independent clones were sent for sequencing. Sequencing of plasmid DNA confirmed the PCR had amplified exon 2 of *TaPIF3*. Two sequencing reads confirmed the sequence of exon 2 of the B homoeologue. Exon 2 of the D homeologue was also confirmed

by two independent sequencing reads. A single sequence read aligned to the A homoeologue. This sequence read supported the 18 bp insertion that is present in exon 2 of the A homoeologue.

PCRs were performed to confirm additional *TaPIF3* sequences including the intron/exon boundaries. Primers TaPIF3_E3_F and TaPIF3_R1 (B) that were designed to amplify exon 3 through to the end of the last exon produced a band of expected size (figure 3.8). As did primers, TaPIF3_E4_F and TaPIF3_R1 (B) (figure 3.8). PCR was performed using Phusion Taq polymerase and a cDNA template isolated from wheat floral tissue. Sequencing of generated clones confirmed amplification of target fragments. Clones containing the exon 3 - exon 5 fragment predominately matched the sequence of the B homoeologues of *TaPIF3*. In total, four sequencing reads aligned to the A homoeologue, seven reads aligned to the B homoeologue and two reads corresponded to the A homoeologue. In contrast, the A and D homoeologues of *TaPIF3* were the most highly represented in the sequencing of the exon 4 - exon 5 product. Of the fourteen sequenced clones, six reads mapped to the A and D homoeologues specifically. The remaining two sequencing reads aligned to the B homoeologue and further confirmed the sequence at this region. Of importance is that, intron sequences were missing in the sequencing reads of both PCR products. Sequencing therefore confirmed the intron/exon boundaries between exon 3 and exon 4 and also between exon 4 and exon 5.

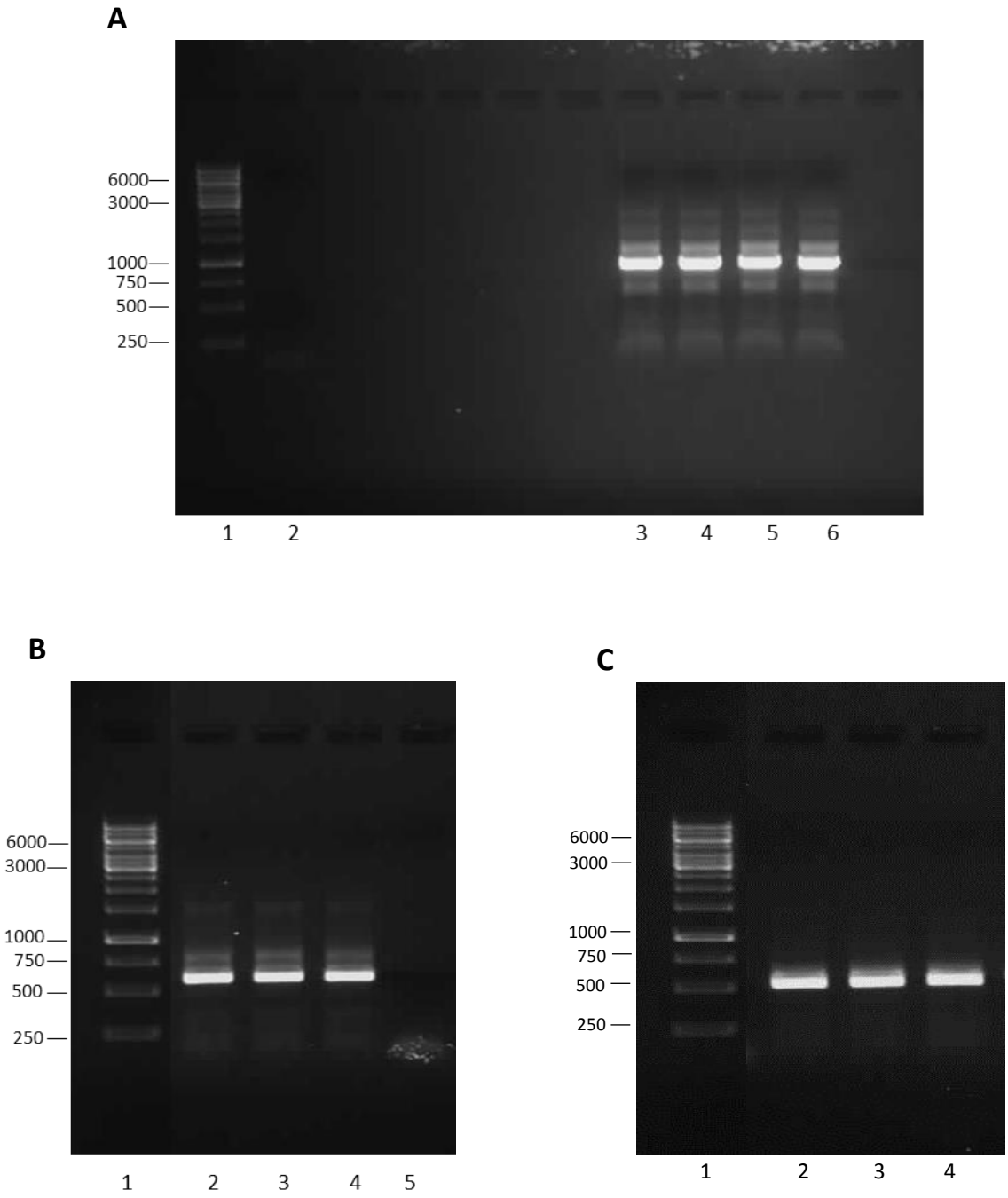


Figure 3.8 Agarose gel electrophoresis of amplified *TaPIF3*. (A) PCR performed to amplify 920 bp of exon 2. Four repeats of PCR using primers TaPIF3_E2_F and TaPIF3_E2_R are loaded into lanes 3-6. 1 kb gene ruler loaded into lane 1, band sizes indicated in bp. A no-template negative control is loaded into lane 2. (B) PCR performed to amplify exon3-exon 5 using primers TaPIF3_E3_F and TaPIF3_R1 loaded into lanes 2-4. Expected amplicon of 580 bp. 1 kb gene ruler loaded into lane 1. A no-template control loaded into lane 5. (C) PCR performed to amplify exon 4-exon5 using primers TaPIF3_E4_F and TaPIF3_R1 loaded into lanes 2-4. Expected amplicon of 495 bp. 1 kb gene ruler loaded into lane 1.

In addition, a fragment of *TaPIF3* was successfully amplified using primers, TaPIF3_E2_F and TaPIF3_E4_R, which were designed to amplify exon 2 through to exon 4. This amplicon therefore represented the majority of the *TaPIF3* cDNA sequence. PCR was performed using Phusion polymerase and a floral cDNA template. Purified PCR product was subsequently cloned and sequenced. Due to the size of the fragment, sequencing was performed from both ends of the insert. This produced two sequencing reads for each clone that overlap in the middle giving coverage of the whole insert. In total, ten sequencing reads were obtained that spanned the target exon 2 – exon 4 product. Sequencing of this PCR product provided confirmation of the intron/exon boundaries between exon 2 and exon 3. In addition, sequencing data further confirmed the intron/exon boundary between exon 3 and exon 4. Of the ten sequencing reads, two reads aligned to the A and D homoeologues of *TaPIF3.* Six reads mapped with 100% sequence identity to the B homoeologue of *TaPIF3.*

3.2.2.3 Amplification of 5′ *TaPIF3*

The sequence and intron/exon architecture of the 5′ region of *TaPIF3* remained unconfirmed. Three additional forward primers, TaPIF3_upstr1, TaPIF3_upstr2 and TaPIF3_upstr3, were designed to anneal to the *TaPIF3* sequence upstream of the translational start codon. RNA-Seq reads map to the scaffold sequence to which these three forward primers should bind (figure 3.9). This region is therefore the putative 5'untranslated region of *TaPIF3* and so a cDNA template could be used for amplification. An additional reverse primer, TaPIF3_5'_R, was designed to be situated in close proximity to the start of exon 2 (figure 3.9).

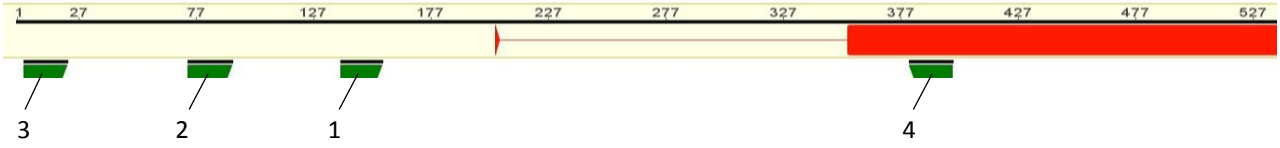


Figure 3.9 Position of primers designed to amplify 5′ region of *TaPIF3*. Position of primers are annotated in green. Primers labelled 1, 2, 3 and 4 correspond to primers TaPIF3_upstr1, TaPIF3_upstr2, TaPIF3_upstr3 and TaPIF3_5'_R respectively. The CDS of *TaPIF3* is shown in red.

To confirm the 5' sequence and architecture of *TaPIF3*, PCR was performed using the described primers. Both a cDNA and a genomic DNA template were used for amplification. It was important to conclude that the genomic sequence at this region in Cadenza is the same or highly similar to the sequence in Chinese Spring and the synthetic W7984. Any differences would have the potential to affect gene annotation of *TaPIF3*. A PCR using a genomic DNA template and primers TaPIF3_upstr3 and TaPIF3_5'_R produced a band of expected size on an agarose gel (figure 3.10). This band was extracted from the gel, cloned and sequenced. Sequencing results illustrated that the PCR had predominantly amplified the D homoeologue of *TaPIF3*. Of the ten sequences obtained, seven sequences aligned to the D homoeologue with 100% sequence identity. At this region, the IWGSC Chinese Spring and W7984 scaffolds for the D homoeologue also share 100% sequence identity. The genomic sequence of *TaPIF3-D1* at this region therefore is conserved between the three wheat varieties.

Primers TaPIF3_upstr3 and TaPIF3_5'_R also yielded a PCR product of interest when a cDNA template was used. Two clear bands could be observed on the gel (figure 3.10). The band representing a product of lower molecular weight matched the expected amplification size of 246 bp. The second band on the gel is presumably due to genomic DNA contamination in the cDNA sample. The PCR product of lower molecular weight amplified from a floral cDNA template was cloned prior to sequencing. The B and D homoeologues of *TaPIF3* were represented in the sequencing data. A total of six sequencing reads mapped to the D homoeologue and three reads mapped to the B homoeologue. Of interest is that the sequencing supported the position of the intron that splits the ATG start codon for the B and D homoeologues. The intron sequence was absent in the product amplified from cDNA as would be expected, but was present in the product amplified from genomic DNA.

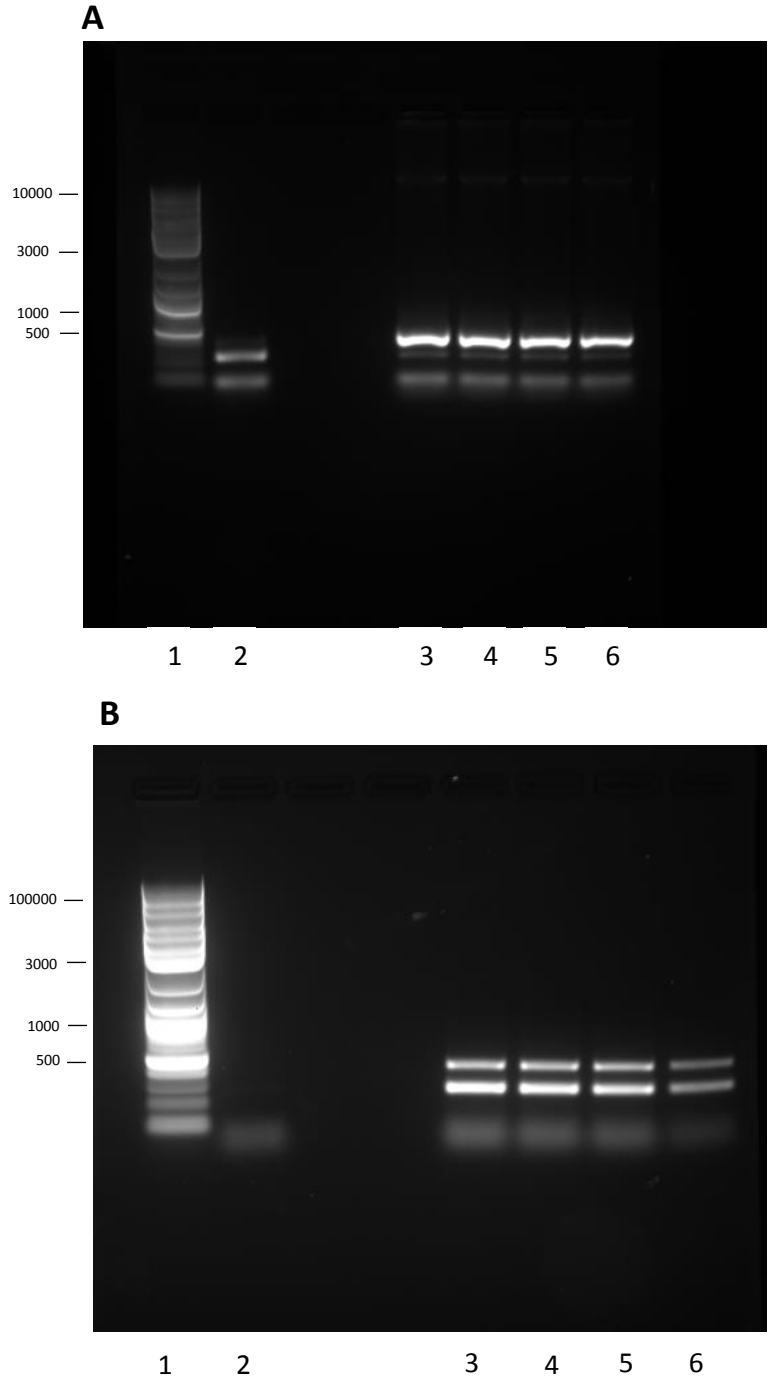


Figure 3.10 Amplification of 5' *TaPIF3*. (A) Agarose gel electrophoresis of PCR performed using primers TaPIF3_upstr3 and TaPIF3_5'_R. A band of expected size, 396bp, was successfully amplified from genomic DNA. Four repeats were performed (lanes 3-6). 2-log DNA ladder was loaded into lane 1 and a no template control loaded into lane 2. (B) Agarose gel electrophoresis of PCR performed using primers TaPIF3_upstr3 and TaPIF3_5'_R. Biomix was used to amplify 5' *TaPIF3* from a cDNA template. The brighter bands of lower molecular weight in lanes 3, 4, 5 and 6 were excised and purified for sequencing. 2-log DNA ladder was run in lane 1. A no template control was run in lane 2.

3.2.2.4 Rapid amplification of cDNA ends

Amplification of the 5' end of *TaPIF3* had confirmed the presence of the intron at the ATG translational start site. To gain evidence that the annotated ATG start codon represents the true translational start site of *TaPIF3*, rapid amplification of cDNA ends (RACE) was used to obtain the full-length sequence of the RNA transcript. To generate a full-length cDNA copy of the *TaPIF3* transcript, the SMARTer RACE cDNA Amplification Kit from Clontech was used. This system relies on the joint action of the SMARTer II A Oligonucleotide and SMARTScribe Reverse Transcriptase. The SMARTScribe Reverse Transcriptase upon reaching the end of the RNA template is able to add 3-5 residues to the end of the first-strand cDNA. The SMARTer II A Oligonucleotide has been designed to anneal to the residues that have been added to the cDNA tail. The Reverse Transcriptase has the ability to switch templates from the RNA molecule to the SMARTer II A oligo, generating a complete cDNA copy of the original transcript. By designing a gene specific primer, combined with a universal primer that anneals to the known sequence at the cDNA tail it is possible to amplify the complete 5' end of the cDNA of interest.

The first step to performing RACE was to generate 5' RACE ready cDNA. 5' RACE cDNA was synthesised using RNA extracted from wheat seedling and wheat floral tissue. In addition, the cDNA synthesis reaction was performed with control mouse heart RNA. When performing the control 5' RACE PCR using the control mouse heart cDNA a band of expected size, 2.1 kb, was present on the gel. The positive control indicated that the cDNA synthesis reactions had been performed correctly. To amplify the 5' sequence of the *TaPIF3* cDNA it was necessary to design an antisense, gene specific primer (GSP). The GSP would be used in combination with the universal primer mix (UPM) (supplied with SMARTer RACE cDNA Amplification Kit) that anneals to the extended cDNA tail to amplify a 5' product. In total seven GSPs were designed (for primer names and sequences see table S.2, for primer positions see figure 3.11). Primers were located close to the 5' end of the gene sequence to reduce the size of the amplified product. Generic primers were designed to anneal to all three homeologues of *TaPIF3* where possible.

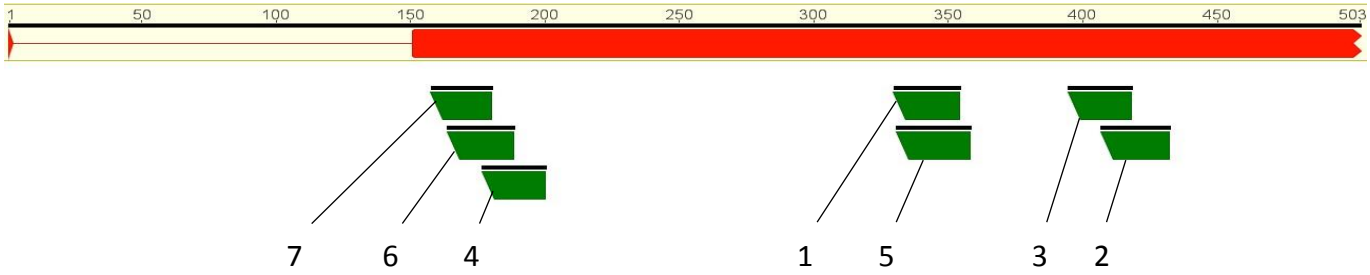


Figure 3.11 Position of *TaPIF3* gene specific 5' RACE primers. Position of antisense gene specific primers designed for 5' RACE. Primers are annotated in green. CDS of *TaPIF3* is annotated in red. A red line indicates non-coding intron sequence. The nucleotide label above the sequence indicates sequence length in relation to the predicted initiating ATG. Primers labelled 1, 2, 3, 4, 5, 6 and 7 correspond to primers TaPIF3_5'Race_1, TaPIF3_5'Race_2, TaPIF3_5'Race_3, TaPIF3_5'Race_4, TaPIF3_5'Race_5, TaPIF3_5'Race_6 and TaPIF3_5'Race_7, respectively.

Initial PCRs performed using 5' RACE ready cDNA derived from wheat seedling tissue and GSPs TaPIF3_5'RACE_2 and TaPIF3_5'RACE_3 did not yield a candidate band for sequencing. However, PCR using UPM and TaPIF3_5'RACE_3 did produce smearing on the agarose gel. The product of this PCR was therefore used as a template for individual nested PCRs using the Nested Universal Primer A (supplied with SMARTer RACE cDNA Amplification Kit) and GSPs TaPIF3_5'RACE1_1, TaPIF3_5'RACE_4 and TaPIF3_5'RACE_5. Gel electrophoresis revealed that no product had been amplified. However, RACE PCR performed using the UPM and TaPIF3_5'RACE_4 produced a faint band of less than 500 bp on an agarose gel. The RACE PCR was repeated using 5' RACE-ready cDNA prepared from wheat seedling and wheat floral RNA. In addition, the number of cycles of the PCR was increased. PCR that used wheat floral 5'RACE cDNA resulted in a clear band of approximately 500 bp on the gel (see figure 3.12). In an attempt to reduce the smearing on the gel, the purified PCR products were used as template for nested PCR. Nested PCR was performed using the Nested Universal Primer A in combination with GSPs TaPIF3_5'RACE_6 and TaPIF3_5'RACE_7. The reactions were analysed on an agarose gel (see figure 3.12). Despite smearing still evident on the gel, visible bands were extracted and purified. The products amplified using GSPs TaPIF3_5'RACE_6 and TaPIF3_5'RACE_7 and the wheat floral RACE reaction as template were sent for sequencing. Both Nested Universal Primer A and the corresponding GSP were used as sequencing primers to generate a full-length sequencing read for both products. The sequence data from both reactions aligned to the B homoeologue of *TaPIF3*. The sequence reads extend

257 bp upstream of the annotated ATG start codon of *TaPIF3*. Within the 257 bp sequence, there are no alternative ATG codons. 5'RACE has therefore provided evidence that the annotated ATG start codon that is split by an intron does represent the true translational start site of *TaPIF3*.

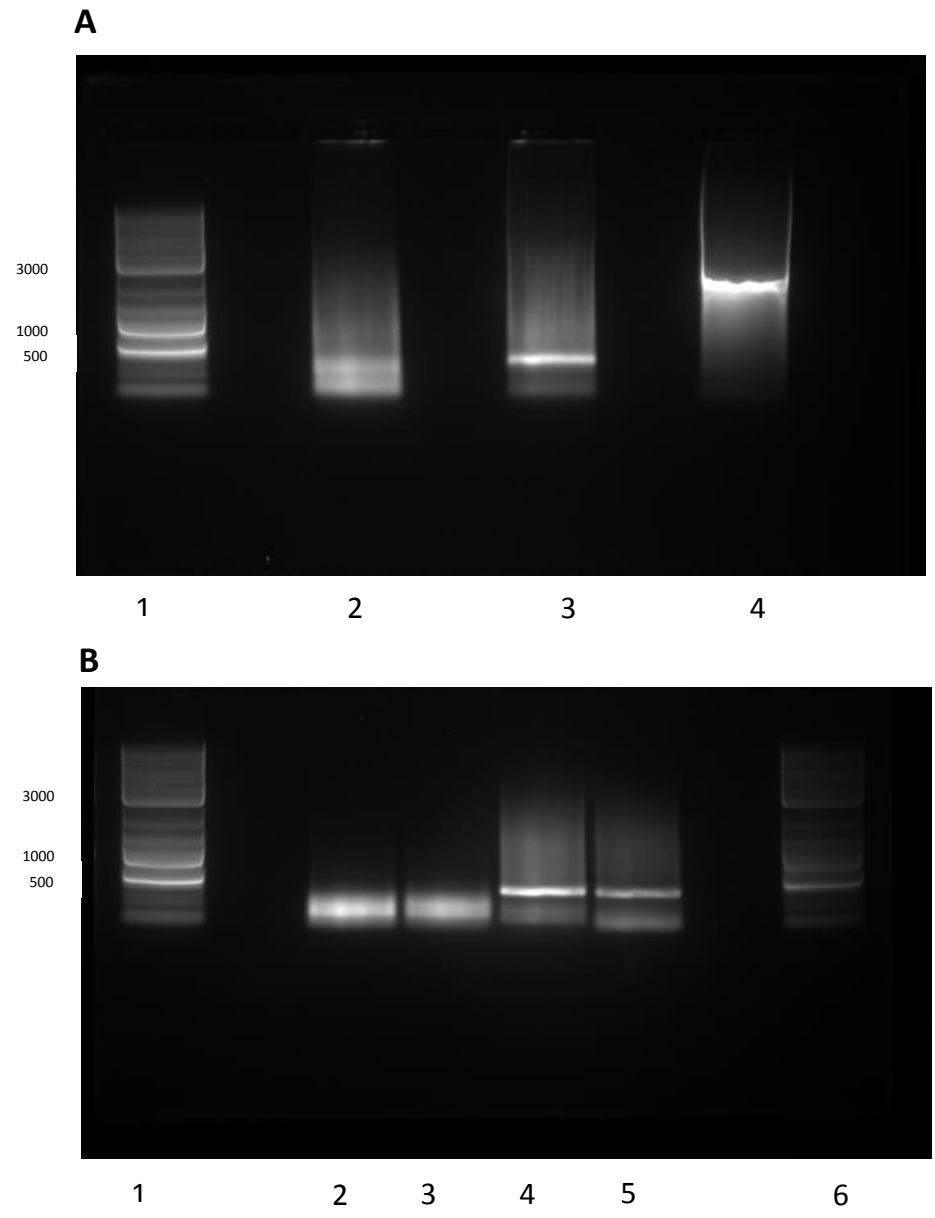


Figure 3.12 Agarose gel electrophoresis of 5' RACE PCR products. (A) Gel electrophoresis of RACE PCR using primers UPM and TaPIF3_5'RACE_4. PCR using wheat seedling RACE ready cDNA was loaded into lane 2. PCR using wheat floral RACE ready cDNA loaded into lane 3. 2-log DNA ladder loaded into lane 1, band sizes indicated in bp. Positive control amplification of 5' mouse transferrin receptor loaded into lane 4. (B) Gel electrophoresis of nested RACE PCR. RACE PCR product amplified from wheat seedling cDNA used as template in lanes 2 and 3. RACE PCR product amplified from wheat floral cDNA used as template in lanes 4 and 5. TaPIF3_5'RACE_6 used as nested GSP in lanes 2 and 4. TaPIF3_5'RACE_7 used as nested GSP in lanes 3 and 5. 2-log DNA ladder loaded into lanes 1 and 6.

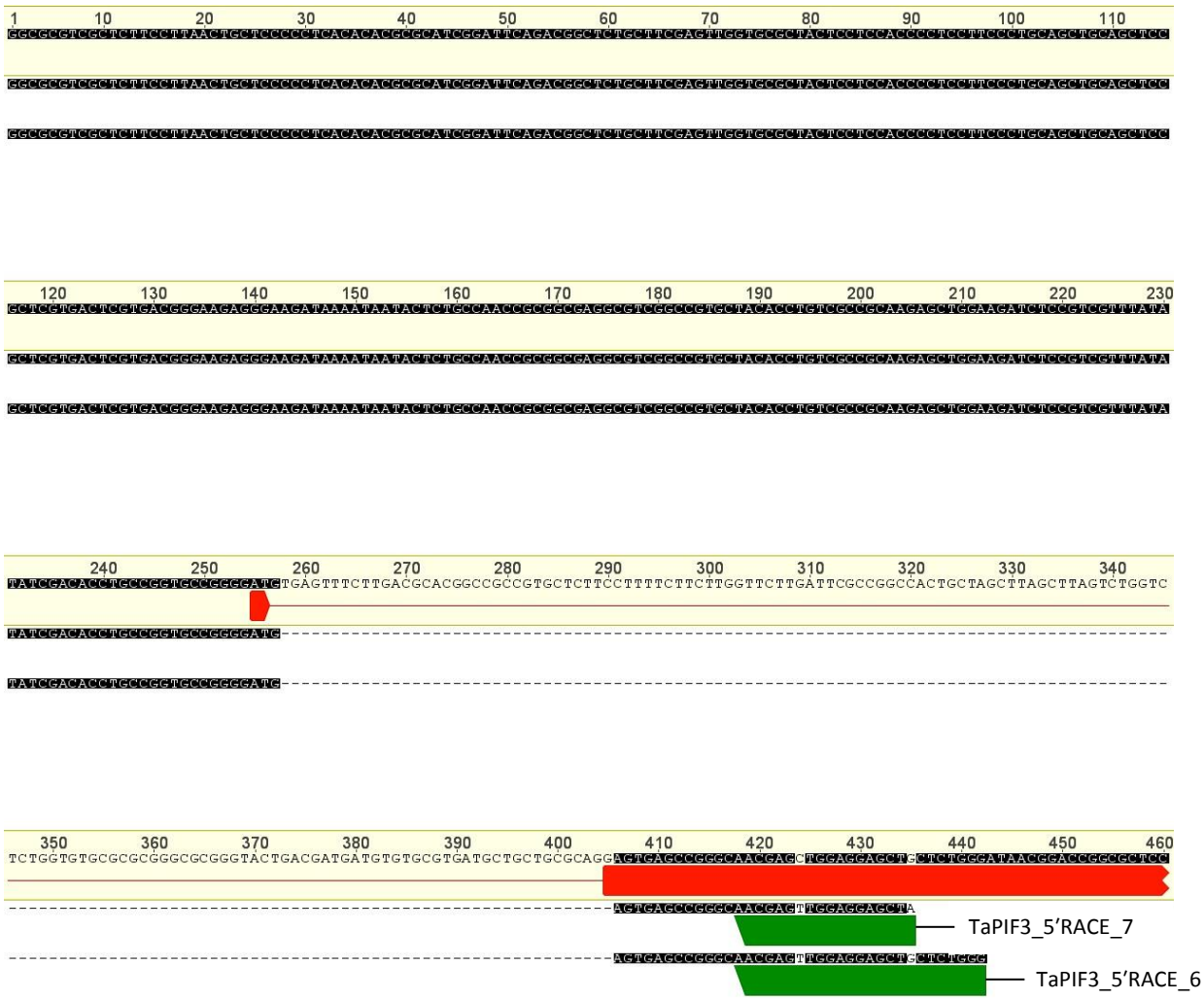


Figure 3.13 Alignment of 5' RACE sequencing. Sequencing reads of nested PCR products amplified using gene specific primers TaPIF3_5'RACE_6 and TaPIF3_5'RACE_7. Sequencing reads are aligned against the *TaPIF3-B1* genomic sequence. CDS of *TaPIF3* is annotated in red. Intron 1 of *TaPIF3* is indicated by a red line. Sequencing reads extend 257 bp upstream of the annotated ATG codon of *TaPIF3*. The nucleotide label above the sequence indicates sequence length in bp in relation to the beginning of the 5'RACE sequencing.

3.2.2.5 Synthesis of the *TaPIF3* coding sequence

The full-length *TaPIF3* sequence was required for overexpression studies in wheat. Amplification of smaller fragments of *TaPIF3* by PCR confirmed the CDS. In addition, 5' RACE has provided evidence to confirm the translational start site for the B homoeologue of *TaPIF3*. PCR amplification of the full-length *TaPIF3* sequence, however, proved unsuccessful. Therefore, a decision was made to synthesise the gene sequence. The B homoeologue of

TaPIF3 was selected for synthesis primarily on the basis of the RACE sequencing data. In addition, sequencing of amplified *TaPIF3* fragments had provided greater coverage of the B homoeologue sequence compared to the A and D homoeologues. The bias of sequencing reads that map to the B homoeologue suggests *TaPIF3-B1* is potentially expressed to a greater level relative to *TaPIF3-A1* and *TaPIF3-D1*. This provided further rationale for selecting the B homoeologue for gene synthesis. The nucleotide sequence encoding for a 5x Myc epitope was added to the 3' end of the *TaPIF3* sequence. In addition, an EcoRI site was added to the 5' end and an EcoRV site added to the 3' end of the *TaPIF3* CDS (figure S.3). Restriction enzyme sites were chosen so that the construct was suitable for downstream cloning applications. The construct was synthesised by GenScript and fused into the cloning vector pUC57 simple. Upon delivery, the *TaPIF3* construct was sequenced to confirm the synthesised sequence.

3.3 Discussion

3.3.1 Identification of a putative *PIF3* orthologue in hexaploid wheat

The first objective of this research project was to identify a *PIF3* orthologue in the wheat genome. In Arabidopsis, PIFs are known to be key regulators in a range of signalling pathways. Perhaps the most prominent role of PIFs is controlling the negative regulation of photomorphogenesis. However, PIFs have also been shown to be important regulators of many other pathways including regulating stem elongation, thermomorphogenesis and responses to biotic and abiotic stress (Pham et al, 2018). Of importance to this project, the function of PIFs is known to be regulated by DELLA proteins as part of the GA signalling pathway (de Lucas et al, 2008, Feng et al, 2008). Identification of *PIF* genes in wheat could therefore provide putative downstream signalling components of the wheat DELLA RHT-1. As the founding member of the PIF family, PIF3 has a prominent role as a negative regulator of seedling de-etiolation. PIF3 is also an important repressor of chloroplast development, including chlorophyll synthesis in etiolated seedlings. Furthermore, PIF3 regulates hypocotyl elongation in response to environmental stimuli. Identification of a *PIF3* orthologue in wheat could therefore provide a specific target to alter growth and development in wheat.

The wheat genome was searched for putative *PIF* orthologues, using the sequences of known Arabidopsis and rice *PIFs*. Three homoeologus genes, termed *TaPIF3,* were identified as being closely related to *PIF3* in a phylogenetic analysis. *TaPIF3* encodes for a protein that contains a PIF-specific APB motif and a bHLH domain, suggesting a conservation of function between *TaPIF3* and Arabidopsis *PIFs*. Sequences for the three homoeologues of *TaPIF3* were obtained from genome sequence assemblies of the Chinese Spring and the synthetic W7984 varieties. To confirm the CDS of *TaPIF3* in the Cadenza background, PCR was used to amplify the cDNA sequence of *TaPIF3* from biological material. However, attempts to amplify the full-length cDNA sequence of *TaPIF3* proved unsuccessful. To this end, shorter, overlapping fragments of the *TaPIF3* sequence were amplified using multiple primer sets. Sequencing of these fragments has provided coverage of the full-length sequence of the B and D homoeologues of *TaPIF3.*

The sequences obtained from amplification of fragments of *TaPIF3* suggests the Cadenza sequence is very similar to the sequences identified from Chinese Spring and W7984. A pairwise alignment of the *TaPIF3-A1* sequence from the Chinese Spring IWGSC RefSeq v1.0 and the W7984 assembly identified a single base pair difference in exon 2. This C-T base pair change is either the result of a varietal difference or an error in one of the assemblies. Sequencing described in this chapter indicates that the Cadenza sequence is identical to the Chinese Spring sequence from the IWGSC RefSeq v1.0 assembly. Sequencing of the amplified PCR products, however, does not provide coverage of the first 27 bp of *TaPIF3-A1*.

An alignment of the *TaPIF3-B1* sequences from the Chinese Spring IWGSC CSS and the W7984 assembly revealed the two sequences share 100% sequence identity. In addition, sequencing of amplified PCR products reveals that the Cadenza sequence is identical to the Chinese Spring and the W7984 sequences. Similar to the A homoeologue of *TaPIF3*, there is a single base pair (C-T) difference in the *TaPIF3-D1* sequence obtained from the Chinese Spring IWGSC CSS and the W7984 assembly. Sequence data generated from overlapping PCR products indicates that the Cadenza sequence is identical to the Chinese Spring sequence with only one base pair difference between the W7984 sequence. Confirming the *TaPIF3* sequence in the Cadenza background has therefore demonstrated high sequence conservation of this gene between the three wheat varieties discussed in this chapter.

Mapping RNA-Seq data to the identified wheat scaffolds indicated that the translational start codon of the *TaPIF3* gene is split by an intron. Searching the literature for genes that contain a start codon split by an intron sequence did not return any further examples. However, in a study, Nielsen and Wernersson (2006) analysed the positions of introns in eukaryotic genes. Previously it had been observed that introns are not uniformly distributed across the gene sequence. Instead, introns tend to be more abundant close to the 5' end (Lin and Zhang, 2005). Introns can be classed into three phases depending on the position within a codon that they occur. A phase 0 intron is positioned between two codons, a phase 1 intron disrupts a codon after the first nucleotide and a phase 2 intron disrupts a codon after the second nucleotide. The intron that disrupts the ATG start codon of *TaPIF3* is therefore a phase 2 intron. In most gene sequences, phase 0 introns are the most common, followed by phase 1 and then phase 2 introns. The bias of phase 0 introns is particularly pronounced in plant genes (Nielsen and Wernersson, 2006). Of interest is that, Nielsen and

Wernersson (2006) observed an overabundance of phase 0 introns at position 2. These introns are therefore positioned between the start codon and codon 2. This phenomenon was especially apparent in vertebrates, arthropods and fungi. However, phase 1 and phase 2 introns at position 1 that disrupt the start codon were seen to be relatively rare, especially in plants (Nielsen and Wernersson, 2006).

Due to concerns regarding the annotation of 5' *TaPIF3*, PCR was performed to amplify this region of the gene sequence. PCR successfully amplified a 5' fragment of *TaPIF3* from a cDNA template. Sequencing of the described PCR product confirmed the intron that splits the ATG codon is spliced. To further substantiate the annotated ATG start codon, 5' RACE was performed. Sequencing results from a nested RACE PCR product provided important transcript information for the *TaPIF3* gene. The sequence data maps to the B homoeologue and indicates that the transcript extends 257 bp upstream of the annotated ATG codon (figure 3.13). Additional ATG codons could not be found within this 257 bp region supported by the RACE sequencing reads. The RACE sequencing also indicates that there are no further exon(s) upstream that could not be identified by RNA-Seq mapping due to relatively short genomic scaffolds. Of course, non-canonical start codons such as CTG could represent potential translational start sites. However, the information that is available to us supports the described annotation of *TaPIF3.*

Due to the inability to clone the full-length *TaPIF3* sequence, the *TaPIF3* coding sequence was synthesised. A nucleotide sequence coding for a 5x Myc tag was added to the 3' end of the gene sequence. The Myc tag has become a popular epitope tag for detecting expression of recombinant proteins. Multiple Myc sequences were added to the 3' end of the *TaPIF3* sequence to increase the likelihood of protein detection. In a previous study, a similar C-terminal 5x Myc tag was used to successfully detect PIF3 protein levels in 35S:6xHis:PIF3:5xMyc transgenic Arabidopsis seedlings (Al-Sady et al, 2006). The abundance of this recombinant PIF3 protein level was reduced after red light pulses therefore indicating that sensitivity to light had not been abolished by the addition of the tag. The synthesis of the *TaPIF3* gene sequence will allow us to generate the *TaPIF3* overexpression line in wheat.

3.3.2 *TaPIF3* is orthologous to *OsPIL15*, *OsPIL16* and *PIF3*

Phylogenetic analysis was performed to identify wheat sequences that are orthologous to Arabidopsis PIFs and rice OsPILs. Arabidopsis PIF proteins especially have well characterised functions, therefore identifying wheat sequences that group closely to specific PIFs could give an indication of function. The three homoeologues of *TaPIF3* grouped closely with Arabidopsis PIF3, OsPIL15 and OsPIL16. This finding is in agreement with phylogenetic analysis conducted by Nakamura et al (2007) who showed that the OsPIL15 and OsPIL16 sequences are highly similar to PIF3. Despite PIF3 grouping with OsPIL15, OsPIL16 and TaPIF3, Arabidopsis, rice and wheat sequences often only grouped with other sequences from the same species. Therefore, additional PIF orthologues in wheat could not be identified using this approach.

In Arabidopsis, a family of PIF proteins are present that play distinct but overlapping roles in light responses. This is most evident as only the quadruple *pifq* mutant displays a completely constitutively photomorphogenic phenotype in the dark (Shin et al, 2009). PIFs have now been studied in species besides Arabidopsis. Tomato encodes a family of eight PIF proteins: *SlPIF1a, SlPIF1b, SlPIF3, SlPIF4, SlPIF7a, SlPIF7b, SlPIF8a, SlPIF8b* (Rosado et al, 2016). The function of SlPIF1a has been characterised and been shown to function in a similar manner to Arabidopsis PIFs. SlPIF1a is degraded in response to red light and is able to restore hypocotyl elongation in dark-grown *pifq* Arabidopsis mutant seedlings to that of the triple mutant lacking PIF3, PIF4 and PIF5 (Llorente et al, 2016). The function of the remaining tomato PIFs are currently unknown. However, the expression profiles of the eight *PIF* genes in tomato differ depending on the developmental stage of the plant, for instance in de-etiolation, senescence and fruit ripening. This suggests that these genes have undergone functional specification (Rosado et al, 2016). *PIF* genes have also been identified in the moss species *Physcomitrella patens* (Possart et al, 2017). *P. patens* contains four putative *PIF* orthologues: PpPIF1, PpPIF2, PpPIF3 and PpPIF4. In contrast, the liverwort *Marchantia polymorpha* has only one *PIF* gene, *Mp-PIF,* which like Arabidopsis PIFs is able to regulate gene expression and is targeted for degradation by phytochrome in response to red light (Inoue et al, 2016). This study indicates that the role of PIFs was established early in land plant evolution. The functional role of TaPIF3 is therefore hard to predict. TaPIF3 could play

a redundant or overlapping role with additional PIFs yet to be characterised in wheat. However, TaPIF3 could also have a distinct role that may be similar to that of PIF3.

Chapter 4. Manipulating the expression of *TaPIF3* in wheat

4.1 Introduction

One objective of this research project was to characterise the function of a *PIF3* orthologue in hexaploid wheat. To determine gene function, the aim was to analyse the effect of knockdown/knockout and overexpression of *TaPIF3* in wheat. In contrast to Arabidopsis, there are fewer resources available to generate loss-of-function mutants in wheat. The polyploid nature of the bread wheat genome introduces complexities when attempting to silence or promote expression of specific genes. For example, most genes in wheat are present as three homoeologous copies. A consequence of this is that the effect of loss-of-function mutations in any single wheat homoeologue is likely to be masked by redundancy of the remaining homoeologues. Therefore, to interrogate gene function, it is often necessary to knock out or silence all homoeologous copies which can be difficult and time consuming. RNAi provides a useful approach to concomitantly knockdown multiple homoeologues. However, this technology has its own limitations such as producing hypomorphic phenotypes and causing off-target gene silencing (Boettcher and McManus, 2015, Birmingham et al, 2006).

4.1.1 Methods of altering gene expression in wheat

4.1.1.1 Targeting Induced Local Lesions IN Genomes (TILLING)

Targeting Induced Local Lesions IN Genomes (TILLING) provides a non-transgenic alternative to generate knockout mutants in wheat. It is a reverse genetics approach that allows sensitive detection of induced point mutations in individuals of a mutagenized population. The method was first described in *Arabidopsis thaliana* (McCallum et al, 2000) and Drosophila (Bentley et al, 2000) and has now been successfully extended to crop species such as rice. Mutation breeding of important species such as wheat however is known to be limited by genetic redundancy as a result of polyploidy (Parry et al, 2009). In a polyploid species, recessive alleles are less likely to demonstrate a clear phenotype. This is due to

complementation of homoeologous copies of the targeted gene. The feasibility of using TILLING as an approach in hexaploid wheat however was demonstrated more than 10 years ago (Slade et al, 2005). In this study, TILLING was used to identify loss-of-function alleles in the Waxy (granule-bound starch synthase I) genes. TILLING identified null alleles in two of the homoeologous genes which were then stacked with a natural null allele in the third homoeologue to generate a triple homozygous mutant that demonstrated clear phenotypic abnormalities in grain development (Slade et al, 2005). This finding demonstrated the potential of TILLING for investigating gene function and its use in crop improvement.

As described, TILLING combines mutagenesis with a sensitive DNA screening technique. A mutagenized population is typically established through use of the chemical mutagen ethyl methanesulfonate (EMS). EMS alkylates guanine residues, producing O^6-ethylguanine (Greene et al, 2003). The resulting O^6-ethylguanine base pairs with T but not C. As a consequence, replication of alkylation damaged DNA will result in a G/C base pair being replaced with an A/T. It is therefore estimated that over 99% of EMS mutations are G-A or C-T substitutions. To detect mutations that have been induced in candidate genes, a number of different platforms have been used in TILLING workflows. Labour intensive/low throughput methods involve amplification of the target gene region. Mismatches are subsequently detected in a heteroduplex amplicon using either the nuclease CJE (celery juice extract) followed by gel electrophoresis or by a high resolution melt analysis (Botticella et al, 2011, Uauy et al, 2015).

To increase the throughput of TILLING, Tsai et al (2011) developed an approach of multidimensional pooling of DNA amplicons. Next generation, Illumina sequencing was subsequently used to detect mutations in up to 40 genes from 768 individuals simultaneously (Tsai et al, 2011). However, the small number of genes that can be targeted in each run limits this approach. In addition, this technology remains labour intensive due to the normalisation of DNA at different stages. An alternative, more efficient strategy involves genomic enrichment technologies that would allow genome wide detection of mutations in the coding part of the genome (exome). Exon capture methods have already been used in the crop species rice (Henry et al, 2014). In addition, exon capture technology has now been used to successfully detect induced mutations in a population of hexaploid wheat (King et al, 2015). King et al (2015) performed a pilot scale evaluation of TILLING in hexaploid wheat.

In this study, an 1831 coding sequence subset from the RIKEN full-length cDNA database was selected and supplemented with 15 genes of interest. A capture array was then designed that covered approximately 2 Mbp of wheat gene coding sequence. Genomic enrichment was performed using genomic DNA from three M5 wheat lines known to contain point mutations in a gene of interest. Following exon capture, the captures were pooled and sequenced using next generation illumina sequencing. The results of the analysis identified 464 putative SNPs of which 453 represented expected EMS-induced changes. In addition, the known mutations in their genes of interest were included in the identified SNPs. This study therefore demonstrated the feasibility of identifying mutations in hexaploid wheat by exon capture followed by next generation sequencing.

An invaluable resource recently made available to the wheat research community is the wheat TILLING populations in tetraploid Kronos durum wheat and Cadenza hexaploid wheat. These TILLING populations were developed as part of a collaboration between Rothamsted Research, University of California Davis, the Earlham Institute and the John Innes Centre. The populations were generated by mutagenesis with EMS. Exome capture and next generation Illumina sequencing was subsequently performed on 1535 Kronos and 1200 Cadenza M3 lines. Sequencing reads were mapped to the IWGSC CSS assembly to identify locations of induced mutations. In total, more than 10 million mutations have been catalogued in the Kronos and Cadenza genomes. This corresponds to an average of 23-24 missense and truncation alleles per gene, with more than 90% of the captured wheat genes in each population containing a mutation with a predicted deleterious effect (Krasileva et al, 2017). In this study, the aim is to use TILLING-by-sequencing to generate a *TaPIF3* knockout mutant in hexaploid wheat.

4.1.1.2 Overexpression

The function of *TaPIF3* in regulating growth and development in wheat will also be studied through the production of overexpression lines. The use of loss-of-function mutants alone has often been insufficient to deduce the function of a gene of interest (Prelich, 2012). Studying the effect of overexpressing a specific gene provides an alternative strategy that is not as vulnerable to functional redundancy. Functional redundancy is a prevalent issue

especially when studying the function of transcription factors. Transcription factors are usually members of large gene families that often have overlapping functions. In some cases, it is therefore necessary to knockout multiple transcription factors in order to observe phenotypic abnormalities (Kumaran et al, 2002). In contrast, the phenotype of a mutant that overexpresses a member of a gene family can often be attributed to a single gene. To generate overexpression lines, cDNAs or ORFs are cloned downstream of a strong promoter resulting in a high level of expression. In Arabidopsis, the *in planta Agrobacterium*-mediated transformation method has made it easy to generate a large collection of overexpression mutants in this species (Kondou et al, 2010).

The use of overexpression lines to characterise novel genes does have some limitations. A constitutive promoter is usually used to drive expression of transgenes in overexpression studies. This can lead to high expression levels of the transgene in non-target tissues or inappropriate developmental contexts. The non-regulated expression of the transgene can therefore cause ectopic expression of endogenous genes leading to a phenotype that is not associated with the function of the transgene (Kondou et al, 2010). For instance, overexpression of phyA in Arabidopsis indicated that phyA had a similar role to phyB in light responses (Boylan and Quail, 1991). However, the function of phyA is dependent on its kinetics, meaning that overexpression interfered with the normal function of the photoreceptor (Reed et al, 1994). Another issue that should be considered is the endogenous expression level of the gene of interest. If a protein is already expressed at a saturating level, overexpression is unlikely to cause a phenotype. Due to the limitations discussed, overexpression lines will be used in parallel with loss-of-function TILLING mutants to investigate *TaPIF3* function.

The aim of this chapter is to manipulate levels of *TaPIF3* expression in wheat in order to characterise *TaPIF3* gene function.

4.2 Results

4.2.1 A TILLING-based approach to generate knockout mutants of *TaPIF3* in hexaploid wheat.

4.2.2 Identification of mutations within the *TaPIF3* sequence

This research project benefitted from an early analysis of the Cadenza TILLING population (described in section 4.1.1.1) by Dr Andy Phillips at Rothamsted Research. Dr Andy Phillips mapped sequence reads generated from exon capture of the mutant population to a subset of gene targets. This subset of gene targets included the three homoeologues of *TaPIF3* identified in chapter 3. Mutations were identified and annotated with the library number, the SNP position in the genomic sequence, the number of supporting variant reads and whether the mutation was heterozygous (het) or homozygous (hom) in the M2 generation. In addition, the effect of each mutation was predicted at Rothamsted Research using the program SnpEff (Cingolani et al, 2012).

The described analysis of the TILLING population provided sequence files of each gene target, annotated with the position of each mutation along with the predicted effect of each mutation. This allowed me to scan the sequence files and select mutations that were likely to cause an effect on *TaPIF3* expression. Mutations that introduced a premature stop codon or disrupted a splice site were most likely to have a knockout effect. A C/T substitution was identified in exon 2 of the A homoeologue of *TaPIF3* which results in a nonsense mutation (Q85*; library 11400). To confirm the mutation, I mapped raw sequencing reads from library 11400 to the A homoeologue sequence. Mapping of sequencing reads confirmed the mutation at position 1037 of the genomic sequence (figure 4.1). The mutation was supported by 10 variant reads compared to two WT reads. The position of the introduced stop codon is located 5’ of the bHLH coding sequence and therefore would be expected to abolish *TaPIF3-A1* function. Further nonsense mutations in *TaPIF3-A1* could not be identified by searching the sequence files.

Analysing mutations that had been introduced in the B homoeologue failed to identify a nonsense mutation. However, searching the sequencing files identified a G/A substitution predicted to disrupt the 3’ splice site of intron 2 (figure 4.1). The splice acceptor mutation was detected in library 10421 at position 1682 of the *TaPIF3-B1* genomic sequence. Mapping sequencing reads generated from this library to the *TaPIF3-B1* sequence reveals

the mutation is supported by 29 variant reads and 0 WT reads. The absence of WT reads indicates that this mutation is homozygous in the M2 generation. The putative splice acceptor mutation is predicted to prevent splicing of intron 2 from the mRNA message causing an altered protein sequence. The exons spanning intron 2 encode for the bHLH domain of the protein. Any disruption at this region would therefore be very likely to disrupt function. The described splice acceptor mutation represented the only mutation in *TaPIF3-B1* with a predicted effect on correct splicing.

In addition to the mutations identified in the A and B homoeologues, a mutation that is predicted to cause loss-of-function was identified in the D homoeologue. A C/T substitution detected in library 9953 results in a Q260* mutation. The point mutation is located within exon 2 of *TaPIF3 D* and lies 5' of the bHLH coding region. The mutation is supported by 11 variant reads and 9 WT reads as identified by mapping raw sequencing reads from library 9953 to the genomic sequence. Further nonsense mutations in the D homoeologue were not detected by TILLING and so the described mutation was selected for this study. Due to the polyploid nature of the bread wheat genome, it would be unlikely to detect a phenotype in plants containing only one of the described alleles. To prevent complementation by homoeologous copies of *TaPIF3*, the aim was to combine all three mutant alleles of *TaPIF3* by crossing to study gene function.

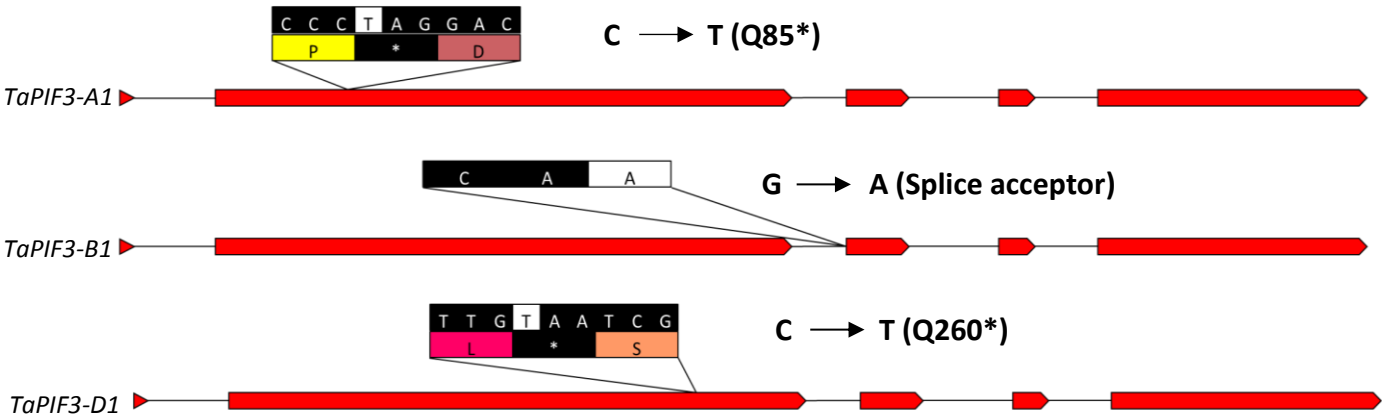


Figure 4.1 Identification of TILLING lines for each homoeologue of *TaPIF3*. Each line carries a mutation that is predicted to cause loss-of-function. The position of the mutation is shown against a schematic of the *TaPIF3* genomic sequences. The single-nucleotide polymorphism identified in each homoeologue is indicated in a white box. The exon sequences are indicated by red arrows.

4.2.3 Genotyping M4 progeny

The library number corresponding to the exon capture of the M2 generation could be converted to a CAD4 identification code (line number). LIB11400, LIB10421 and LIB9953 corresponds to CAD4-0649, CAD4-1803 and CAD4-0256 respectively. M4 seed from these lines was made available to this project prior to the release to the wider scientific community. Six M4 seeds from each of the identified lines were planted. Leaf tissue was sampled from each of the 18 plants for genomic DNA extraction and subsequent genotyping by sequencing. To genotype for the desired mutation, it was necessary to amplify and sequence the homoeologue-specific region flanking the mutation. Homoeologue-specific primers were designed to allow amplification of a 200-500 bp region containing the mutation of interest in each of the three *TaPIF3* genes. Primers typically contain at least two mismatches at the 3' end to the two non-target homoeologues.

To amplify the target region, PCR was performed using different forward and reverse primer combinations. Initially, two forward and reverse primers were designed for each homoeologue. However, due to problems amplifying the target sequence from lines CAD4-0649 and CAD4-0256, additional primers were designed (table S.3). Primer combinations were tested using an annealing temperature gradient PCR. Genomic DNA samples from individual plants of each TILLING line was used as template in a BioMix PCR reaction. Primers 1AS_forward_till_4 and 1AS_reverse_till_1 successfully amplified the 307bp region of *TaPIF3-A1* containing the mutation of interest. PCR products were purified and sent for sequencing. Sequencing results indicated that plants 1, 2 and 5 of line CAD4-0649 were heterozygous for the A homoeologue mutation as shown by two distinct peaks on the sequencing chromatogram. Plants 3 and 6 were wild type for the mutation whilst plant 4 was homozygous for the mutation (figure 4.2).

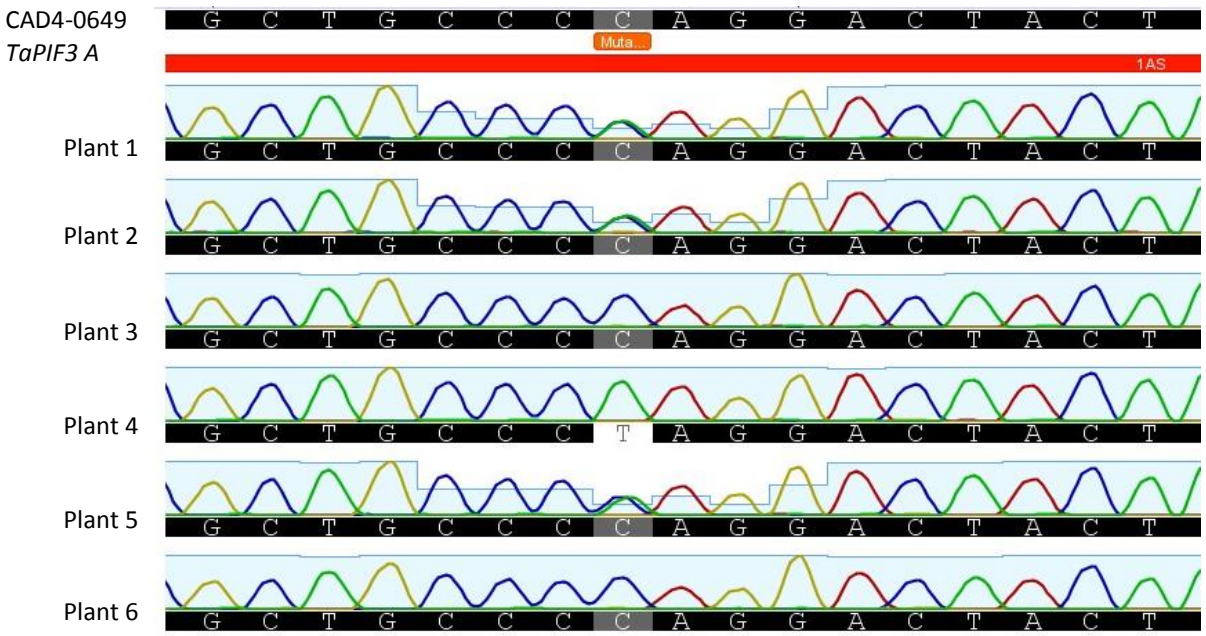


Figure 4.2 Genotyping by sequencing of line CAD4-0649. Genomic DNA was extracted from six plants of line CAD4-0649. A homoeologue specific PCR was performed to amplify the region of *TaPIF3 A* containing the identified TILLING mutation. PCR products were sequenced and aligned to the *TaPIF3 A* sequence. The coding sequence of *TaPIF3* is indicated in red, the site of the mutation is annotated in orange. The chromatogram of each sample is shown.

PCR performed using primers 1BS_forward_till_1 and 1BS_reverse_till_1 successfully amplified a product of expected size from CAD4-1803 genomic samples. The PCR products corresponding to six individual plants were purified and sequenced. Results of the sequencing indicated plants 2-6 of line CAD4-1803 were homozygous for the mutation. As previously mentioned the absence of wild type reads from the sequencing of M2 DNA had implied that this mutation was homozygous. The presence of a second peak in the chromatogram of plant 1 could therefore be an artefact of the PCR or sequencing (figure 4.3)

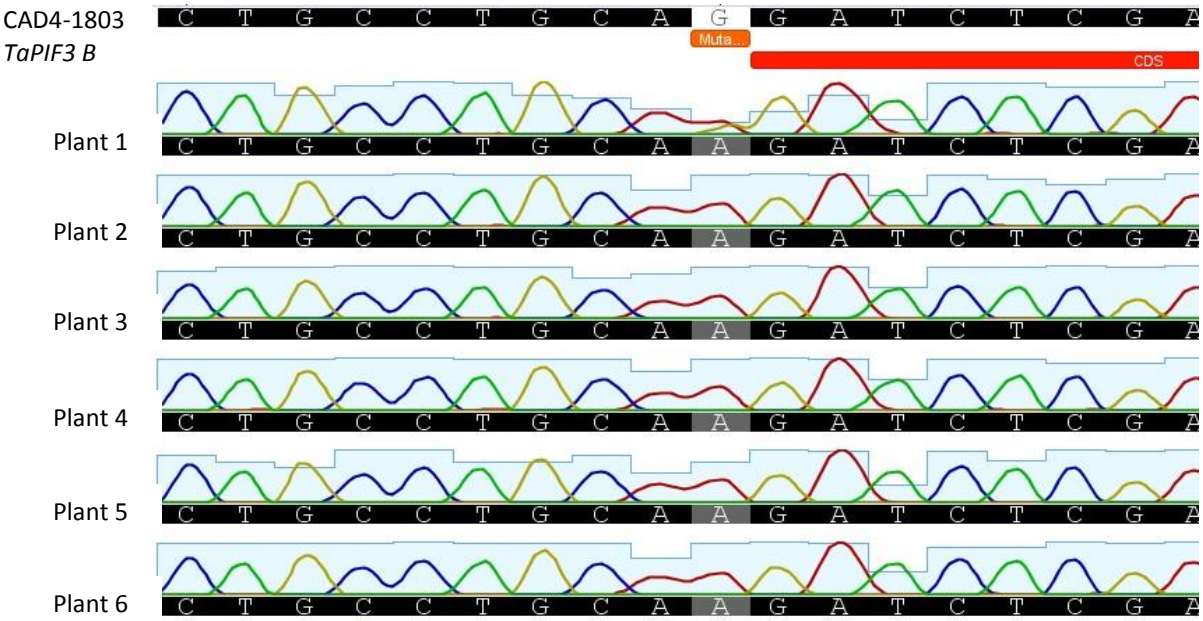


Figure 4.3 Genotyping by sequencing of line CAD4-1803. Genomic DNA was extracted from six plants of line CAD4-1803. A homoeologue specific PCR was performed to amplify the region of *TaPIF3 B* containing the identified TILLING mutation. PCR products were sequenced and aligned to the *TaPIF3 B* sequence. The coding sequence of *TaPIF3* is indicated in red, the site of the mutation is annotated in orange. The chromatogram of each sample is shown. Sequencing data indicates that plants 2-6 are homozygous for the mutation.

The sequence flanking the identified mutation in the D homoeologue of *TaPIF3* was successfully amplified from genomic DNA samples of line CAD4-0256 using primers 1DS_forward_till_4 and 1DS_reverse_till_4. Sequencing of the amplified products indicated that plants 1 and 3 of line CAD4-0256 were heterozygous for the D homoeologue mutation. Plant 2 was shown to be wild type for the mutation whilst sequencing indicated that plants 4, 5 and 6 were homozygous for the C-T mutation (figure 4.4).

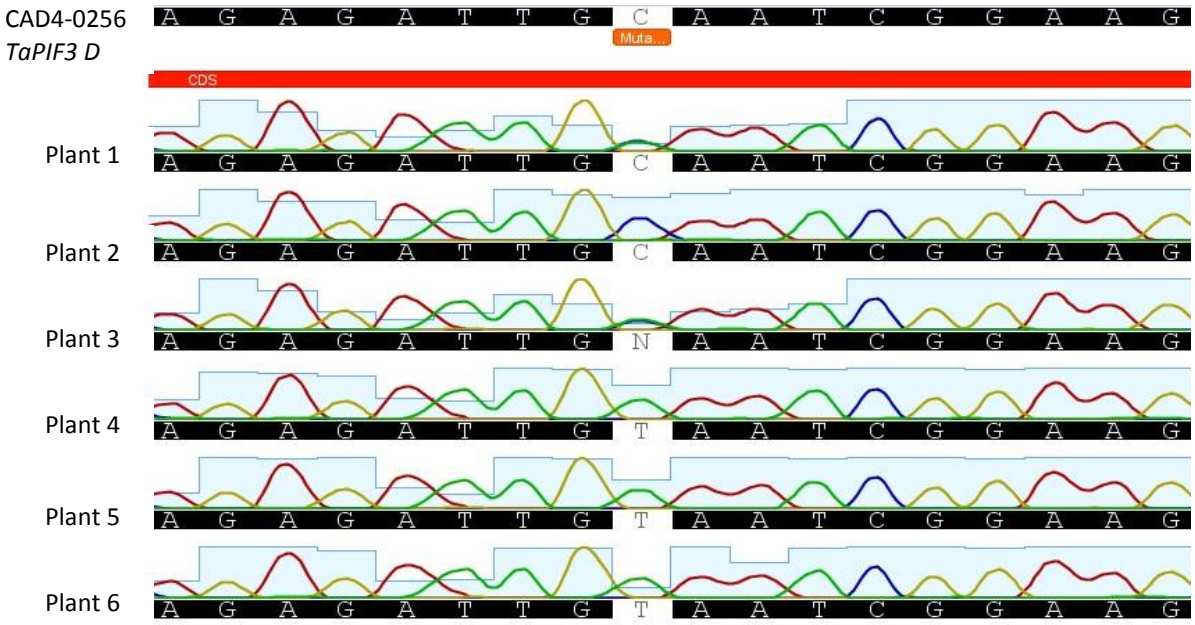


Figure 4.4 Genotyping by sequencing of line CAD4-0256. Genomic DNA was extracted from six plants of line CAD4-0256. A homoeologue specific PCR was performed to amplify the region of *TaPIF3 D* containing the identified TILLING mutation. PCR products were sequenced and aligned to the *TaPIF3 D* sequence. The coding sequence of *TaPIF3* is indicated in red, the site of the mutation is annotated in orange. The chromatogram of each sample is shown. Sequencing data indicates that plants 4, 5 and 6 are homozygous for the mutation.

4.2.4 Stacking of *TaPIF3* alleles

Six plants from each of the identified lines, CAD4-0649, CAD4-1803 and CAD4-0256 were grown to the onset of anthesis. Due to problems with PCR amplification, genotyping of plants from A homoeologue line CAD4-0649 could not be completed in time for crossing. To combine the identified B and D alleles of *TaPIF3,* plants homozygous for each mutation were crossed to produce F_1 progeny containing both alleles. The resulting F_1 progeny was planted and leaf tissue harvested for genomic DNA extraction. Four F_1 plants were selected for genotyping. PCR was performed using the same combinations of homoeologue specific primers used to genotype the M4 generation. Sequencing of the amplified PCR products confirmed all four F_1 plants were heterozygous for the B and D mutations. To combine all three mutant alleles, BxD F_1 plants were set to be crossed with A homoeologue mutants.

However, it became apparent that plants of line CAD4-0649 exhibited a late flowering phenotype, potentially due to mutations elsewhere in the genome. Plants carrying the B and D mutant alleles reached head emergence and commencement of anthesis approximately 3 weeks prior to CAD4-0649 genotyped plants. Due to the delay in flowering, performed crosses proved unsuccessful.

Plants heterozygous for the B and D mutations were therefore allowed to self-pollinate to identify lines homozygous for the two alleles. The resulting progeny was planted and twenty individuals were genotyped. Genotyping identified two individuals that were homozygous for both mutations. These plants were crossed with plants homozygous for the A mutation to stack all three alleles. The resulting heterozygous AaBbDd mutant was identified by sequencing and allowed to self-pollinate to generate F_2 grain. According to Mendelian genetics, one in sixty-four plants could be expected to be the triple mutant. To avoid screening a large number of individual lines in order to identify the triple mutant, a smaller number of plants were grown and genotyped with the aim to identify a line containing five mutant alleles. A total of 30 individuals were genotyped. Genotyping successfully identified a single line, plant 24, containing five mutant alleles aaBbdd. This plant was left to self-pollinate in order to detect homozygous triple mutants in the next generation. Of importance, a WT segregant line was also identified. In addition, all three double combinations of mutant alleles could be identified.

The progeny of plant 24 aaBbdd was grown at both Rothamsted Research and the University of Southampton. Plants were genotyped to identify individuals that were homozygous for the B mutation. Nine individuals were confirmed to be homozygous for the B mutation. Lines identified to be homozygous for the B mutant allele were subsequently genotyped to confirm the A and D mutant alleles (figure 4.5).

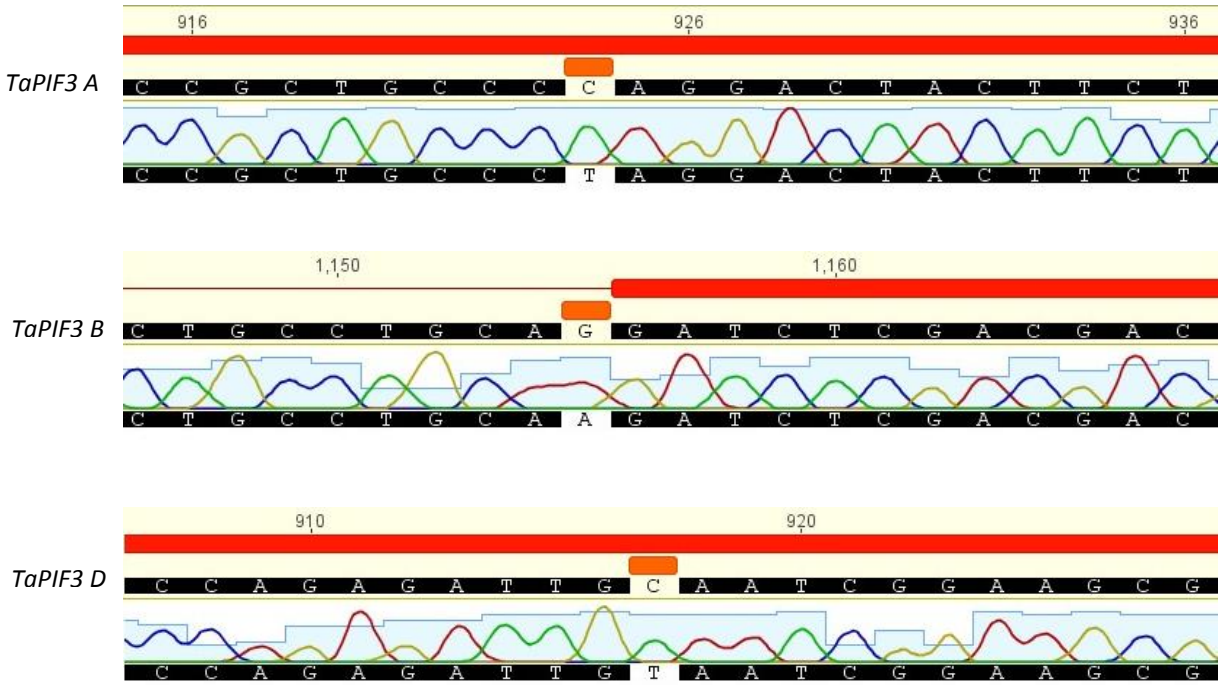


Figure 4.5 Genotyping by sequencing to identify a *TaPIF3* triple mutant. Genomic DNA was extracted from the progeny of a selfed line carrying five mutant alleles aaBbdd. Homoeologue specific PCR was used to confirm the identified mutation in each homoeologue of *TaPIF3*. PCR products were sequenced and aligned to the *TaPIF3* sequence. Figure represents sequencing of a single identified triple mutant. The coding sequence is indicated in red, position of identified mutations are annotated in orange. The genomic positions are shown above the coding sequence. The chromatograms of the sequencing reads are presented.

Of the nine *TaPIF3* homozygous triple mutants identified, seven *TaPIF3* triple mutants were identified at Rothamsted Research and two *TaPIF3* triple mutants were identified at the University of Southampton. Interestingly, the two *TaPIF3* homozygous triple mutants grown at the University of Southampton demonstrated a reduced height phenotype compared to four WT segregant plants (figure 4.6). To investigate the observed reduction in plant height, the length of the following for the six tallest tillers of each plant was measured: tiller, flag leaf, ear, emergence above flag leaf, peduncle, internode 2, internode 3 and internode 4. The measurements from the six tallest tillers were then averaged to obtain one value per plant for each measurement. Measurements taken from these lines confirmed a significant reduction in total tiller length, peduncle length and length of internode 3 in the triple mutant (figure 4.6). These observations suggest that *TaPIF3* may have a role in promoting stem elongation in wheat.

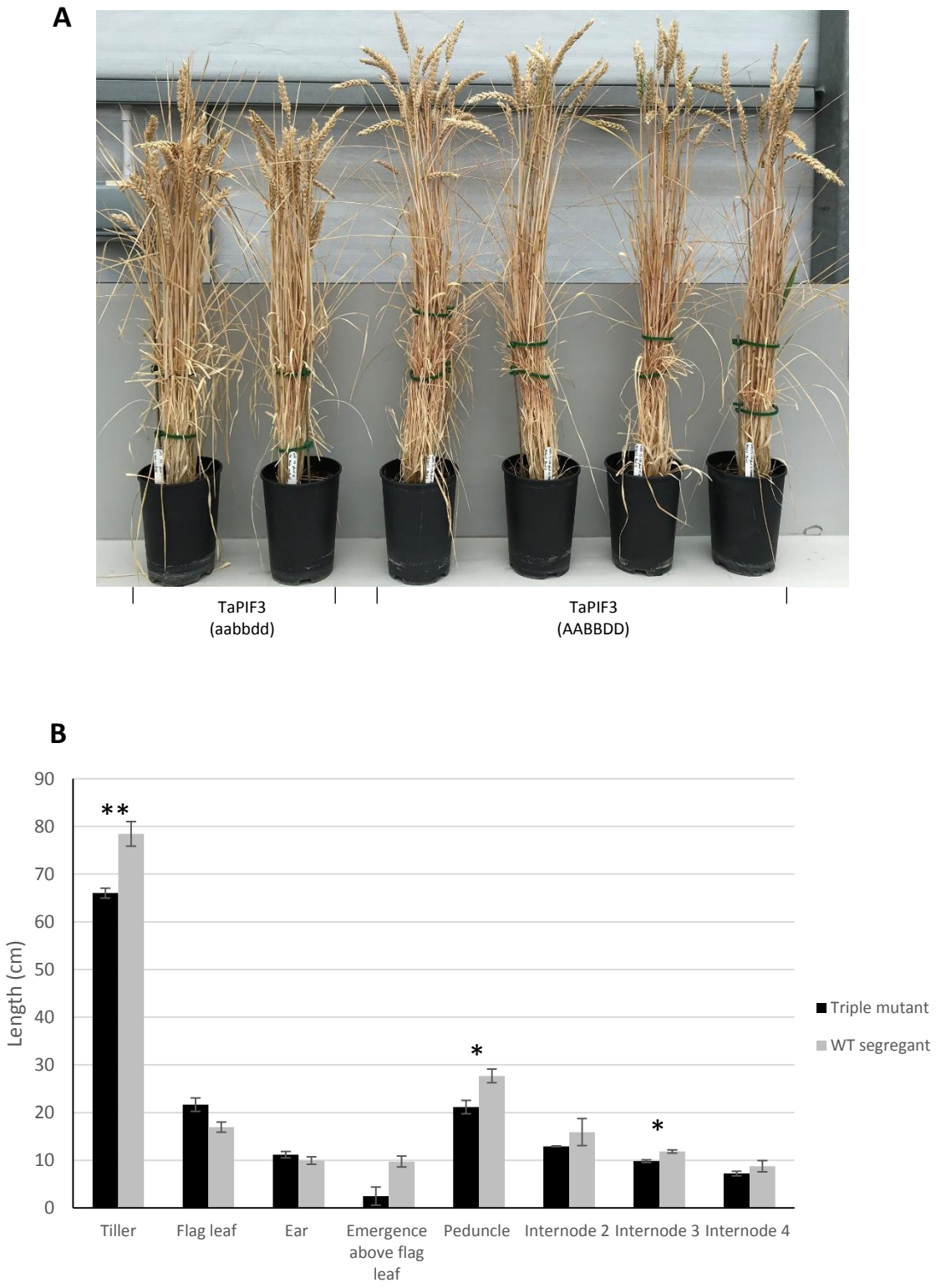


Figure 4.6 Phenotype of *TaPIF3* triple mutant. **(A)** Mature plant phenotypes of two *TaPIF3* triple TILLING mutants (aabbdd) left and four *TaPIF3* WT segregants (AABBDD) right. The triple mutant was observed to have a reduced height phenotype. **(B)** Mature plant measurements of *TaPIF3* triple mutant and *TaPIF3* WT segregant. Black bars represent mean values of two triple mutant plants (n=2). Grey bars represent mean values of four WT segregants (n=4). Error bars represent 1 standard deviation ± mean. * = P = <0.05, ** = P = <0.01.

4.2.5 Overexpression of *TaPIF3* in wheat

One approach to determine the function of *TaPIF3* was to investigate the effect of increasing expression levels of *TaPIF3* in wheat. The pRRes14.125 vector was selected for transformation of wheat cv. Cadenza. The described vector contains the promoter sequence, 5' untranslated region and first intron of the maize ubiquitin gene for stable, constitutive expression of the target gene (figure 4.7). The maize ubiquitin promoter has previously been shown to drive expression of transgenes in wheat (Weeks et al, 1993). The multiple cloning site of pRRes14.125 contains a number of restriction sites for insertion of the gene of interest. A sequence of 14 nucleotides that included the EcoRI restriction site of pRRes14.125 had been added to the 5' end of the synthesised *TaPIF3-B1* sequence. The position of the EcoRI site relative to the ATG start codon would allow insertion of *TaPIF3* in the correct reading frame. In addition, an EcoRV site was added 3' of the *TaPIF3* sequence for cloning into pRRes14.125.

To clone the TaPIF3-5xMyc sequence into the pRRes14.125 vector, TaPIF3-5xMyc was first digested from the pUC57 simple vector in a restriction digest reaction using EcoRI and EcoRV. The products of the digest were analysed on an agarose gel. A band of approximately 1700bp that represented the digested TaPIF3-5xMyc sequence was excised from the gel and purified. To enable ligation of the insert into the overexpression vector, a digestion of pRRes14.125 was performed with EcoRI and EcoRV simultaneously. The digest reaction of pRRes14.125 was run on an agarose gel to separate the vector backbone and the digested fragment of the multiple cloning site. The digestion of pRRes14.125 left the vector compatible for the ligation with TaPIF3-5xMyc. Once cloned into pRRes14.125, the *TaPIF3* overexpression construct was sent for sequencing using both a forward and reverse primer that flanked the multiple cloning site of the pRRes14.125 vector. Sequencing confirmed the TaPIF3-5xMyc sequence had been cloned correctly into the overexpression vector. For wheat transformation, 50µg of the plasmid at a concentration of 1µg/µL was required. To increase the quantity of plasmid DNA, seven individual minipreps that had been verified by sequencing were pooled. Ethanol precipitation was performed on the pooled minipreps to give a final volume of 50µL plasmid at a concentration of 2.05µg/µL. The *TaPIF3* overexpression construct was subsequently sent to the Rothamsted Research

transformation facility who used particle bombardment to transform the construct into wheat (cv Cadenza) embryos (Sparks & Jones, 2014).

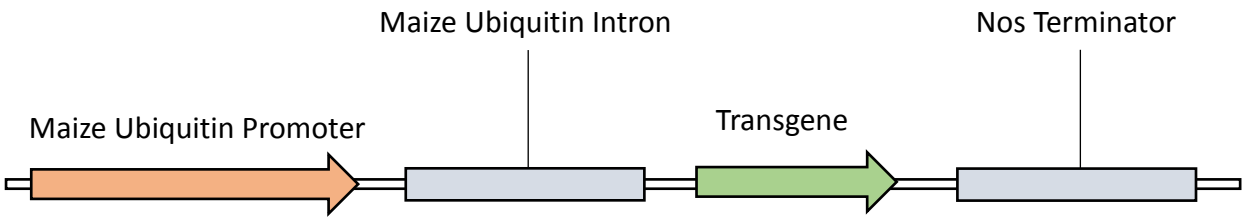


Figure 4.7 Schematic of *TaPIF3* transgene. *TaPIF3* was cloned into the pRRES14.125 vector. The maize ubiquitin promoter indicated in orange drives expression of the transgene. Transgene is indicated in green. The maize ubiquitin intron and Nos terminator are indicated by a grey box.

Two independent transformation experiments assigned B3602 and B3612 were performed. Transformation B3602 led to the identification of 39 independent transgenic lines and a control. Transformation B3612 generated 12 independent overexpression lines and a control. Control lines in both experiments were not transformed, but underwent the same tissue culture regeneration procedure as the transformed lines. A gene of interest PCR, performed at the transformation facility using a forward primer that anneals to the maize ubiquitin promoter and a reverse primer that anneals to the *TaPIF3* sequence, confirmed the presence of the *TaPIF3* transgene in the T_0 generation. PCR identified a single line from each experiment that does not contain the *TaPIF3* transgene but does contain the *bar* selection marker gene. These lines therefore provide additional controls.

T_1 seed was provided by the Rothamsted transformation facility. Seven transgenic lines were selected from the B3602 experiment in addition to the designated control and the *bar* only control. Furthermore, three transgenic lines and two controls were selected from the B3612 experiment. In both cases, five seeds per line were planted. Leaf material was harvested from each seedling for qRT-PCR based genotyping to determine transgene copy

number. Time limitations unfortunately meant that the necessary genotyping could not be completed. Three plants per line however were grown to maturity. Of interest, B3602 lines displayed a consistent increased plant height phenotype compared to controls. The increase in mature plant height was evident in two T_1 lines in particular, B3602 R6P1 and B3602 R7P1 (figure 4.8).

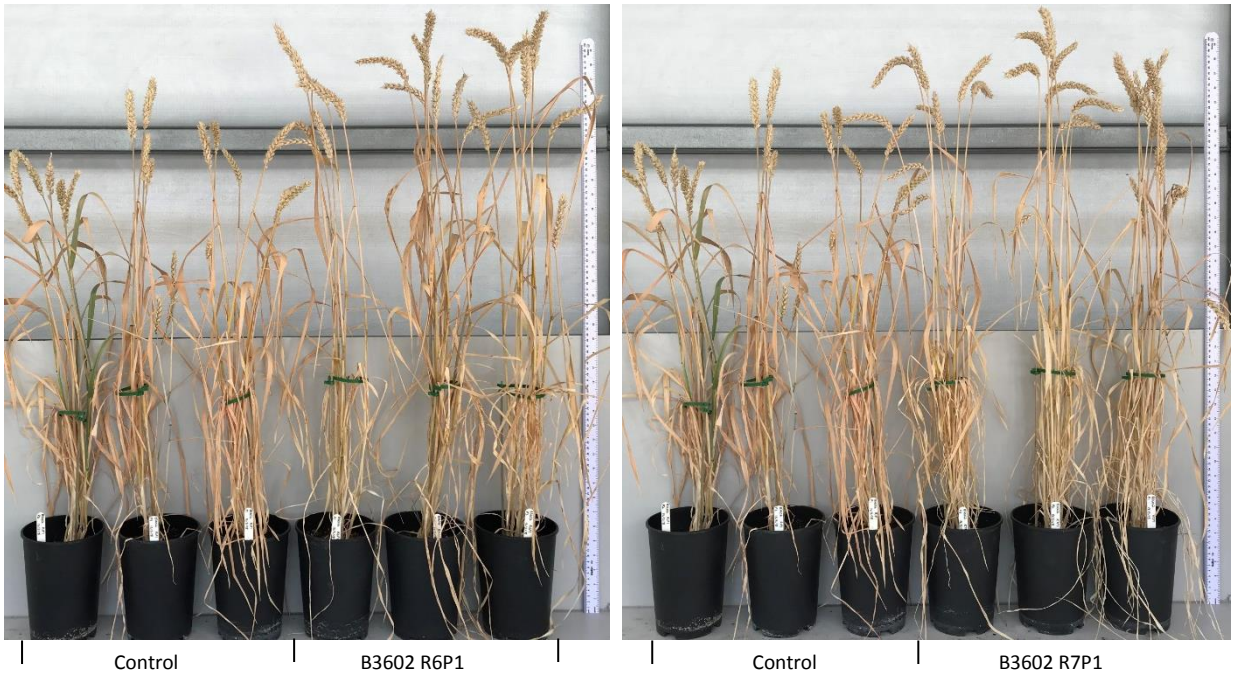


Figure 4.8 Mature plant phenotype of *TaPIF3* overexpression lines B3602 R6P1 and B3602 R7P1. T_1 plants were grown to maturity in the glasshouse. Overexpression lines B3602 R6P1 and B3602 R7P1 were observed to have a tall mature plant phenotype compared to controls. Control lines have not been transformed but have been regenerated through tissue culture at the same time.

It is not feasible to correlate an increase in plant height as a consequence of increased *TaPIF3* function without characterising these lines for copy number or expression level. One concern, is that plants demonstrating an increase in mature plant height could represent null segregants of the *TaPIF3* transgene. However, transformation by particle bombardment has been reported to result in high copy numbers (Jackson et al, 2013). High copy numbers can be the result of multiple insertions at different genomic loci or tandem repeats. Cheng

et al (1997) transformed wheat using both *Agrobacterium*-mediated transformation and particle bombardment and compared transgene copy number. Using *Agrobacterium*-mediated transformation, more than one third of transformants contained a single T-DNA insert. In contrast, from the population of plants transformed by particle bombardment only 17% contained a single copy of the transgene. Lines with high copy numbers due to multiple insertions would be less likely to produce null segregants in the T1 generation. A second concern is that overexpression lines are not expressing the *TaPIF3* transgene. The observed tall phenotype could therefore be due to other factors such as environmental variation within the glasshouse.

4.2.6 Phenotyping of TILLING knockout mutant and overexpression lines

A comparative phenotype experiment was set up to further investigate the phenotype of plants with altered *TaPIF3* expression levels. Previously, plants containing all six mutant alleles had demonstrated a reduced height phenotype compared to WT segregant controls. Furthermore, overexpression lines of *TaPIF3* had exhibited an opposite effect on plant height, with two lines in particular displaying an evident tall phenotype. Eight plants of the *TaPIF3* triple mutant and the appropriate WT segregant control were grown to maturity in the glasshouse. In addition, six plants of each overexpression line, B3602 R6P1, B3602 R7P1, the bar only control B3602 R2P3 and Cadenza wild-type were included in the experiment. Plants were set up in a completely randomised design that included latinisation to ensure an even distribution of the six lines. A randomised design was necessary to account for environmental variation within the glasshouse. The number of days to anthesis was recorded for the first tiller on each plant. Once plants had reached maturity, the tiller number was recorded. In addition, the length of the following for the three tallest tillers of each plant was measured: tiller, flag leaf, ear, emergence above flag leaf, peduncle, internode 2, internode 3 and internode 4. The ears of the three tallest tillers were harvested and spikelet number was recorded.

The measurements from the three tallest tillers of each plant were then averaged to obtain one value per plant for each measurement (table 4.1). Univariate analysis of variance (ANOVA) was applied to the data to consider the overall significance of differences between

the lines using the F test. The least significant difference (LSD) values were then used to compare means of interest at the 5% level of significance. Due to unequal replication between the lines, three LSD values were calculated. One for comparing lines with greatest replication (Triple mutant and null control), one for comparing lines with eight replicates versus lines with six replicates and a third for comparing lines with six replicates. Mean values were compared between lines of biological interest using the appropriate LSD value. Mean values for the *TaPIF3* triple mutant were compared to those of Cadenza wild-type and the WT segregant (table 4.2). Mean values for overexpression lines R6P1 and R7P1 were compared to the control line R2P3 and to Cadenza wild-type (table 4.3).

Table 4.1 A comparison of architectural traits for *TaPIF3* TILLING lines and overexpression lines. The mean value for each measurement from six lines is shown: Cadenza wild-type, *TaPIF3* triple mutant (TM), TILLING WT segregant (WT_S), overexpression line R6P1, overexpression line R7P1 and overexpression line control R2P3. An analysis of variance (ANOVA) was applied to the measurements. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM versus WT_S. LSD 2 is used to compare lines with n=6 versus lines with n=8. LSD 3 is used to compare means between lines with n=6. The degrees of freedom used for this comparison are 5 and 34.

Measurement	Cadenza (n=6)	TM (n=8)	WT_S (n=8)	R6P1 (n=6)	R7P1 (n=6)	R2P3 Control (n=6)	LSD 1	LSD 2	LSD 3
Anthesis (Days)	54.00	55.25	61.13	53.83	55.17	56.17	2.33	2.52	2.69
Tiller number	11.00	9.63	9.75	9.00	12.00	11.50	2.27	2.45	2.62
Tiller (cm)	73.27	67.09	67.27	71.66	69.29	69.74	3.81	4.11	4.40
Flag leaf (cm)	27.58	28.10	28.28	28.42	26.68	24.82	3.81	4.12	4.40
Ear (cm)	9.62	9.63	10.34	9.67	9.21	8.28	0.59	0.64	0.68
Spikelet number	19.39	19.33	18.67	18.83	18.61	17.56	1.20	1.30	1.39
Emergence above flag leaf (cm)	12.62	10.23	10.29	9.55	10.68	10.96	1.79	1.93	2.06
Peduncle (cm)	31.31	28.45	29.13	28.78	28.63	28.89	1.91	2.06	2.20
Internode 2 (cm)	15.80	14.30	13.66	15.04	14.89	15.67	1.35	1.45	1.55
Internode 3 (cm)	9.86	8.86	8.32	9.85	9.55	9.58	0.76	0.82	0.88
Internode 4 (cm)	5.77	5.24	4.93	6.46	6.09	5.83	0.96	1.04	1.11

Table 4.2 A comparison of architectural traits for *TaPIF3* TILLING lines. Mean values from three lines are shown: Cadenza wild-type, *TaPIF3* triple mutant (TM) and WT segregant (WT_S). An analysis of variance (ANOVA) was applied to the measurements. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM vs WT_S. LSD 2 is used to compare means between Cadenza and TILLING lines (n=6 vs n=8). The degrees of freedom used for this comparison are 5 and 34. Means shown in bold are significantly different from the WT_S control. Means shown underlined are significantly different from Cadenza.

Measurement	Cadenza (n=6)	WT_S (n=8)	TM (n=8)	LSD 1 (n=8 vs n=8)	LSD 2 (n=6 vs n=8)
Anthesis (Days)	54.00	61.13	**55.25**	2.33	2.52
Tiller number	11.00	9.75	9.63	2.27	2.45
Tiller (cm)	73.27	67.27	67.09	3.81	4.11
Flag leaf (cm)	27.58	28.28	28.10	3.81	4.12
Ear (cm)	9.62	10.34	**9.63**	0.59	0.64
Spikelet number	19.39	18.67	19.33	1.20	1.30
Emergence above flag leaf (cm)	12.62	10.29	10.23	1.79	1.93
Peduncle (cm)	31.31	29.13	28.45	1.91	2.06
Internode 2 (cm)	15.80	13.66	14.30	1.35	1.45
Internode 3 (cm)	9.86	8.32	8.86	0.76	0.82
Internode 4 (cm)	5.77	4.93	5.24	0.96	1.04

Table 4.3 A comparison of architectural traits for *TaPIF3* overexpression lines. Mean values from four lines are shown: Cadenza wild-type, line R6P1, line R7P1 and control line R2P3 C. An analysis of variance (ANOVA) was applied to the measurements. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 3 is used to compare means of all four lines. The degrees of freedom used for this comparison are 5 and 34. Means shown in bold are significantly different from the R2P3 Control. Means shown underlined are significantly different from Cadenza.

Measurement	Cadenza (n=6)	R2P3 C (n=6)	R6P1 (n=6)	R7P1 (n=6)	LSD 3 (n=6 vs n=6)
Anthesis (Days)	54.00	56.17	53.83	55.17	2.69
Tiller number	11.00	11.50	9.00	12.00	2.62
Tiller (cm)	73.27	69.74	71.66	69.29	4.40
Flag leaf (cm)	27.58	24.82	28.42	26.68	4.40
Ear (cm)	9.62	8.28	**9.67**	**9.21**	0.68
Spikelet number	19.39	17.56	18.83	18.61	1.39
Emergence above flag leaf (cm)	12.62	10.96	9.55	10.68	2.06
Peduncle (cm)	31.31	28.89	28.78	28.63	2.20
Internode 2 (cm)	15.80	15.67	15.04	14.89	1.55
Internode 3 (cm)	9.86	9.58	9.85	9.55	0.88
Internode 4 (cm)	5.77	5.83	6.46	6.09	1.11

The triple mutant produced significantly shorter tillers than Cadenza wild-type (table 4.2). Despite observations made in section 4.2.4, there was no significant difference in tiller length between the triple mutant and the WT segregant. Triple mutant plants had significantly shorter ears than the WT segregant but there was no difference in spikelet number. Furthermore, the triple mutant had significantly fewer days to anthesis than the WT segregant. When compared to wild-type Cadenza, triple mutants had significantly shorter emergence above the flag leaf (table 4.2). In addition, wild-type plants had a significantly longer peduncle compared to the triple mutant. At internodes 2 and 3, Cadenza wild-type was significantly longer than the triple mutant. However, no significant difference in internode lengths was detected between the triple mutant and WT segregant plants.

Comparisons were also made between overexpression lines R6P1 and R7P1 to the control line R2P3 and Cadenza wild-type. Despite observations made in section 4.2.5 no significant difference in tiller length could be detected between either of the overexpression lines and the control. Similarly, there was no difference in tiller length between the overexpression lines and Cadenza wild-type. Lines R6P1 and R7P1 had significantly longer ears than R2P3 control plants (table 4.3). When compared to wild-type, overexpression line R6P1 had a shorter emergence above the flag leaf. Furthermore, the two overexpression lines R6P1 and R7P1 had significantly shorter peduncles than wild-type Cadenza (table 4.3). No other measurements showed a significant difference.

The number of days to anthesis and tiller number for all lines are compared in figure 4.9. Figure 4.9 shows that the WT segregant has a greater number of days to anthesis than the triple mutant. Overexpression line R6P1 demonstrates a reduced tiller number compared to its control line R2P3 and also to wild-type Cadenza (figure 4.9). The difference between these lines however is not significant at the 5% level. Total tiller length and flag leaf length are compared in figure 4.10 which shows that the triple mutant and the WT segregant have the shortest tillers. Overexpression line R6P1 shows a small increase in tiller length compared to the R2P3 control line. Figure 4.11 clearly demonstrates *TaPIF3* triple mutants have a reduced ear length compared to the WT segregant. In contrast, overexpression lines R6P1 and R7P1 have significantly longer ears than their control line. Despite the differences in ear length, figure 4.11 shows no difference in spikelet number between the lines. Emergence above the flag leaf and peduncle length are compared in figure4.12**,** which

shows wild-type Cadenza plants have the greatest emergence above the flag leaf. Wild-type plants also have the longest peduncle. The length of internodes 2-4 for all lines are shown in figure 4.13 which shows that the triple mutant demonstrates a reduction in length of internode 2 and internode 3 when compared to WT.

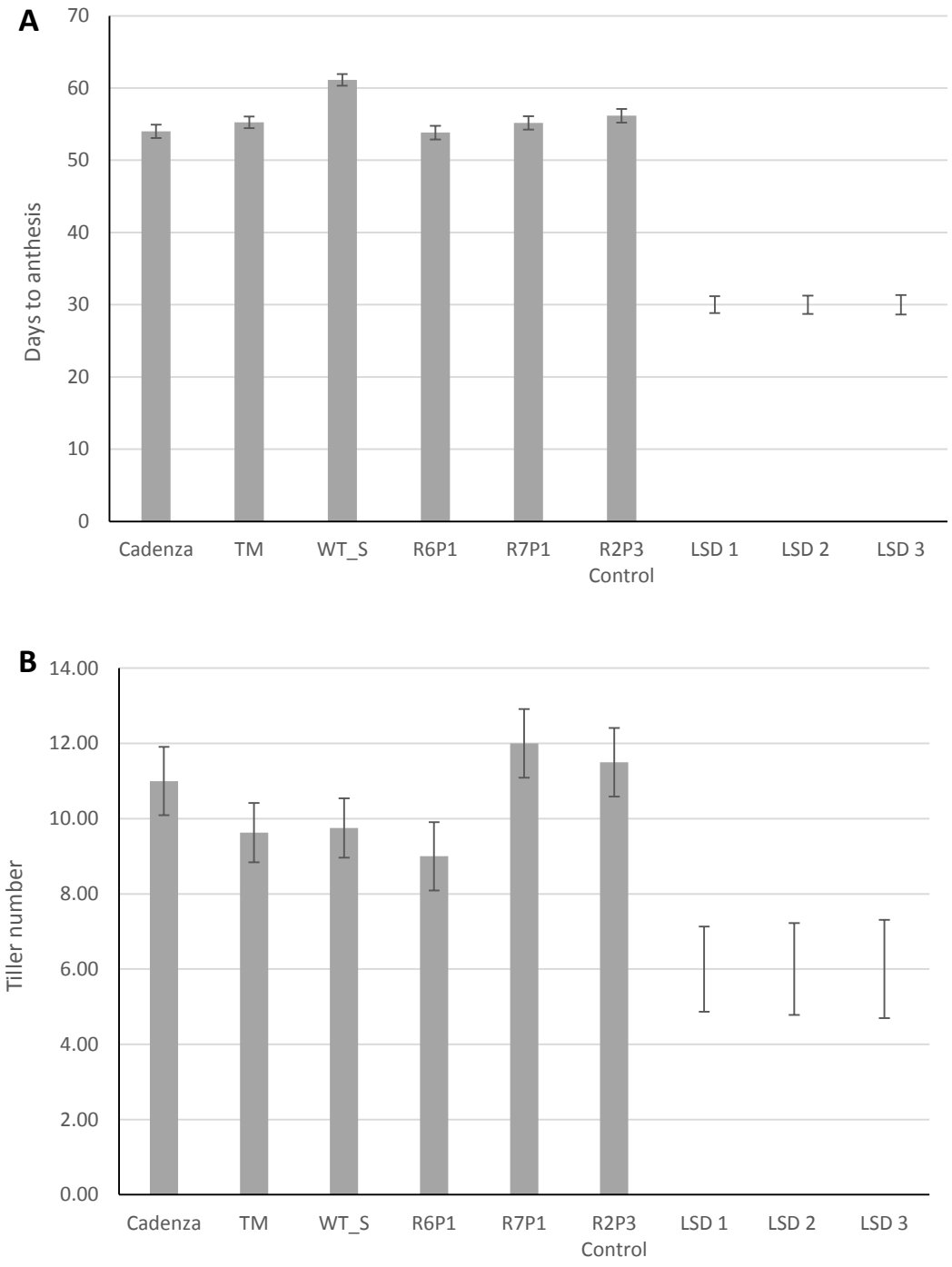


Figure 4.9 Days to anthesis and tiller number in *TaPIF3* TILLING mutant and overexpression lines. The mean values for days to anthesis from date of sowing **(A)** and tiller number **(B)** for Cadenza wild-type, *TaPIF3* TILLING triple mutant (TM), *TaPIF3* TILLING WT segregant (WT_S), overexpression lines R6P1 and R7P1 and overexpression line control R2P3 are shown. Error bars refer to the standard error. Analysis of variance was applied to measurements from all six lines. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM versus WT_S. LSD 2 is used to compare means of TM and WT_S versus others. LSD 3 is used to compare means between all other lines. The degrees of freedom used for this comparison are 5 and 34.

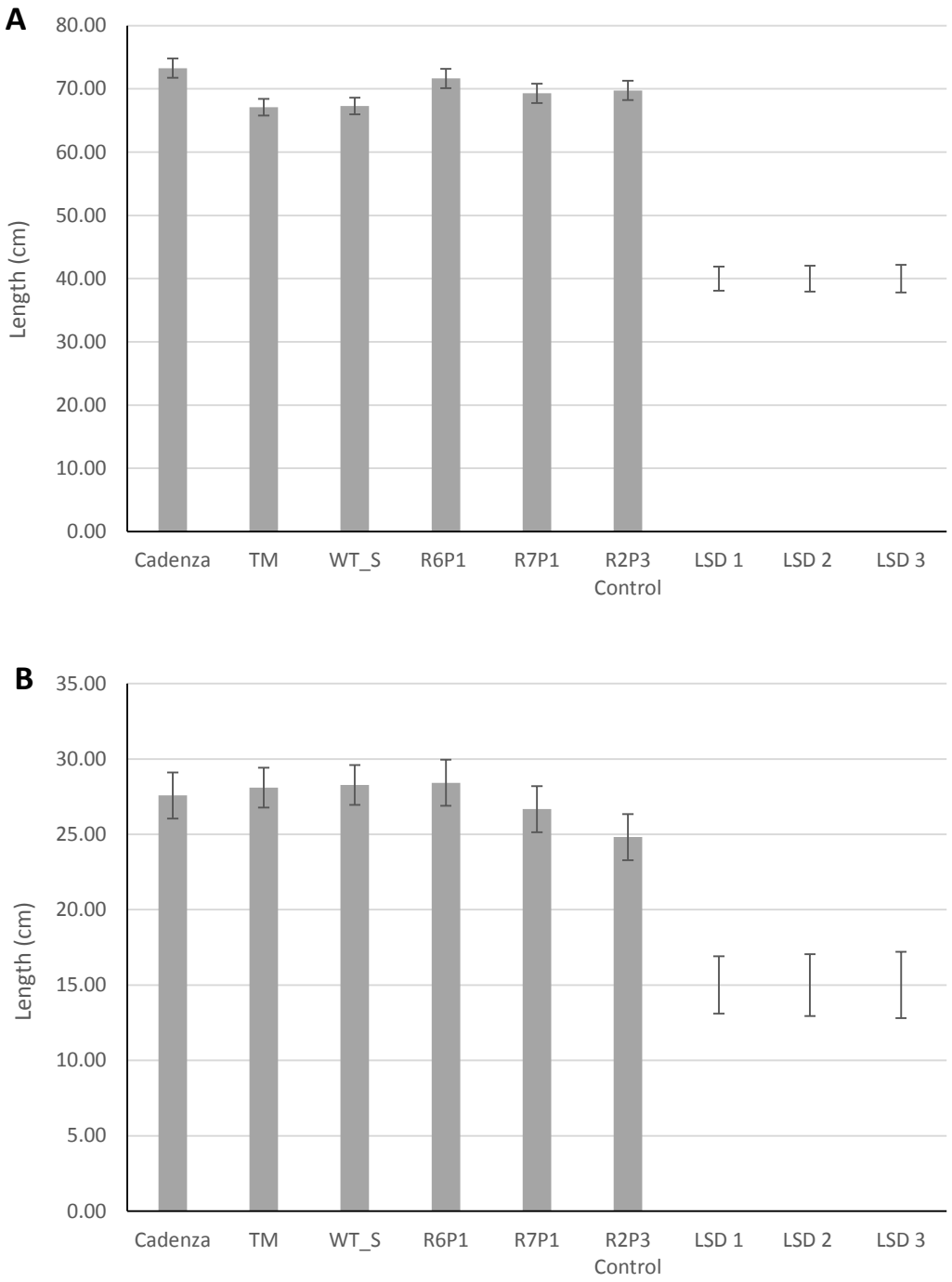


Figure 4.10 Tiller length and flag leaf length in *TaPIF3* TILLING mutant and overexpression lines. The mean values for tiller length **(A)** and flag leaf length **(B)** for Cadenza wild-type, *TaPIF3* TILLING triple mutant (TM), *TaPIF3* TILLING WT segregant (WT_S), overexpression lines R6P1 and R7P1 and overexpression line control R2P3 are shown. Error bars refer to the standard error. Analysis of variance was applied to measurements from all six lines. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM versus WT_S. LSD 2 is used to compare means of TM and WT_S versus others. LSD 3 is used to compare means between all other lines. The degrees of freedom used for this comparison are 5 and 34.

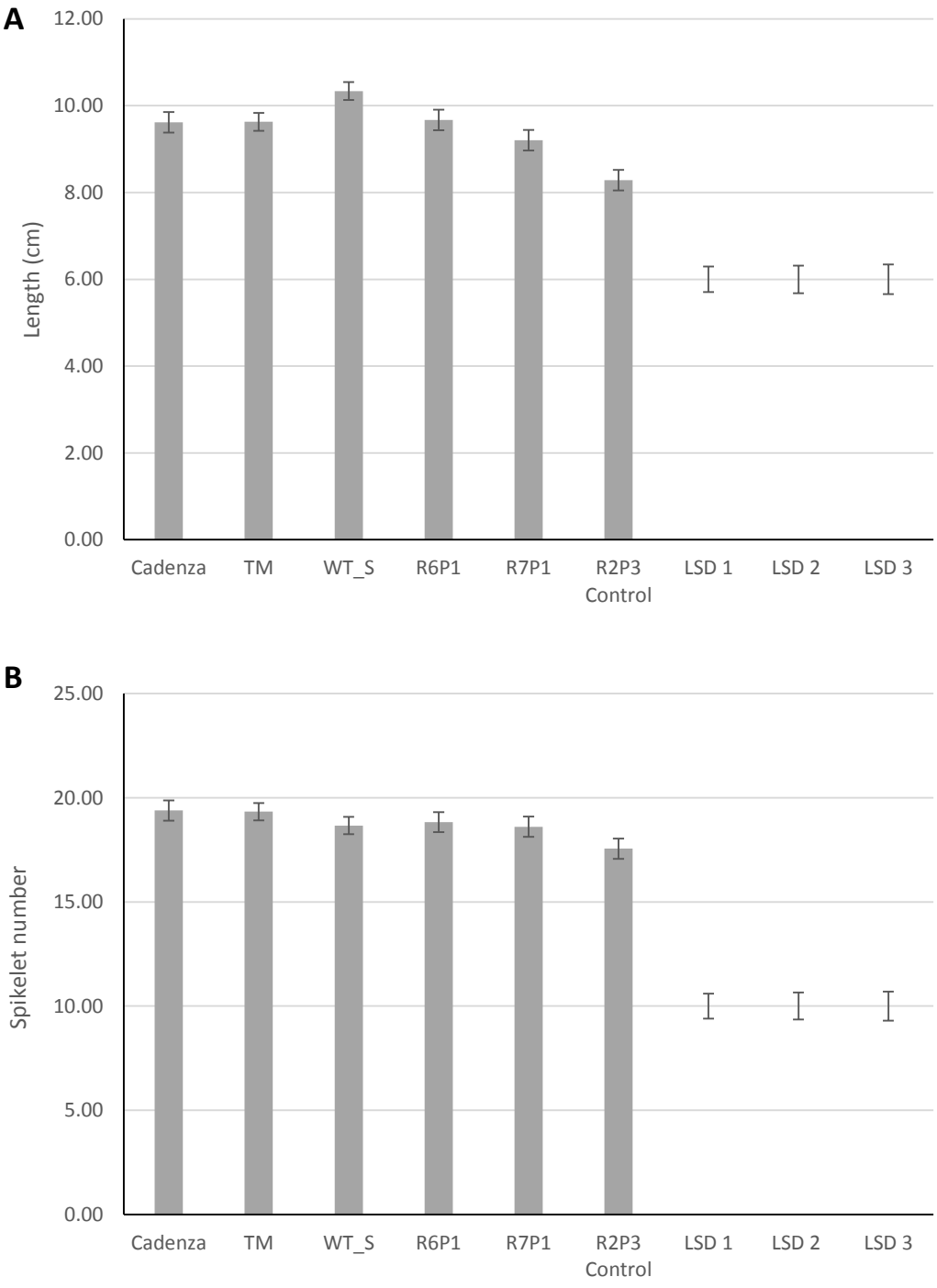


Figure 4.11 Ear length and spikelet number in *TaPIF3* TILLING mutant and overexpression lines. The mean values for ear length **(A)** and spikelet number **(B)** for Cadenza wild-type, *TaPIF3* TILLING triple mutant (TM), *TaPIF3* TILLING WT segregant (WT_S), overexpression lines R6P1 and R7P1 and overexpression line control R2P3 are shown. Error bars refer to the standard error. Analysis of variance was applied to measurements from all six lines. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM versus WT_S. LSD 2 is used to compare means of TM and WT_S versus others. LSD 3 is used to compare means between all other lines. The degrees of freedom used for this comparison are 5 and 34.

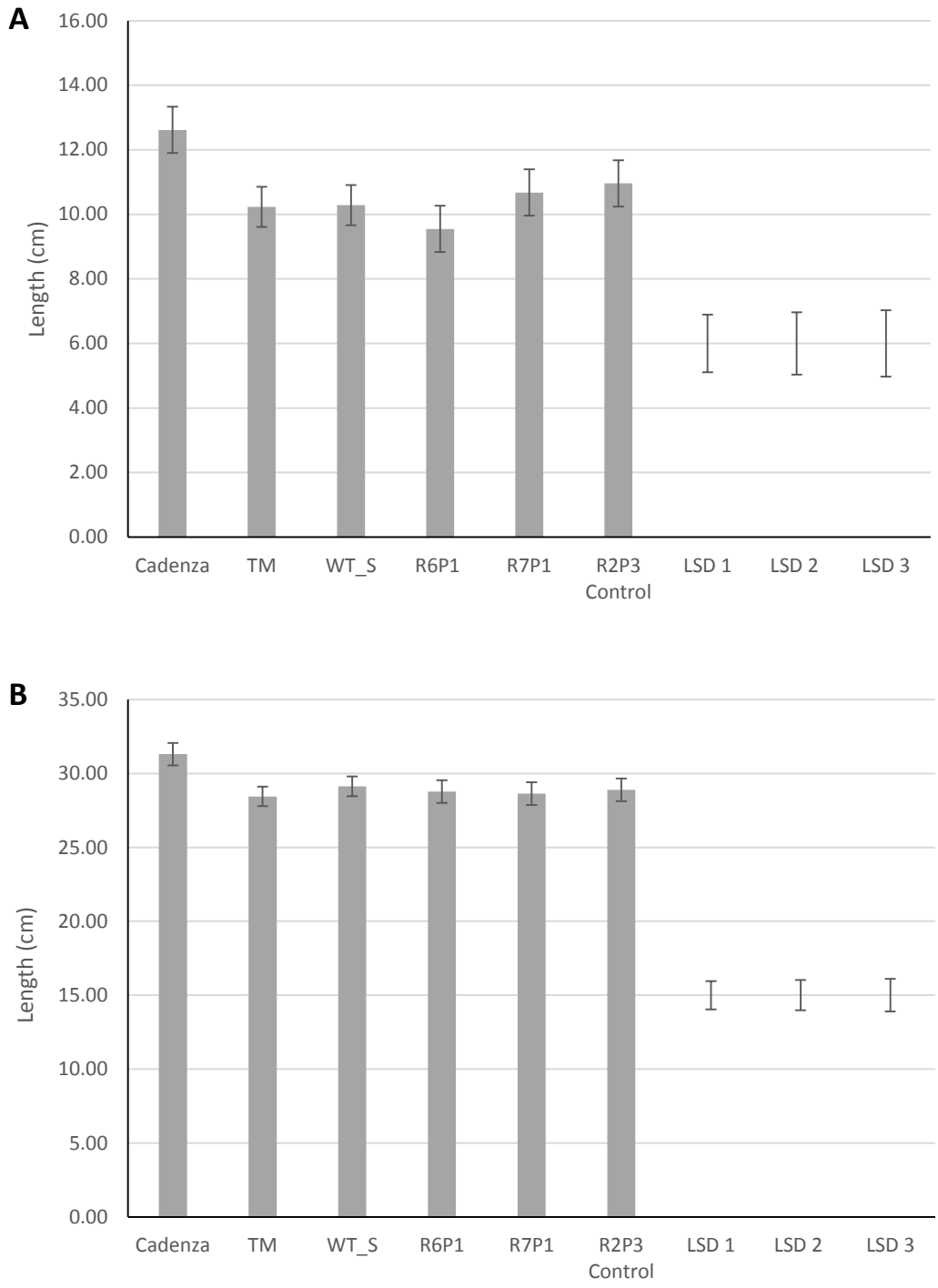
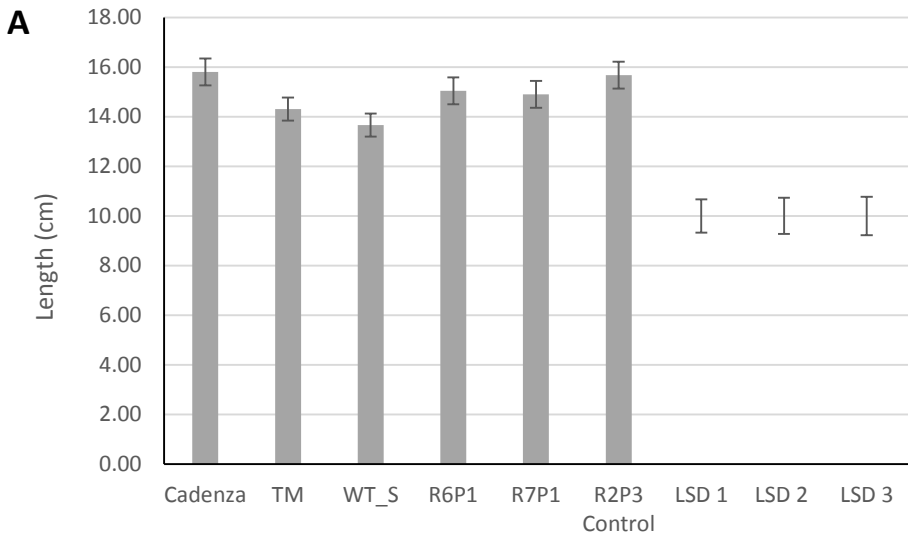
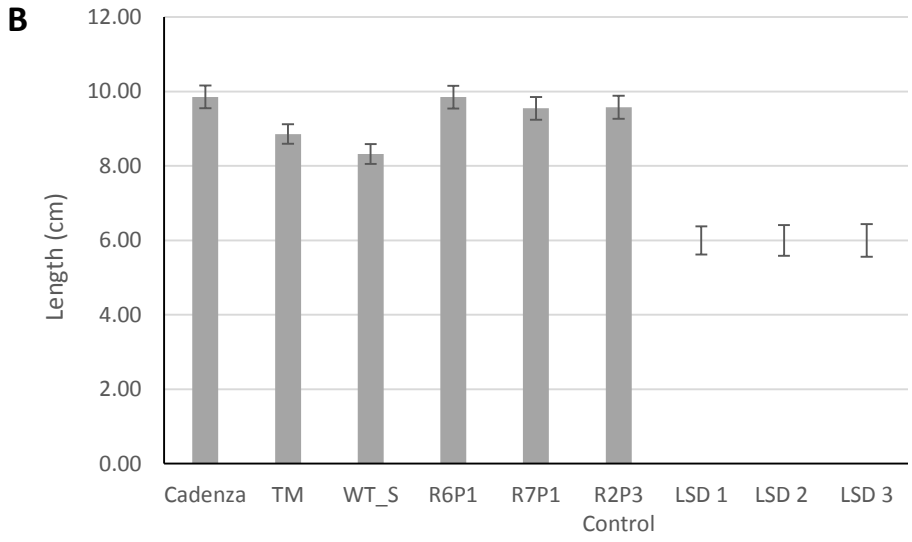


Figure 4.12 Emergence above flag leaf and peduncle length in *TaPIF3* TILLING mutant and overexpression lines. The mean values for emergence above flag leaf **(A)** and peduncle length **(B)** for Cadenza wild-type, *TaPIF3* TILLING triple mutant (TM), *TaPIF3* TILLING WT segregant (WT_S), overexpression lines R6P1 and R7P1 and overexpression line control R2P3 are shown. Error bars refer to the standard error. Analysis of variance was applied to measurements from all six lines. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM versus WT_S. LSD 2 is used to compare means of TM and WT_S versus others. LSD 3 is used to compare means between all other lines. The degrees of freedom used for this comparison are 5 and 34.


A
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Length (cm)
Cadenza
TM
WT_S
R6P1
R7P1
R2P3
Control
LSD 1
LSD 2
LSD 3


B
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Length (cm)
Cadenza
TM
WT_S
R6P1
R7P1
R2P3
Control
LSD 1
LSD 2
LSD 3

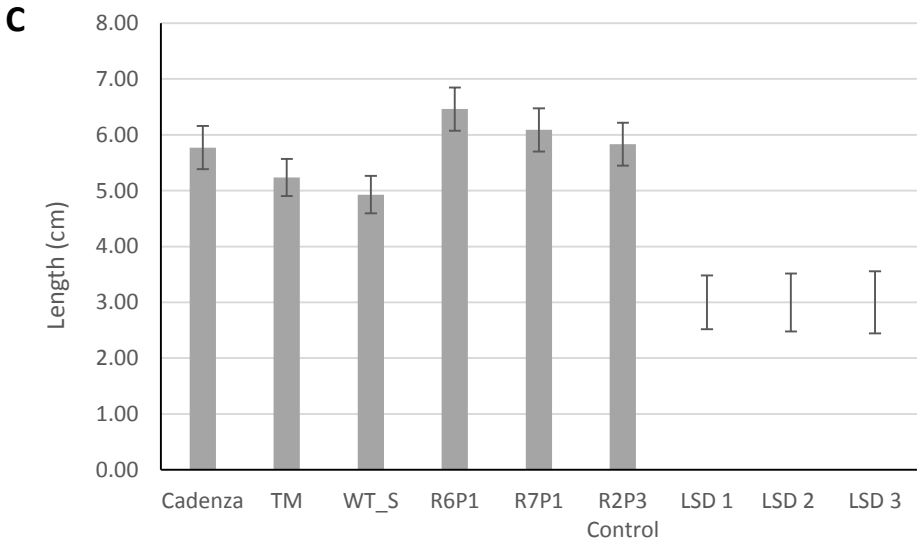

C
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
Length (cm)
Cadenza
TM
WT_S
R6P1
R7P1
R2P3
Control
LSD 1
LSD 2
LSD 3

Figure 4.13 Lengths of internodes 2, 3 and 4 in *TaPIF3* TILLING mutant and overexpression lines. The mean values for internode 2 **(A),** internode 3 **(B)** internode 4 **(C)** for Cadenza wild-type, *TaPIF3* TILLING triple mutant (TM), *TaPIF3* TILLING WT segregant (WT_S), overexpression lines R6P1 and R7P1 and overexpression line control R2P3 are shown. Error bars refer to the standard error. Analysis of variance was applied to measurements from all six lines. Least significant difference values (LSD) at the 5% level of significance are shown. LSD 1 is used to compare means of TM versus WT_S. LSD 2 is used to compare means of TM and WT_S versus others. LSD 3 is used to compare means between all other lines. The degrees of freedom used for this comparison are 5 and 34.

Images of the six lines selected for the comparative phenotype experiment are shown in figure 4.14 and figure 4.15. The reduction in tiller length demonstrated by the triple mutant compared to Cadenza wild-type is clearly shown in figure 4.14, in which triple mutant plants are photographed alongside wild-type controls. Figure 4.14 illustrates the small reduction in plant height of the triple mutant compared to the WT segregant. This small reduction is in contrast to the evident difference in height between the two lines previously shown in figure 4.6.

Images of the two overexpression lines, R6P1 and R7P1, are shown in figure 4.15**.** Figure 4.15 illustrates a clear increase in tiller length between overexpression line R6P1 and the control line R2P3. This difference however is not significant at the 5% level of significance. Figure 4.15 shows that there is no difference in height between overexpression line R7P1 and the control R2P3.

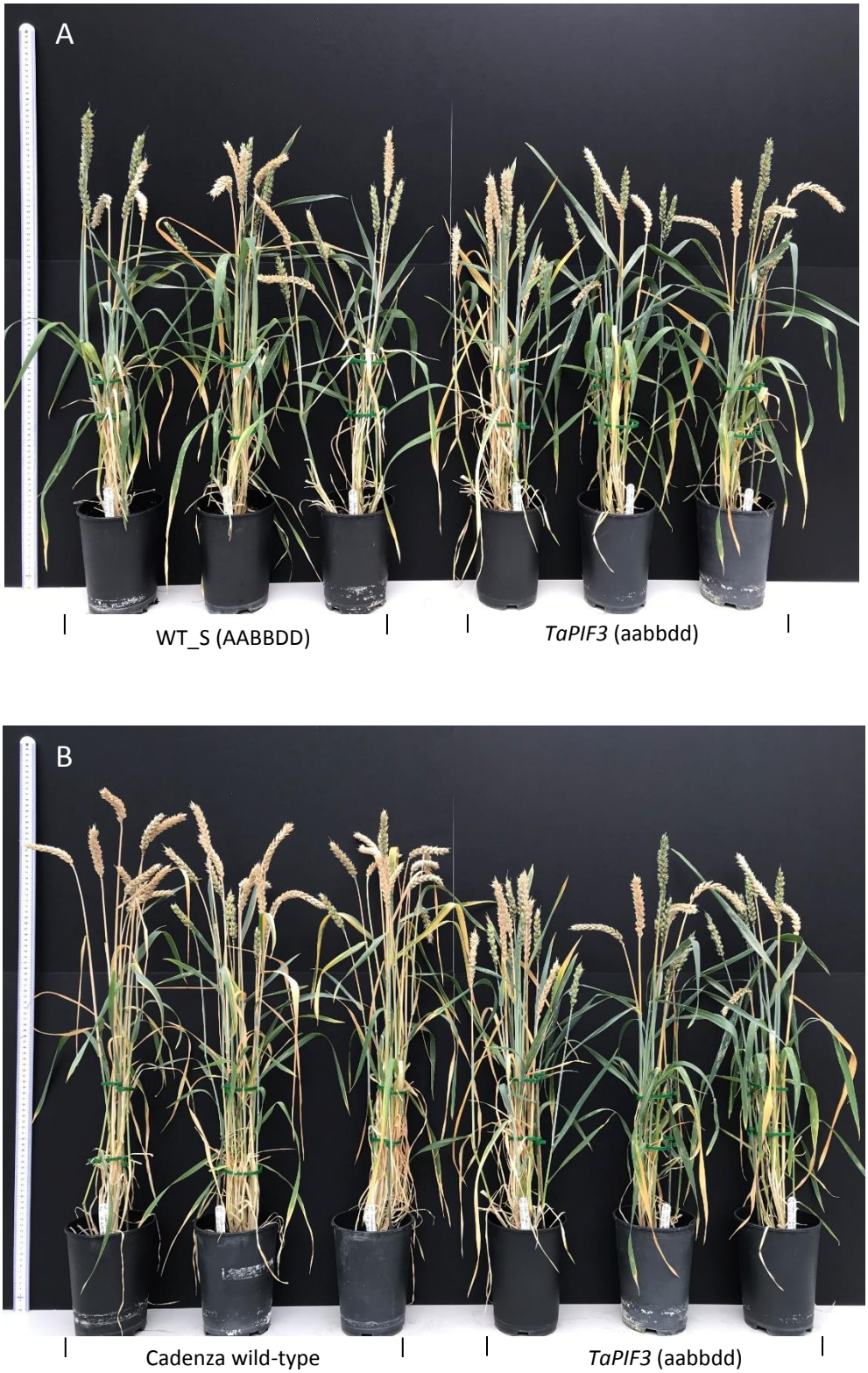


Figure 4.14 Phenotypic comparison of the *TaPIF3* triple mutant. *TaPIF3* triple mutant plants from the comparative phenotype experiment were photographed with control lines. Plants were photographed approximately 4 weeks post anthesis (A) *TaPIF3* triple mutants with the WT segregant controls (WT_S). (B) *TaPIF3* triple mutants with Cadenza wild-type plants. A meter rule is displayed on the left in each photograph.

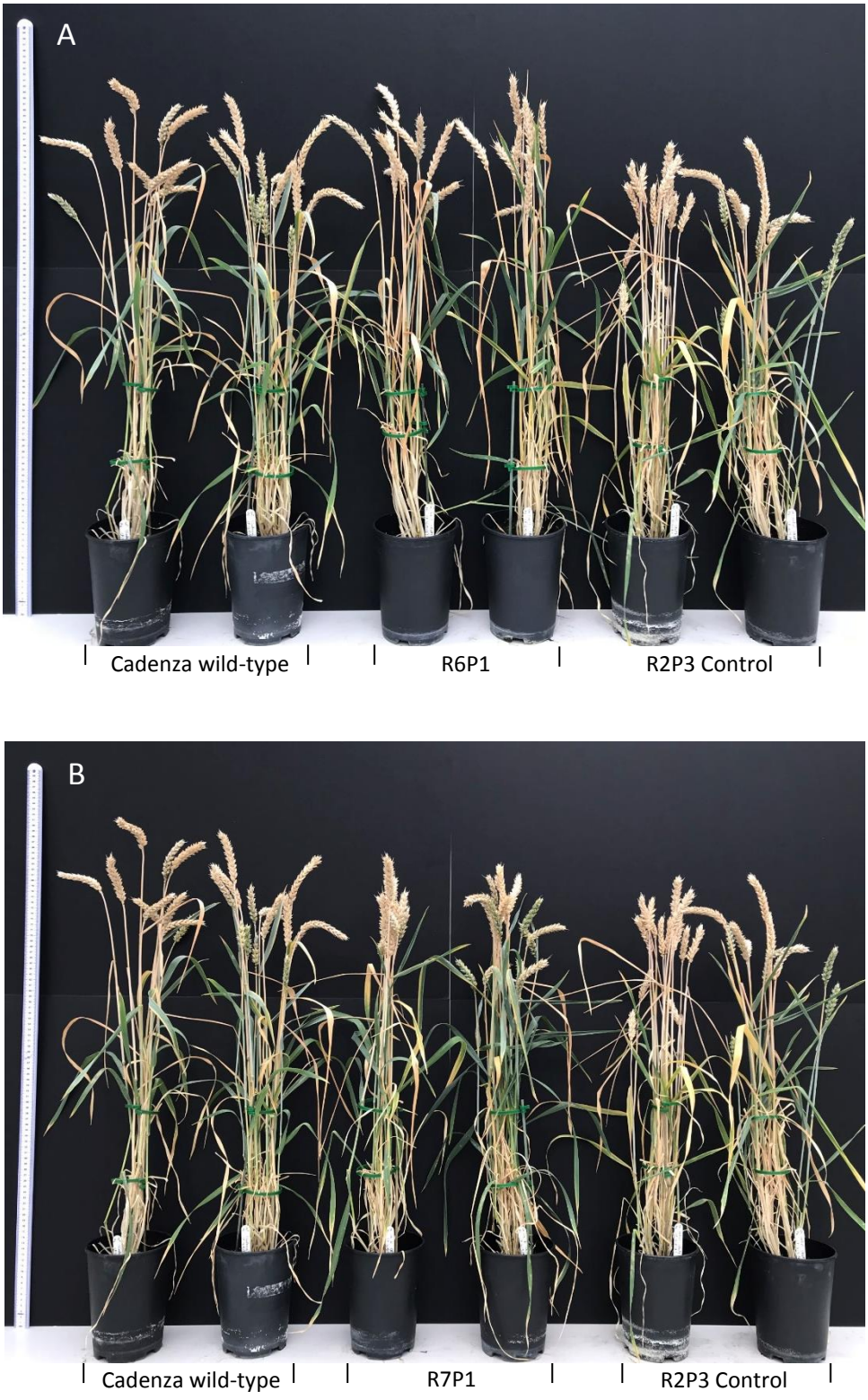


Figure 4.15 Phenotypic comparison of *TaPIF3* overexpression lines. *TaPIF3* overexpression lines from the comparative phenotype experiment were photographed with control lines. Plants were photographed approximately 4 weeks post anthesis (A) Overexpression line R6P1 with control line R2P3 and wild-type Cadenza. (B) Overexpression line R7P1 with control line R2P3 and wild-type Cadenza. A meter rule is displayed on the left in each photograph.

4.2.7 Interaction between TaPIF3 and the GA signalling pathway

The results presented in this chapter have indicated that TaPIF3 has a potential role in regulating stem elongation in wheat. In Arabidopsis, PIFs are negatively regulated through the direct physical association of DELLA proteins (de Lucas et al, 2008, Feng et al, 2008). In view of the fact that DELLA proteins are negative regulators of GA-dependent stem growth, it is conceivable that in wheat, DELLA proteins may bind to and inhibit the activity of TaPIF3 to prevent stem elongation. To date, an interaction between DELLAs and PIFs has only been documented in Arabidopsis. Establishing whether an interaction occurs between RHT-1 and TaPIF3 could provide evidence that a conserved signalling mechanism occurs in wheat.

In addition, an interaction between TaPIF3 and a protein termed TaPIL2 will be examined. *TaPIL2* was identified as a putative PHYTOCHROME INTERACTING FACTOR-LIKE gene in wheat due to its sequence similarity with Arabidopsis and rice *PIF* sequences (Wallace, 2017). TaPIL2 contains a conserved bHLH domain and critically a conserved APB motif. Specifically, *TaPIL2* was identified as a putative orthologue of OsPIL1/OsPIL13. In Arabidopsis, PIFs have been documented to form homodimers and heterodimers (Pham et al, 2018). The ability of TaPIF3 and TaPIL2 to interact could therefore provide some information on how these two proteins function.

Protein interaction studies with TaPIF3 were performed using the Invitrogen ProQuest Two-Hybrid System. This system uses the expression of the reporter genes *LacZ* and *HIS3* to indicate an interaction between two proteins. The bait and prey plasmids contain selective marker genes required for growth on media lacking leucine and tryptophan respectively. Prior to performing a specific yeast two-hybrid assay, the protein fused to the DNA binding domain of GAL4 must be tested for self-activation. Self-activation can occur if the protein of interest, when fused to the DNA binding domain, can activate transcription of the reporter gene. Previously, self-activation has been a problem when using DELLAs as a bait protein due to their proposed function of being transcriptional activators. This problem has been avoided in past studies by only expressing the C terminal GRAS domain (de Lucas et al, 2008). The degree of self-activation can be determined by quantifying the concentration of the HIS3 inhibitor, 3-amino-1,2,4-triazole (3-AT) needed to repress growth on media lacking histidine.

4.2.7.1 Cloning of *TaPIF3* constructs used in a yeast two-hybrid assay

To test for a specific interaction between RHT-1 and TaPIF3, the *TaPIF3-B1* CDS was cloned into the pDEST32 bait plasmid and the pDEST22 prey plasmid. Firstly, the *TaPIF3-B1* sequence was amplified from a pENTR1A-TaPIF3-5xMyc entry clone generated in section 5.2.1. To remove the 5xMyc sequence from the 3' end of *TaPIF3*, a reverse primer was designed that anneals to the 3' end of the *TaPIF3* CDS. Three bases complementary to the TAA stop codon of *TaPIF3-B1* were added to the 5' end of the reverse primer thereby excluding the 5xMyc sequence. In addition, the 25 bp attB1 and attB2 sites were added to the 5' end of the forward and reverse primers respectively to make the PCR product compatible with the Gateway recombination cloning system (Invitrogen, CA, USA).

Amplification of the *TaPIF3* CDS was performed using primers TaPIF3_attB1_F1 and primers TaPIF3_attB2_R1 in a Phusion PCR reaction (see table S.1 for primer sequences). A product of expected size, 1.5 kb, was excised from the gel and purified. A BP recombination reaction was subsequently performed using the attB-flanked PCR product and the Gateway Donor vector pDONR221 to generate an entry clone. A restriction digest reaction and subsequent sequencing confirmed the presence of the *TaPIF3* sequence in the entry vector. To generate the *TaPIF3-B1* ait and prey plasmids, the *TaPIF3* sequence was recombined into the pDEST32 bait plasmid and the pDEST22 prey plasmid in independent LR recombination reactions. Bait and prey plasmids were analysed by a restriction digest reaction and sequencing to confirm correct insertion of the *TaPIF3-B1* gene.

4.2.7.2 RHT-1 does not interact with TaPIF3 in yeast two-hybrid assays

A specific yeast two-hybrid assay was performed by Dr Stephen Thomas to test for an interaction between RHT-1 and TaPIF3. Due to the full-length RHT-D1A protein causing a high level of self-activation, a C-terminal truncated RHT-D1A product was used to detect an interaction with *TaPIF3*. The described truncated RHT-1 protein consists of 401 amino acids, lacking 217 amino acids from the N-terminus. The truncated product contains the C-terminal domains required for protein-protein interactions.

The truncated RHT-D1A coding sequence was cloned into the pDEST32 bait plasmid. The resulting RHT-D1A bait plasmid was transformed into Mav203 competent yeast cells with the pDEST22-TaPIF3 prey plasmid. In addition, the pDEST32-TaPIF3-B1 ait construct was transformed into yeast with a TaPIL2 bait plasmid. Yeast was also transformed with empty bait and prey plasmids for use as negative controls. The yeast strain Mav203 used in this experiment expresses a basal level of *HIS3*. Furthermore, it is common for most bait proteins to contain some level of transcriptional activity. As such, even negative controls will initiate some transcription of *HIS3*. The yeast strains shown in table 4.4 were spotted on –Leu/-Trp/-His media containing concentrations of 3-AT ranging from 0 mM to 100Mm. Plates were incubated for 72 hours and then assessed for growth of yeast strains.

Table 4.4 Yeast strains used to detect protein interactions with TaPIF3. The bait and prey constructs transformed into each yeast strain are shown.

Strain	Bait	Prey	
1	Empty	Empty	Negative control
2	Empty	TaPIF3	Negative control
3	RHT-D1A	Empty	Self-activation control
4	RHT-D1A	TaPIF3	Test
5	Empty	Empty	Negative control
6	Empty	TaPIF3	Negative control
7	TaPIF3	Empty	Self-activation control
8	TaPIF3	TaPIL2	Test

At 0 mM 3-AT all eight yeast strains grew on –Leu/-Trp/-His media including the negative controls strains (figure 4.16). Increasing the 3-AT concentration to 10 mM reduced the growth of the negative control strains containing empty bait and prey plasmids. By 25 mM 3-AT, the empty bait and prey vector controls and the RHT-D1A self-activation control showed very low levels of detectable growth. Of importance strain 4, transformed with the RHT-D1A bait plasmid and the TaPIF3 prey plasmid, showed a comparable level of growth to

the negative control strains and the RHT-D1A self-activation control which suggests that the truncated RHT-D1 and TaPIF3 proteins do not interact in yeast (figure 4.16). Increasing the 3-AT concentration to 50 mM almost completely abolishes the growth of strain 4 (figure 4.16). This preliminary study indicates that RHT-1 and TaPIF3 do not interact.

In contrast, the level of growth of the TaPIF3 self-activation control and the TaPIF3-TaPIL2 strain remained strong up to 100 mM 3-AT (figure 4.16). This demonstrates that the level of self-activation of TaPIF3 is very high and it is therefore an unsuitable bait construct for testing interactions in yeast two-hybrid assays. Consequently, no conclusions can be made about an interaction between TaPIF3 and TaPIL2.

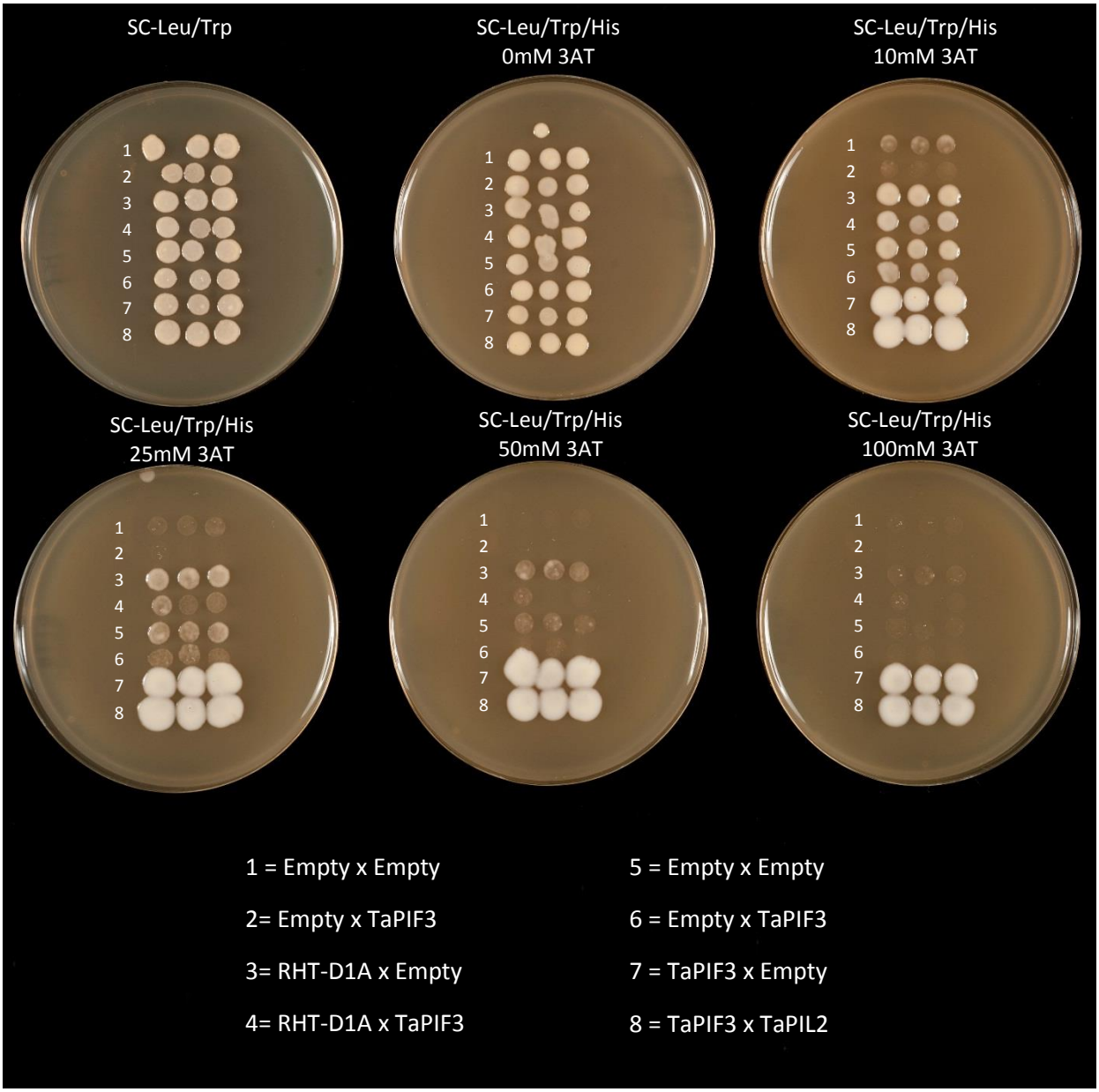


Figure 4.16 Yeast two-hybrid assay investigating an interaction between RHT-D1A and TaPIF3. Mav203 competent yeast cells were transformed with GAL4 DNA-binding bait plasmids and GAL4 activation domain prey plasmids. Expression of the reporter gene *HIS3* was detected by spotting strains on –Leu/-Trp/-His media containing concentrations of 3-amino-1,2,4-triazole (3-AT) ranging from 0-100 mM. Strains were spotted on each plate with three technical replicates. Strains used in the assay are labelled below indicating the bait protein and the prey protein expressed in each strain (Bait x Prey).

4.3 Discussion

The aim of this chapter was to investigate the function of *TaPIF3-B1* by exploiting a reverse genetics approach. Two methods, overexpression and knockout TILLING lines, have been utilised in order to robustly characterise the role of *TaPIF3*. However, production of these tools has been time-consuming, largely due to the hexaploid nature of the bread wheat genome. Nevertheless, these approaches have led to the production of novel genetic resources that have allowed us to investigate the phenotype of lines that have manipulated levels of TaPIF3

4.3.1 Generation of genetic resources to characterise *TaPIF3* function

TILLING lines that contain introduced nonsense premature stop codons or splice site mutations within the homoeologous sequences of *TaPIF3* have been identified and stacked to generate a *TaPIF3* knockout mutant. A single mutation with a putative knockout effect was selected for each homoeologue of *TaPIF3* (figure 4.1). Mutations were subsequently stacked to produce a homozygous triple mutant by crossing independent TILLING lines. The A homoeologue mutation of *TaPIF3* encodes for a premature stop codon located within exon 2 of the coding sequence. The location of the mutation is 5' of the bHLH coding region. A putative N-terminal truncated protein product would therefore be absent of the bHLH domain, disrupting DNA binding ability and dimerization of *TaPIF3*. The splice site mutation within the B homoeologue is located at the 3' acceptor site of intron 2. The G/A substitution disrupts the conserved dinucleotide AG acceptor site that is important for correct gene splicing. Exon 3 that flanks the B homoeologue mutation encodes for the bHLH domain. Alteration of the coding sequence at this region would therefore be very likely to disrupt function. The D homoeologue mutation that is predicted to cause loss-of-function causes a premature stop codon within exon 2. In a similar manner to the A homoeologue, the nonsense mutation is located 5' of the bHLH coding sequence and so is expected to abolish *TaPIF3* function. Crossing of described *TaPIF3* TILLING lines identified lines homozygous for all three alleles (aabbdd). In addition, a WT segregant line was identified (AABBDD) along with all combinatorial mutants of *TaPIF3.*

However, despite all three mutant alleles of *TaPIF3* having a putative knockout effect, the impact of each mutation *in vivo* has not yet been confirmed. Attempts have been made to confirm the effect of the B splice acceptor mutation using PCR. Homoeologue specific primers were designed to amplify a short amplicon that spans intron 2 of the B homoeologue using a cDNA template. If exon 2 is retained in the mRNA message, a PCR product of increased size would be expected. In addition, primer 1BS_forward_till_1 that was successfully used to genotype the B mutation (section 4.2.3) was tested with a homoeologue specific reverse primer. Primer 1BS_forward_till_1 anneals to the non-coding sequence of intron 2. Therefore, amplification of a PCR product from a cDNA template would only be expected if intron 2 had been retained as an effect of the splice acceptor mutation. Amplification of *TaPIF3* at this region however has proved unsuccessful. One reason for this could be the high GC content of the *TaPIF3* sequence as discussed in chapter 3**.** A second reason may be due to the limitations imposed by designing homoeologue specific primers. Designing primers that are specific to a single homoeologue limits the position and sequence of potential primers. An alternative strategy to confirm the effect of the B splice acceptor mutation would be to amplify a generic *TaPIF3* product and then carry out next generation sequencing of the PCR product pool. This approach would allow more flexibility in designing primers.

In addition, the *TaPIF3* triple mutant and the WT segregant line have not yet been backcrossed to wild-type due to time constraints. Backcrossing of these lines is important to remove additional mutations introduced into the genome by EMS that could cause or exacerbate a phenotype. Ideally, these lines would have been backcrossed prior to a comprehensive phenotype analysis of BC_2F_3 material. However, a robust phenotypic experiment has been performed with the described TILLING lines, which has identified some interesting phenotypic differences.

4.3.2 TaPIF3 has a potential role in stem elongation

Initial comparisons between two *TaPIF3* homozygous triple mutants and the WT segregant indicated that knockout of *TaPIF3* leads to a decrease in plant height (figure 4.6). Measurements of these plants revealed a decrease in peduncle length and reduced

emergence above the flag leaf in the triple mutant. Observations that *TaPIF3* may regulate stem elongation suggests a conserved function of PIFs in Arabidopsis and wheat. In Arabidopsis, members PIF4 and PIF5 are known to have a prominent role in regulating hypocotyl elongation as part of the shade avoidance response (Lorrain et al, 2008). PIF4 is also implicated in promoting elongation growth in response to high temperature (Franklin et al, 2011). PIF3 has been shown to be required for ethylene-induced hypocotyl elongation in the light (Zhong et al, 2012). In addition, PIL proteins in rice have also been shown to regulate stem elongation. Todaka et al (2012) demonstrated that overexpression of OsPIL1 leads to an increase in stem elongation and mature plant height.

The role of *TaPIF3* in regulating elongation in wheat was further scrutinised by a mature plant phenotype experiment that included the *TaPIF3* triple mutant and the WT segregant line with increased replication. Despite previous observations, the triple mutant did not show a reduced height phenotype compared to the WT segregant. Figure 4.10 shows triple mutant lines display a very slight reduction in final plant height. This reduction was not significant at the 5% level of significance.

One explanation for why a difference in plant height between the triple mutant and the WT segregant was not observed in the phenotypic experiment is due to variable environmental conditions that were experienced when growing them in a glasshouse. Triple mutant lines and WT segregants photographed in figure 4.6, when a difference in plant height was first observed, were grown during the winter months. These plants, grown between the months of November and March, would have been subjected to reduced light intensity, reduced day length and reduced temperatures. In contrast, the comparative phenotype experiment was sown in May 2018 and harvested in August 2018. An increase in light intensity and light duration would be expected to influence the activity of PIFs. In the light, phytochrome will predominantly exist in its active Pfr form due to a high R/FR light ratio. Phytochrome in its Pfr form is able to translocate to the nucleus and interact with PIFs causing subsequent degradation of the PIF (Pham et al, 2018). The level of PIF protein and PIF activity would therefore be expected to remain minimal in control plants grown under high light conditions. This may provide an explanation as to why the plant height phenotype of the WT segregant control mimics that of the triple mutant. Furthermore, during the growth period of the phenotype experiment, the UK witnessed one of the hottest summers on record (Met

Office, 2018). Temperatures inside the glasshouse reached 34°C, which may have affected the growth of these plants. Therefore repeating this experiment under controlled environmental conditions would be important in confirming a role of TaPIF3 in the regulation of stem elongation.

The triple mutant however did demonstrate a significant reduced height phenotype when compared to Cadenza wild-type (figure 4.10). Total tiller length was significantly reduced when compared to wild-type. A reduction in total tiller length was due to a significant reduction in the emergence above the flag leaf, peduncle length and lengths of internodes 2 and 3. These results provide evidence that *TaPIF3* does have a role in regulating stem elongation in wheat.

4.3.3 Knockout of *TaPIF3* causes changes in days to anthesis and ear length

The *TaPIF3* triple mutant had significantly fewer days to anthesis than the control WT segregant line (figure 4.9), suggesting that *TaPIF3* may have a role in regulating the developmental timing of flowering. However, a difference in days to anthesis was not observed between the triple mutant and Cadenza wild-type. Furthermore, figure 4.9 shows that the WT segregant line had the greatest number of days to anthesis out of the six genotypes. It is therefore possible that a delay in flowering is due to non-target mutations in this line. A further phenotypic effect of manipulating TaPIF3 levels was an alteration in ear length. The *TaPIF3* triple mutant displayed a significantly shorter ear length than the WT segregant. A difference in ear length could be attributed to an alteration in rachis internode elongation, which occurs during spike development. The rachis is the main axis of the ear and has a single spikelet attached at each node. Previously, rachis internode elongation has been shown to be regulated by GA in barley (Nicholls, 1978). This result implicates a possible role of *TaPIF3* in the regulation of ear length.

4.3.4 Overexpression of *TaPIF3*

Overexpression was used as a second approach to characterise the function of *TaPIF3* in wheat. Independent overexpression lines were generated and T1 lines were grown at the

University of Southampton. Of interest, overexpression lines demonstrated a tall mature plant phenotype compared to controls. This phenotype is opposite to the phenotype first observed in the TILLING knockout mutant. Two lines specifically, B3602 R6P1 and B3602 R7P1, displayed an exaggerated difference in height compared to the control (figure 4.8). Due to this phenotype, overexpression lines B3602 R6P1 and B3602 R7P1 were included in the comparative phenotype experiment described in section 4.2.6. Line B3602 R2P3 that was genotyped to be positive for the *bar* selection marker gene but negative for the *TaPIF3* transgene was included as a control.

Overexpression line R6P1 demonstrated a small increase in tiller length compared to the control line R2P3. The difference in plant height between these two lines, photographed in figure 4.15, however was not significant at the 5% level of significance. Overexpression line R7P1 showed no increase in plant height compared to the control. These findings fail to replicate an increase in stem elongation observed previously in these lines. As mentioned, a difference in growing conditions within the glasshouse due to time of year may account for the altered phenotypes of these plants. In addition, the expression level of *TaPIF3* in these lines is unknown.

Despite overexpression lines R6P1 and R7P1 having no effect on total tiller length, both lines showed an increase in ear length compared to the control. The effect on ear length in these lines is opposite to the effect demonstrated by the TILLING knockout mutant. These findings therefore propose that TaPIF3 is likely to have a role in regulating ear length. Compared to Cadenza wild-type, overexpression line R6P1 showed a reduction in the emergence above the flag leaf. In addition, both lines R6P1 and R7P1 displayed a reduction in peduncle length compared to wild-type.

To further investigate the role of TaPIF3 in regulating growth and development in wheat, described overexpression lines need to be analysed for copy number and expression level. This would identify suitable lines to perform a comprehensive phenotypic study. Any observed phenotypes could then be directly associated with TaPIF3 function.

4.3.5 TaPIF3 does not interact with the wheat DELLA protein RHT-1 in yeast

One aim of this project was to elucidate whether the wheat DELLA protein RHT-1 interacts with a wheat PIF3 orthologue to regulate growth and development. So far, interactions between DELLAs and PIFs have only been demonstrated in Arabidopsis (de Lucas et al, 2008, Feng et al, 2008). In this chapter, TaPIF3 has been implicated as having a role in promoting stem elongation in wheat. To investigate whether TaPIF3 is a target for controlling GA-regulated responses including stem elongation, the ability of TaPIF3 to interact with RHT-1 was examined.

The yeast two-hybrid approach used in this study indicates that TaPIF3 does not interact with RHT-D1A. The growth of the yeast strain testing an interaction between RHT-D1A and TaPIF3 was equivalent to the RHT-D1A self-activation control and also the empty vector controls (figure 4.16). In this assay, a C-terminal fragment of RHT-1 was used to test an interaction with TaPIF3. This C-terminal truncated protein lacks the N-terminal DELLA/TVHYNP motif which has been implicated in conferring transactivation activity (Hirano et al, 2012). The truncated RHT-D1A bait plasmid used in this study however still shows some level of self-activation up to 50mM 3AT. This suggests that this C-terminal fragment possesses some transactivation activity.

TaPIF3 was also shown to self-activate in a yeast two-hybrid assay (figure 4.16). TaPIF3 caused a high level of self-activation even up to 100 mM 3-AT presumably because of its activity as a transcription factor. It was therefore not possible to repeat the assay using TaPIF3 as bait and RHT-1 as prey. To confirm that TaPIF3 does not interact with RHT-1, ideally an alternative method such as a pull-down assay would be used. Alternatively, an interaction *in vivo* could be tested using Bimolecular Florescence Complementation or immunoprecipitations. This preliminary finding however suggests that TaPIF3 does not interact with RHT-1 although further studies are required before any firm conclusion can be drawn.

Chapter 5. Heterologous overexpression of *TaPIF3* in *Arabidopsis thaliana*

5.1 Introduction

The results described in Chapter 3 demonstrate that TaPIF3 shares a highly conserved bHLH domain with Arabidopsis PIFs. In addition, TaPIF3 contains a conserved N-terminal APB motif that has been shown to be crucial for PIF function in Arabidopsis. These findings indicate potential conservation in physiological function between TaPIF3 and members of the Arabidopsis PIF subfamily. Investigation of *TaPIF3* function in wheat has been addressed in Chapter 4. However, due to the complexity and time-scale of generating knockout mutants or overexpression lines in wheat, further characterization of these lines is needed to determine gene function. Meanwhile, an alternative approach is the characterization of Arabidopsis transgenic lines, in which *TaPIF3* is over-expressed ectopically.

In contrast to many plants species, transformation of Arabidopsis to generate transgenic lines can be performed with relative ease. Transformation of Arabidopsis can be achieved using the 'floral dip' method, where an appropriate strain of *Agrobacterium tumefaciens* is applied to flowering Arabidopsis plants. Due to the physiology of Arabidopsis and some related Brassicaceae, a sufficient population of *Agrobacterium* can colonize the interior of developing ovaries during the early stages of floral development. This is where female gametophyte cell lineages are transformed (Bent, 2006). Of importance is that, this method requires minimal labour and is being used in laboratories world-wide. Perhaps the most significant application of this transformation method was in the generation of T-DNA insertion mutants. This means researchers can now obtain a gene knockout line from stock centres for most Arabidopsis genes (Alonso et al, 2003).

Transformation of Arabidopsis via the floral dip method will be used to investigate *TaPIF3* function. Specifically, *TaPIF3* will be transformed into a range of *pif* mutant backgrounds and the ability of *TaPIF3* expression to complement known *pif* mutant phenotypes will be examined. For instance, the *pifq* mutant exhibits a constitutive photomorphogenic phenotype in the dark that is characterised by short hypocotyls and open cotyledons (Leivar

et al, 2008, Shin et al, 2009). A heterologous complementation approach can therefore indicate whether TaPIF3 shares conservation of function to characterised PIFs in Arabidopsis.

5.1.1 Heterologous expression of rice PILs and maize PIFs in Arabidopsis

A heterologous expression approach has previously been used to characterise PIL and PIF sequences identified in the monocot species rice and maize. Nakamura et al (2007) identified a set of six highly conserved OsPIL sequences in rice. Five of these *OsPIL* genes were transformed into Arabidopsis under the control of the strong cauliflower mosaic virus (CaMV) 35S promoter. The ability of each OsPIL to regulate hypocotyl elongation in response to light was determined. This study concluded that OsPIL proteins have the functional ability to interfere with the light signal transduction pathway in a manner presumably homologous to that of PIF5 (Nakamura et al, 2007).

In an independent study, the physiological functions of two maize proteins, ZmPIF4 and ZmPIF5, were partially characterised through overexpression in Arabidopsis (Shi et al, 2018). The ability of ZmPIF4 and ZmPIF5 to complement the constitutive photomorphogenic phenotype of the *pifq* mutant in the dark was determined. Overexpression of *ZmPIF4* and *ZmPIF5* resulted in partial complementation of this phenotype with significantly elongated hypocotyls in transgenic lines relative to *pifq.* In addition, overexpression of *ZmPIF4* in Arabidopsis caused a constitutive shade avoidance phenotype further indicating that ZmPIF4 and ZmPIF5 have functions similar to those of PIF4 and PIF5 (Shi et al, 2018). A heterologous expression approach in Arabidopsis has therefore been successfully used to characterise the function of PIF sequences in monocot species.

5.2 Results

5.2.1 Production of *TaPIF3* overexpression lines

To express *TaPIF3* in Arabidopsis, the TaPIF3-5xMyc sequence, synthesised in section 3.2.2.5, was cloned into the Gateway binary vector pGWB502Ω (figure 5.1). The pGWB502Ω vector carries a modified CaMV 35S promoter in which the enhancer regions are duplicated and which further incorporates the Ω translational enhancer from tobacco mosaic virus. This promoter is therefore designed for strong, constitutive expression of protein-coding genes (Nakagawa et al, 2007). In addition, the pGWB502Ω vector contains the hygromycin phosphotransferase gene (HPT) conferring resistance to hygromycin. The hygromycin resistance marker allows selection of successfully transformed plants.

The Gateway recombination cloning system (Invitrogen, CA, USA) was used to clone the TaPIF3-5xMyc sequence into the pGWB502Ω expression vector. To facilitate cloning of TaPIF3-5xMyc into the required destination vector, TaPIF3-5xMyc was first cloned into the Gateway entry vector pENTR1A. TaPIF3-5xMyc was digested from the pUC57-Simple vector using restriction enzymes EcoRI and EcoRV. In addition, a digestion of pENTR1A was performed using enzymes EcoRI and EcoRV simultaneously. The products of each digest reaction were confirmed on an agarose gel and the digested bands were extracted and purified. The digested TaPIF3-5xMyc product was subsequently ligated into the digested pENTR1A vector in a T4 DNA ligase reaction. The ligation product was transformed into competent *E.coli* DH5α cells and plasmid DNA was extracted from positive transformants using a miniprep protocol. A diagnostic restriction digest and subsequent sequencing of plasmid DNA confirmed correct insertion of TaPIF3-5xMyc into pENTR1A.

Generation of an entry clone meant Gateway LR recombination technology could be used to transfer TaPIF3-5xMyc to the pGWB502Ω expression vector. The product of the LR recombination reaction was transformed into chemically competent DH5α cells. Plasmid DNA was extracted from transformed cells and restriction digests were used to confirm the presence of the insert. Sequencing of plasmid DNA confirmed that TaPIF3-5xMyc had been inserted in the correct reading frame for expression of the transgene as shown in the plasmid map in figure 5.1.

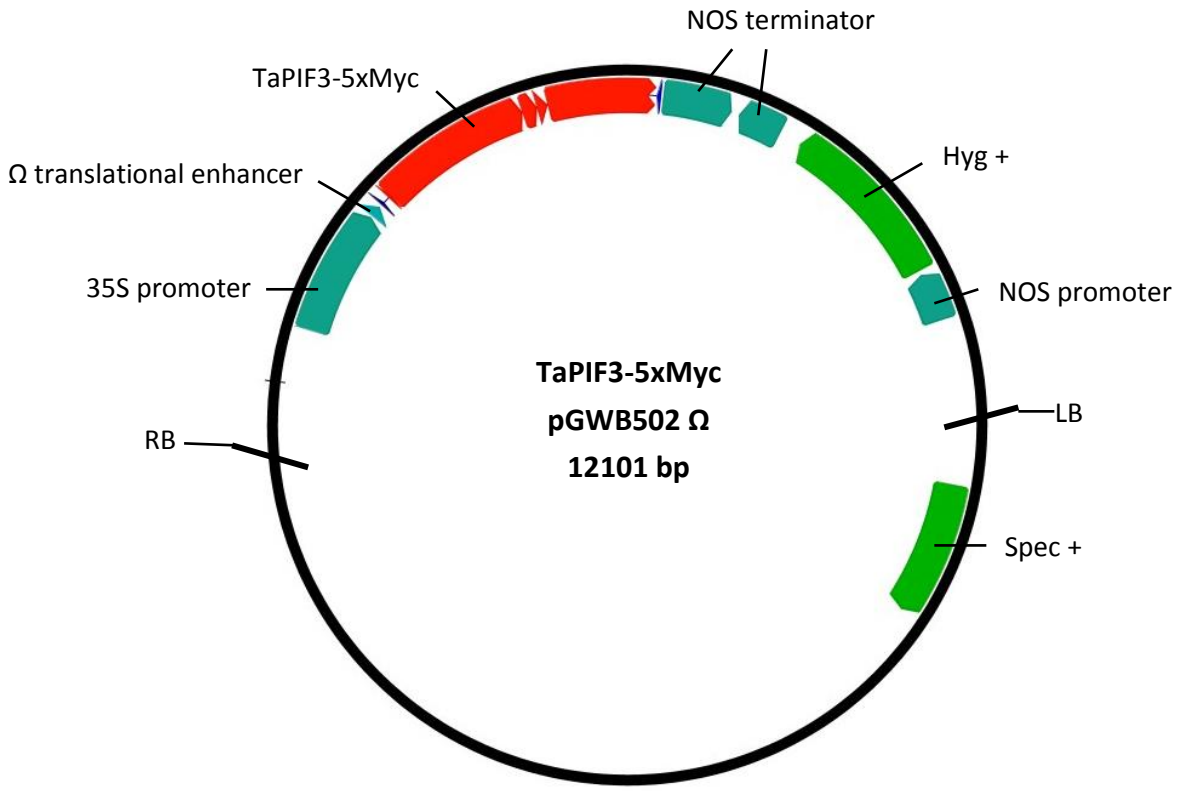


Figure 5.1 pGWB502Ω-TaPIF3 construct for *TaPIF3* overexpression in Arabidopsis. Genetic material between the left border (LB) and right border (RB) was transformed into Arabidopsis using *Agrobacterium*-mediated transformation. The TaPIF3-5xMyc sequence was cloned downstream of a modified 35S CaMV promoter with the Ω translational enhancer from tobacco mosaic virus for high level expression. The hygromycin (Hyg+) resistance gene allows selection of successfully transformed plants. The spectinomycin (Spec+) resistance gene, which is not transformed into the plant, allows selection of positive transformants in *E.coli* and *A.tumefaciens*.

The pGWB502Ω-TaPIF3 construct was transformed into *Agrobacterium tumefaciens* GV3101 cells. Positive transformants were selected on antibiotic selection plates. Two colonies that grew on selection plates were isolated and cultured for *Agrobacterium*-mediated transformation of Arabidopsis. The presence of the pGWB502Ω-TaPIF3 construct in these cells was confirmed in a colony PCR reaction (figure 5.2). A forward primer, termed pGWB_F, that anneals to the 35S promoter sequence of the T-DNA insert was designed. pGWB_F was used in combination with reverse primer TaPIF3_5'_R in a PCR reaction to

confirm the presence of the expression vector in *A. tumefaciens* (for primer sequences see table S.1). Transformation of Arabidopsis was subsequently performed using the floral dip method. Six plants of each of the following genotypes were transformed: wild-type (WT), *pif1*, *pif3*, *pif1pif3* and *pifq* (*pif1pif3pif4pif5*). Transformed plants were allowed to mature and produce seed.

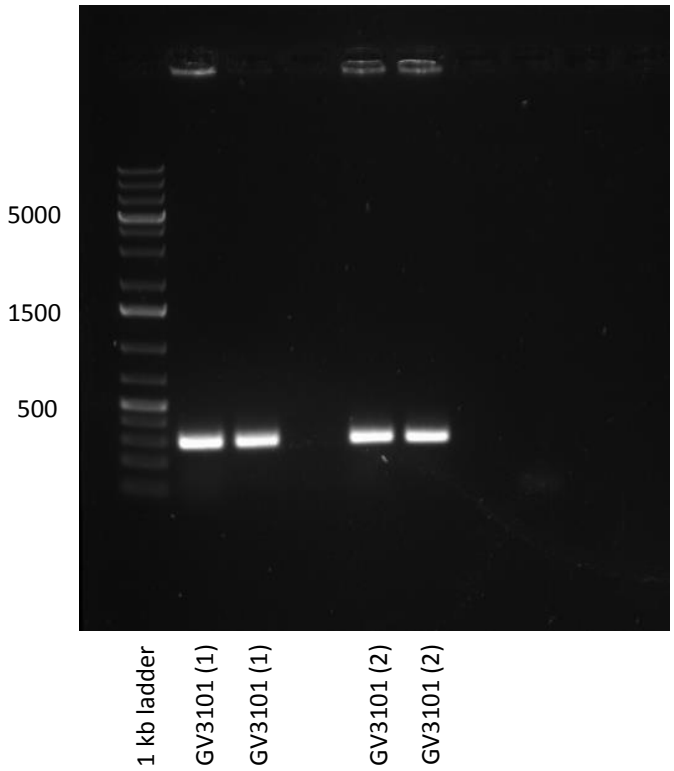


Figure 5.2 Confirmation of pGWB502Ω-TaPIF3 construct in GV3101 *A. tumefaciens* cells. The pGWB502Ω-TaPIF3 construct was transformed into *Agrobacterium tumefaciens* for *Agrobacterium*-mediated transformation of Arabidopsis.A 300 bp fragment of the pGWB502Ω-TaPIF3 construct was amplified using primers pGWB_F and TaPIF3_5'_R in a colony PCR reaction to confirm successful transformation of *Agrobacterium*. Two colonies, GV3101 (1) and GV3101 (2), that grew on antibiotic selection plates were used as template. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

The resulting T1 offspring were screened for positive transformants on selective media containing hygromycin antibiotic. Seeds sewn on selective media were treated to a three day cold treatment followed by a six hour WL treatment to induce germination. Seeds were

then grown in the dark for six days. Seedlings containing the hygromycin resistance gene were identified by elongation of the hypocotyl. From the population of positive transformants a total of twenty (T1) seedlings per background genotype were transferred to soil to self-fertilise. Surviving progeny from all six T0 plants were chosen to be transferred to soil. T1 lines were labelled according to the following nomenclature X:35STaPIF3-Y-Z where X refers to the background genotype, 35S describes the promoter of the transgene, Y refers to the T0 plant number and T1 denotes the T1 plant number.

To confirm the presence of the TaPIF3-5xMyc transgene in the T1 generation, genomic DNA was extracted from each plant transferred to soil for gene of interest PCR. Primers pGWB_F and TaPIF3_5'_R were selected to amplify a small region of the transgene and thereby confirm the presence of the T-DNA in T1 material (figures 5.3-5.7).

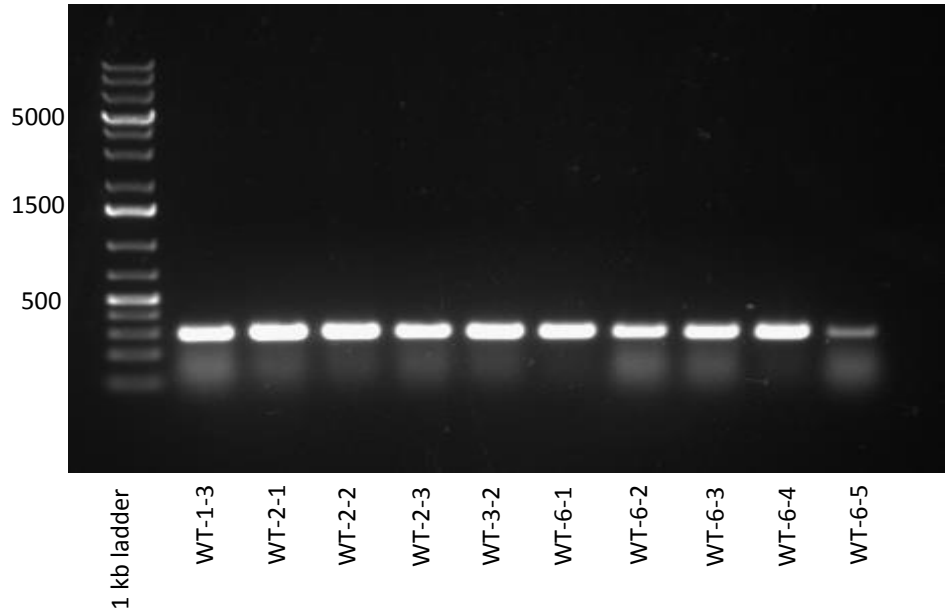


Figure 5.3 Amplification of *TaPIF3* transgene in WT Arabidopsis T1 overexpression lines. A 300 bp fragment of the *TaPIF3* transgene was amplified using primers pGWB_F and TaPIF3_5'_R. Genomic DNA was extracted from wild-type background T1 lines. Samples are labelled according to the background genotype, T0 plant number and T1 plant number. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

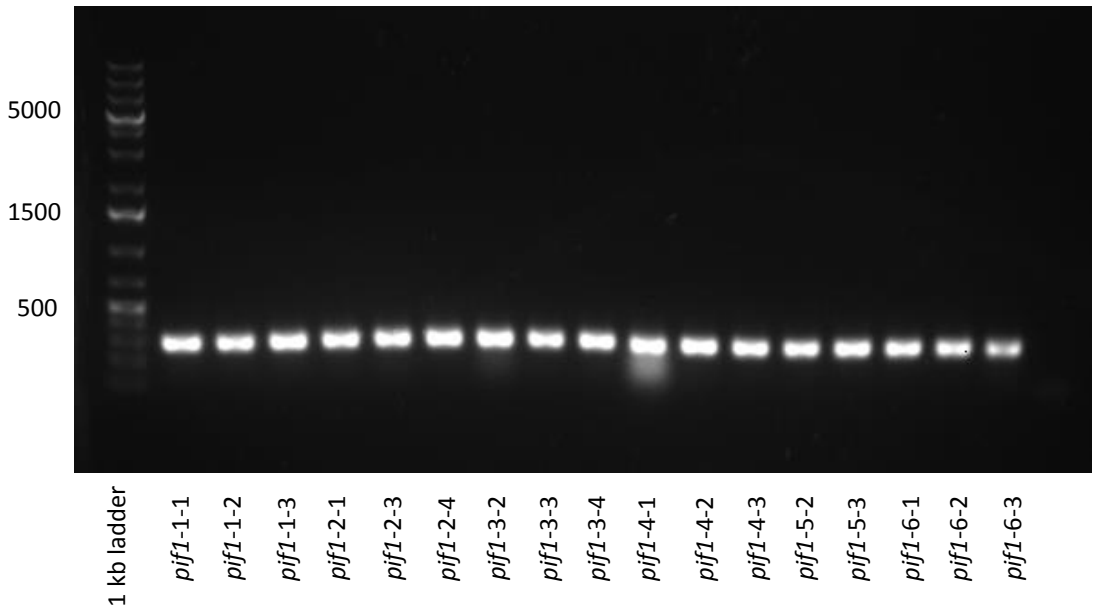


Figure 5.4 Amplification of *TaPIF3* transgene in Arabidopsis T1 overexpression lines in a *pif1* mutant background. A 300 bp fragment of the *TaPIF3* transgene was amplified using primers pGWB_F and TaPIF3_5'_R. Genomic DNA was extracted from *pif1* background T1 lines. Samples are labelled according to the background genotype, T0 plant number and T1 plant number. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

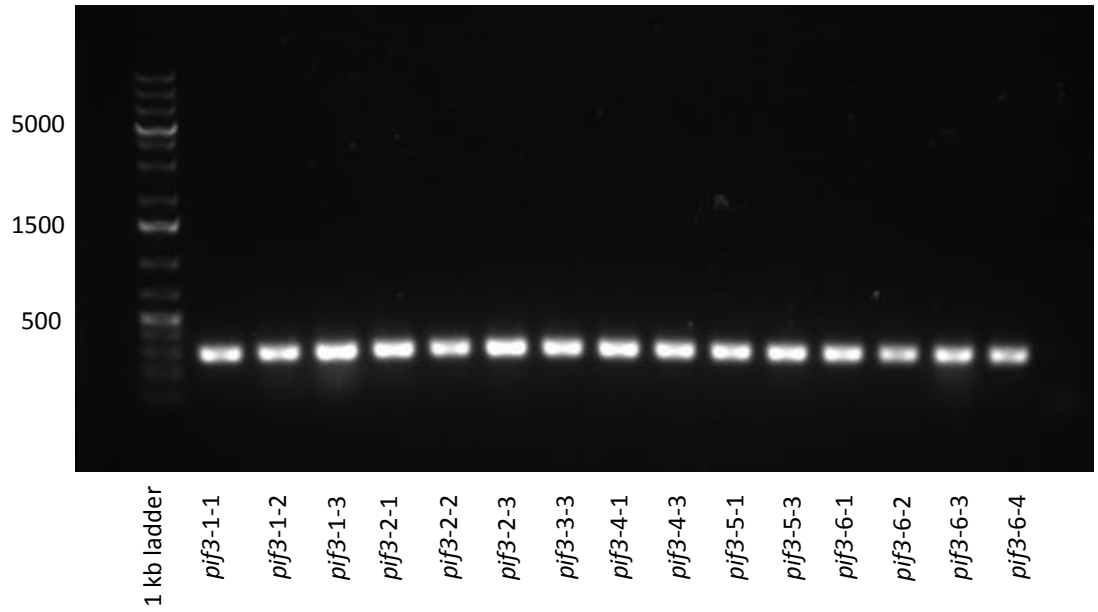


Figure 5.5 Amplification of *TaPIF3* transgene in Arabidopsis T1 overexpression lines in a *pif3* mutant background. A 300 bp fragment of the *TaPIF3* transgene was amplified using primers pGWB_F and TaPIF3_5'_R. Genomic DNA was extracted from *pif3* background T1 lines. Samples are labelled according to the background genotype, T0 plant number and T1 plant number. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

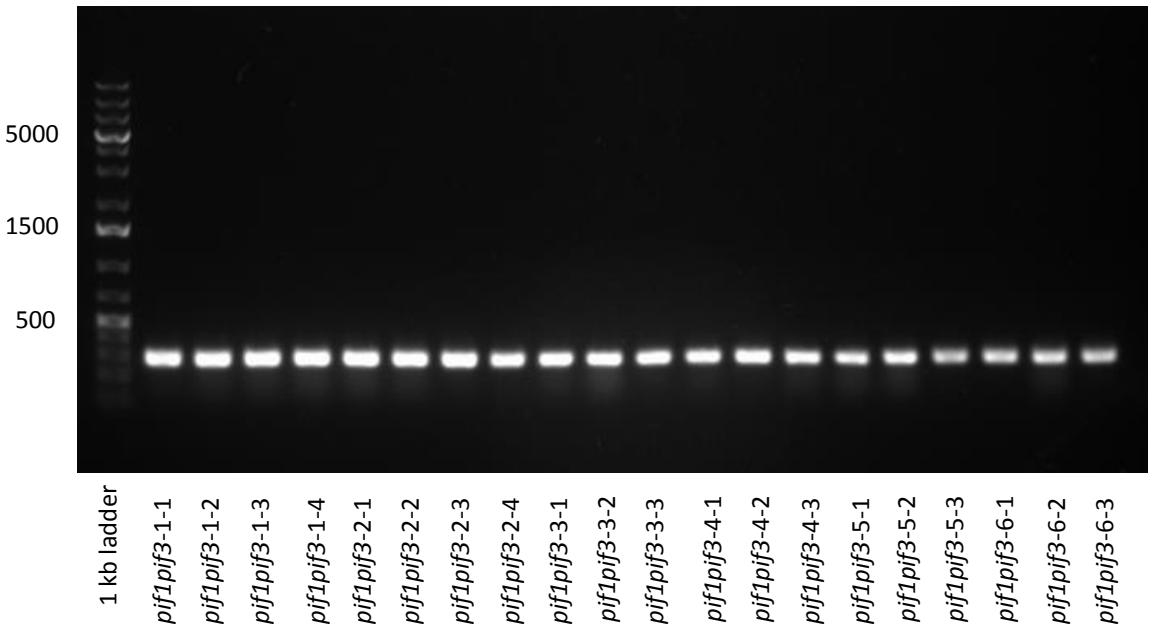


Figure 5.6 Amplification of *TaPIF3* transgene in Arabidopsis T1 overexpression lines in a *pif1pif3* mutant background. A 300 bp fragment of the *TaPIF3* transgene was amplified using primers pGWB_F and TaPIF3_5'_R. Genomic DNA was extracted from *pif1pif3* background T1 lines. Samples are labelled according to the background genotype, T0 plant number and T1 plant number. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

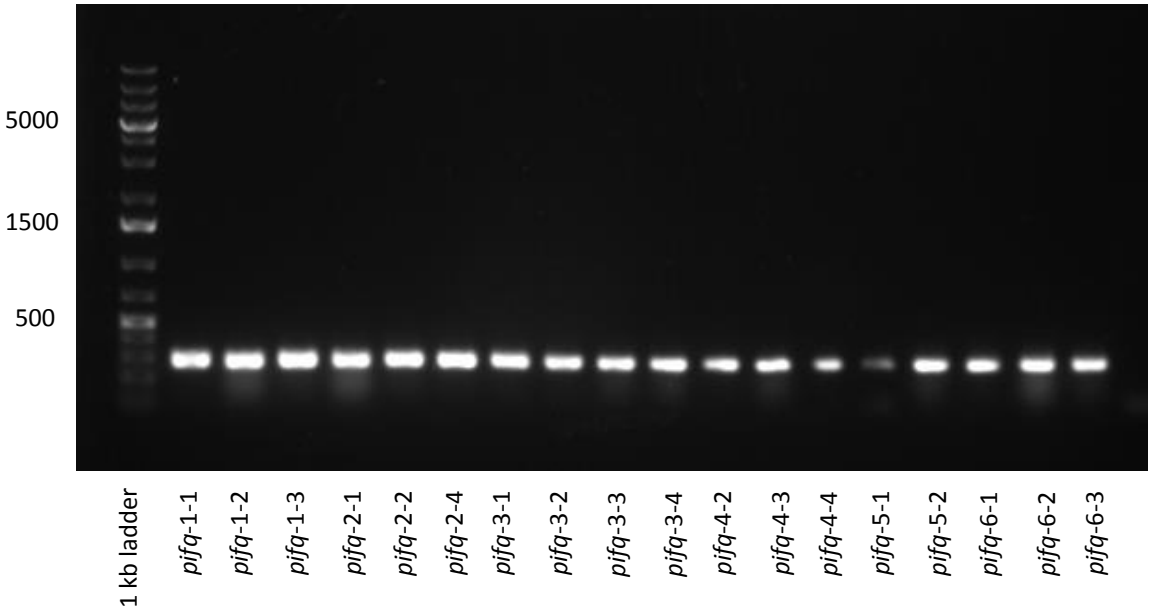


Figure 5.7 Amplification of *TaPIF3* transgene in Arabidopsis T1 overexpression lines in a *pifq* mutant background. A 300 bp fragment of the *TaPIF3* transgene was amplified using primers pGWB_F and TaPIF3_5'_R. Genomic DNA was extracted from *pifq* background T1 lines. Samples are labelled according to the background genotype, T0 plant number and T1 plant number. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

Amplification of a small fragment of the transgene indicated T1 plants were positive for the *TaPIF3* gene and the regulatory promoter sequence. T2 offspring was subsequently screened on selective media containing hygromycin. Lines that contain a single insertion could be identified with a 75% survival rate when exposed to the antibiotic. Initially eight independent lines per background genotype were screened with the aim of identifying six single insertion lines. The survival rates of T2 lines are shown in tables 5.1-5.5. Six seedlings per transgenic line selected as a single copy event were transferred to soil and allowed to self-fertilise. The single insertion lines transferred to soil for T3 seed are also shown in tables 5.1-5.5.

Table 5.1 Survival of T2 *TaPIF3* overexpressing lines in a WT background on hygromycin selective media. Survival rates and predicted insert number of *TaPIF3* overexpression lines in a wild-type background. Lines highlighted in bold were transferred to soil to produce T3 seed.

Plant line	Died	Survived	% Survived	Predicted Insert number
WT:35STaPIF3-1-3	32	89	73.5	1
WT:35STaPIF3-2-1	1	68	98.6	2+
WT:35STaPIF3-2-3	**24**	**68**	**73.9**	**1**
WT:35STaPIF3-6-1	**27**	**98**	**78.4**	**1**
WT:35STaPIF3-6-2	**37**	**89**	**70.6**	**1**
WT:35STaPIF3-6-3	**37**	**92**	**71.3**	**1**
WT:35STaPIF3-6-4	**32**	**86**	**72.8**	**1**
WT:35STaPIF3-6-5	**36**	**88**	**71.0**	**1**

Table 5.2 Survival of T2 *TaPIF3* overexpressing lines in a *pif1* background on hygromycin selective media. Survival rates and predicted insert number of *TaPIF3* overexpression lines in a *pif1* background. Lines highlighted in bold were transferred to soil to produce T3 seed.

Plant line	Died	Survived	% Survived	Predicted Insert number
***pif1*:35STaPIF3-1-1**	**35**	**89**	**71.8**	**1**
***pif1*:35STaPIF3-1-2**	**37**	**89**	**70.6**	**1**
pif1:35STaPIF3-2-1	10	118	92.0	2+
pif1:35STaPIF3-2-3	7	107	93.9	2+
***pif1*:35STaPIF3-3-4**	**37**	**84**	**69.4**	**1**
***pif1*:35STaPIF3-4-1**	**32**	**91**	**74.0**	**1**
***pif1*:35STaPIF3-5-3**	**27**	**99**	**78.6**	**1**
***pif1*:35STaPIF3-6-3**	**34**	**93**	**73.2**	**1**

Table 5.3 Survival of T2 *TaPIF3* overexpressing lines in a *pif3* background on hygromcyin selective media. Survival rates and predicted insert number of *TaPIF3* overexpression lines in a *pif3* background. Lines highlighted in bold were transferred to soil to produce T3 seed.

Plant line	Died	Survived	% Survived	Predicted Insert number
***pif3*:35STaPIF3-1-1**	**29**	**97**	**76.9**	**1**
***pif3*:35STaPIF3-1-2**	**35**	**90**	**72.0**	**1**
***pif3*:35STaPIF3-2-1**	**31**	**95**	**75.4**	**1**
***pif3*:35STaPIF3-2-2**	**34**	**93**	**73.2**	**1**
***pif3*:35STaPIF3-3-3**	**27**	**93**	**77.5**	**1**
pif3:35STaPIF3-4-1	7	122	94.5	2+
pif3:35STaPIF3-5-1	3	125	97.6	2+
pif3:35STaPIF3-6-1	10	117	92.0	2+

Table 5.4 Survival of T2 *TaPIF3* overexpressing lines in a *pif1pif3* background on hygromycin selective media. Survival rates and predicted insert number of *TaPIF3* overexpression lines in a *pif1pif3* background. Lines highlighted in bold were transferred to soil to produce T3 seed.

Plant line	Died	Survived	% Survived	Predicted Insert number
***pif1pif3*:35STaPIF3-1-2**	**27**	**96**	**78.0**	**1**
pif1pif3:35STaPIF3-1-3	9	114	92.7	2+
pif1pif3:35STaPIF3-2-1	20	109	84.5	1/2
pif1pif3:35STaPIF3-2-2	19	104	84.6	1/2
pif1pif3:35STaPIF3-3-1	9	118	93.0	2+
***pif1pif3*:35STaPIF3-4-1**	**32**	**79**	**71.2**	**1**
pif1pif3:35STaPIF3-5-1	1	126	99.2	2+
pif1pif3:35STaPIF3-6-1	5	124	96.1	2+
pif1pif3:35STaPIF3-1-4	7	123	94.6	2+
pif1pif3:35STaPIF3-2-3	1	129	99.2	2+
***pif1pif3*:35STaPIF3-3-2**	**29**	**65**	**69.1**	**1**
***pif1pif3*:35STaPIF3-4-2**	**42**	**86**	**67.2**	**1**
pif1pif3:35STaPIF3-5-2	1	127	99.2	2+
***pif1pif3*:35STaPIF3-6-2**	**43**	**82**	**65.6**	**1**
pif1pif3:35STaPIF3-2-4	6	105	94.6	2+
pif1pif3:35STaPIF3-3-3	20	101	83.5	1/2

Table 5.5 Survival of T2 *TaPIF3* overexpressing lines in a *pifq* background on hygromycin selective media. Survival rates and predicted insert number of *TaPIF3* overexpression lines in a *pifq* background. Lines highlighted in bold were transferred to soil to produce T3 seed.

Plant line	Died	Survived	% Survived	Predicted Insert number
***pifq*:35STaPIF3-1-1**	**35**	**93**	**72.7**	**1**
pifq:35STaPIF3-1-2	23	102	81.6	1
***pifq*:35STaPIF3-2-1**	**30**	**94**	**75.8**	**1**
***pifq*:35STaPIF3-3-1**	**43**	**82**	**65.6**	**1**
***pifq*:35STaPIF3-3-2**	**30**	**82**	**73.2**	**1**
***pifq*:35STaPIF3-4-2**	**24**	**67**	**73.6**	**1**
pifq:35STaPIF3-4-3	8	109	93.2	2+
***pifq*:35STaPIF3-5-1**	**31**	**94**	**75.2**	**1**

To identify homozygote transgenic lines, T3 seed was sown on hygromycin selective media and the survival rate of seedlings was recorded. A survival rate of 100% indicated a single insertion homozygote line had been identified. In most cases, a homozygote derived from an independent transgenic line could be identified (see tables 5.6-5.9). When more than one homozygote was identified, a single line was chosen at random for transgene expression analysis. Three seedlings from each homozygous line selected were transferred to soil to bulk T4 seed. Due to the number of lines generated, T3 selection of *TaPIF3* overexpression lines in the *pif1pif3* mutant background could not be completed.

Table 5.6 Survival of T3 *TaPIF3* overexpressing lines in a wild-type background on hygromycin selection media. Lines highlighted in bold were transferred to soil to bulk T4 seed. The genotype of each line is indicated, Het = heterozygous, Hom = homozygous, DNS = did not survive.

Plant line	Died	Survived	% Survived	Result
WT:35sTaPIF3-2-3-1	0	63	100	Hom
WT:35sTaPIF3-2-3-2	1	62	98.4	Hom
WT:35sTaPIF3-2-3-3	17	47	73.4	Het
WT:35sTaPIF3-2-3-4	0	64	100	Hom
WT:35sTaPIF3-2-3-5	16	47	74.6	Het
WT:35sTaPIF3-2-3-6	**0**	**64**	**100**	**Hom**
WT:35sTaPIF3-6-1-1	17	42	71.2	Het
WT:35sTaPIF3-6-1-2	14	47	77	Het
WT:35sTaPIF3-6-1-3	18	42	70	Het
WT:35sTaPIF3-6-1-4	**0**	**59**	**100**	**Hom**
WT:35sTaPIF3-6-1-5	23	41	64.1	Het
WT:35sTaPIF3-6-1-6	19	42	68.9	Het
WT:35sTaPIF3-6-2-1	11	46	80.7	Het
WT:35sTaPIF3-6-2-2	**0**	**61**	**100**	**Hom**
WT:35sTaPIF3-6-2-3	14	49	77.8	Het
WT:35sTaPIF3-6-2-4	18	46	71.9	Het
WT:35sTaPIF3-6-2-5	16	47	74.6	Het
WT:35sTaPIF3-6-2-6	14	50	78.1	Het
WT:35sTaPIF3-6-3-1	**0**	**64**	**100**	**Hom**
WT:35sTaPIF3-6-3-2	1	63	98.4	Hom
WT:35sTaPIF3-6-3-3	15	46	75.4	Het
WT:35sTaPIF3-6-3-4	0	64	100	Hom
WT:35sTaPIF3-6-3-5	8	56	87.5	Het
WT:35sTaPIF3-6-3-6	18	46	71.9	Het
WT:35sTaPIF3-6-4-1	0	62	100	Hom
WT:35sTaPIF3-6-4-2	19	44	69.8	Het
WT:35sTaPIF3-6-4-3	**0**	**64**	**100**	**Hom**
WT:35sTaPIF3-6-4-4	2	62	96.9	Hom
WT:35sTaPIF3-6-4-5	12	50	80.6	Het
WT:35sTaPIF3-6-4-6	20	43	68.3	Het
WT:35sTaPIF3-6-5-1	**0**	**64**	**100**	**Hom**
WT:35sTaPIF3-6-5-2	0	63	100	Hom
WT:35sTaPIF3-6-5-3	12	48	80	Het
WT:35sTaPIF3-6-5-4	0	61	100	Hom
WT:35sTaPIF3-6-5-5	16	44	73.3	Het
WT:35sTaPIF3-6-5-6	17	46	73	Het

Table 5.7 Survival of T3 *TaPIF3* overexpressing lines in a *pif1* background on hygromycin selection media. Lines highlighted in bold were transferred to soil to bulk T4 seed. The genotype of each line is indicated, Het = heterozygous, Hom = homozygous, DNS = did not survive.

Plant line	Died	Survived	% Survived	Result
pif1:35sTaPIF3-1-1-1	17	47	73.4	Het
pif1:35sTaPIF3-1-1-2	23	40	63.5	Het
***pif1*:35sTaPIF3-1-1-3**	**0**	**64**	**100**	**Hom**
pif1:35sTaPIF3-1-1-4	19	44	70	Het
pif1:35sTaPIF3-1-1-5	0	64	100	Hom
pif1:35sTaPIF3-1-1-6	22	41	65.1	Het
pif1:35sTaPIF3-1-2-1	15	47	75.8	Het
***pif1*:35sTaPIF3-1-2-2**	**0**	**63**	**100**	**Hom**
pif1:35sTaPIF3-1-2-3	1	63	98.4	Hom
pif1:35sTaPIF3-1-2-4	DNS	DNS	-	-
pif1:35sTaPIF3-1-2-5	3	61	95.3	Hom/Het
pif1:35sTaPIF3-1-2-6	DNS	DNS	-	-
***pif1*:35sTaPIF3-3-4-1**	**0**	**63**	**100**	**Hom**
pif1:35sTaPIF3-3-4-2	0	64	100	Hom
pif1:35sTaPIF3-3-4-3	0	64	100	Hom
pif1:35sTaPIF3-3-4-4	1	63	98.4	Hom
pif1:35sTaPIF3-3-4-5	16	47	74.6	Het
pif1:35sTaPIF3-3-4-6	0	64	100	Hom
pif1:35sTaPIF3-4-1-1	13	50	79.4	Het
pif1:35sTaPIF3-4-1-2	11	50	82	Het
pif1:35sTaPIF3-4-1-3	8	55	87.3	Het
pif1:35sTaPIF3-4-1-4	13	51	79.7	Het
***pif1*:35sTaPIF3-4-1-5**	**0**	**64**	**100**	**Hom**
pif1:35sTaPIF3-4-1-6	16	46	74.2	Het
***pif1*:35sTaPIF3-5-3-1**	**0**	**64**	**100**	**Hom**
pif1:35sTaPIF3-5-3-2	12	51	81	Het
pif1:35sTaPIF3-5-3-3	0	62	100	Hom
pif1:35sTaPIF3-5-3-4	17	45	72.6	Het
pif1:35sTaPIF3-5-3-5	1	61	98.4	Hom
pif1:35sTaPIF3-5-3-6	17	47	73.4	Het
pif1:35sTaPIF3-6-3-1	21	39	65	Het
pif1:35sTaPIF3-6-3-2	DNS	DNS	-	-
pif1:35sTaPIF3-6-3-3	2	62	96.9	Hom
***pif1*:35sTaPIF3-6-3-4**	**0**	**63**	**100**	**Hom**
pif1:35sTaPIF3-6-3-5	20	41	67.2	Het
pif1:35sTaPIF3-6-3-6	19	44	69.8	Het

Table 5.8 Survival of T3 *TaPIF3* overexpressing lines in a *pif3* background on hygromycin selection media. Lines highlighted in bold were transferred to soil to bulk T4 seed. The genotype of each line is indicated, Het = heterozygous, Hom = homozygous, DNS = did not survive.

Plant line	Died	Survived	% Survived	Result
pif3:35sTaPIF3-1-1-1	0	64	100	Hom
pif3:35sTaPIF3-1-1-2	15	48	76.2	Het
pif3:35sTaPIF3-1-1-3	16	47	74.6	Het
***pif3*:35sTaPIF3-1-1-4**	**0**	**63**	**100**	**Hom**
pif3:35sTaPIF3-1-1-5	16	48	75	Het
pif3:35sTaPIF3-1-1-6	12	51	80.9	Het
pif3:35sTaPIF3-1-2-1	DNS	DNS	-	-
pif3:35sTaPIF3-1-2-2	0	64	100	Hom
pif3:35sTaPIF3-1-2-3	19	45	70.3	Het
pif3:35sTaPIF3-1-2-4	15	46	75.4	Het
pif3:35sTaPIF3-1-2-5	16	46	74.2	Het
***pif3*:35sTaPIF3-1-2-6**	**0**	**64**	**100**	**Hom**
pif3:35sTaPIF3-2-1-1	20	43	68.3	Het
pif3:35sTaPIF3-2-1-2	1	62	98.4	Hom
pif3:35sTaPIF3-2-1-3	16	47	74.6	Het
***pif3*:35sTaPIF3-2-1-4**	**0**	**64**	**100**	**Hom**
pif3:35sTaPIF3-2-1-5	0	64	100	Hom
pif3:35sTaPIF3-2-1-6	0	64	100	Hom
pif3:35sTaPIF3-2-2-1	15	46	75.4	Het
pif3:35sTaPIF3-2-2-2	11	51	82.3	Het
pif3:35sTaPIF3-2-2-3	16	47	74.6	Het
***pif3*:35sTaPIF3-2-2-4**	**0**	**64**	**100**	**Hom**
pif3:35sTaPIF3-2-2-5	19	42	68.9	Het
pif3:35sTaPIF3-2-2-6	0	62	100	Hom
pif3:35sTaPIF3-3-3-1	0	63	100	Hom
pif3:35sTaPIF3-3-3-2	1	63	98.4	Hom
pif3:35sTaPIF3-3-3-3	11	53	82.8	Het
pif3:35sTaPIF3-3-3-4	2	59	97.6	Hom
pif3:35sTaPIF3-3-3-5	5	58	92.1	Hom/Het
***pif3*:35sTaPIF3-3-3-6**	**0**	**64**	**100**	**Hom**

Table 5.9 Survival of T3 *TaPIF3* overexpressing lines in a *pifq* background on hygromycin selection media. Lines highlighted in bold were transferred to soil to bulk T4 seed. The genotype of each line is indicated, Het = heterozygous, Hom = homozygous, DNS = did not survive.

Plant line	Died	Survived	% Survived	Result
pifq:35sTaPIF3-1-1-1	2	61	96.8	Hom
pifq:35sTaPIF3-1-1-2	20	44	68.8	Het
***pifq*:35sTaPIF3-1-1-3**	**0**	**64**	**100**	**Hom**
pifq:35sTaPIF3-1-1-4	0	62	100	Hom
pifq:35sTaPIF3-1-1-5	0	61	100	Hom
pifq:35sTaPIF3-1-1-6	DNS	DNS	-	-
pifq:35sTaPIF3-2-1-1	1	63	98.4	Hom
pifq:35sTaPIF3-2-1-2	1	60	98.4	Hom
pifq:35sTaPIF3-2-1-3	1	63	98.4	Hom
pifq:35sTaPIF3-2-1-4	18	46	71.9	Het
***pifq*:35sTaPIF3-2-1-5**	**0**	**64**	**100**	**Hom**
pifq:35sTaPIF3-2-1-6	3	60	95.2	Hom/Het
pifq:35sTaPIF3-3-1-1	22	42	65.6	Het
***pifq*:35sTaPIF3-3-1-2**	**0**	**64**	**100**	**Hom**
pifq:35sTaPIF3-3-1-3	19	45	70.3	Het
pifq:35sTaPIF3-3-1-4	2	60	96.8	Hom
pifq:35sTaPIF3-3-1-5	13	50	79.4	Het
pifq:35sTaPIF3-3-1-6	19	44	69.8	Het
pifq:35sTaPIF3-3-2-1	17	46	73	Het
pifq:35sTaPIF3-3-2-2	0	57	100	Hom
pifq:35sTaPIF3-3-2-3	1	56	98.2	Hom
pifq:35sTaPIF3-3-2-4	0	64	100	Hom
***pifq*:35sTaPIF3-3-2-5**	**0**	**64**	**100**	**Hom**
pifq:35sTaPIF3-3-2-6	11	62	82.3	Het
pifq:35sTaPIF3-4-2-1	11	52	82.5	Het
pifq:35sTaPIF3-4-2-2	17	46	73	Het
***pifq*:35sTaPIF3-4-2-3**	**0**	**63**	**100**	**Hom**
pifq:35sTaPIF3-4-2-4	DNS	DNS	-	-
pifq:35sTaPIF3-4-2-5	13	48	78.7	Het
pifq:35sTaPIF3-4-2-6	DNS	DNS	-	-
pifq:35sTaPIF3-5-1-1	11	51	82.3	Het
pifq:35sTaPIF3-5-1-2	13	50	79.4	Het
pifq:35sTaPIF3-5-1-3	19	45	70	Het
pifq:35sTaPIF3-5-1-4	21	43	67.2	Het
pifq:35sTaPIF3-5-1-5	19	44	69.8	Het
pifq:35sTaPIF3-5-1-6	11	49	81.7	Het

5.2.2 Confirmation of *TaPIF3* expression in Arabidopsis transgenic lines

Leaf material was harvested from each homozygote line transferred to soil and used for RNA extraction and subsequent cDNA synthesis. To confirm expression of the transgene in these plants, the presence of the *TaPIF3* transcript was tested by PCR using a cDNA template. Due to previous issues amplifying the *TaPIF3* sequence, a shorter fragment of the *TaPIF3* gene was chosen to confirm expression.

PCR was performed using primers TaPIF3_E4_F and TaPIF3_R1 (table S.1) designed to amplify a 485 bp product of *TaPIF3*. Initially, *TaPIF3* overexpression lines in a wild-type and *pifq* mutant background were analysed in a Phusion PCR reaction. Gel electrophoresis confirmed a PCR product of expected size had been amplified (figure 5.8). However, a band of the same size was also present in a non-transformed wild-type control reaction. A positive result in the wild-type control indicated possible contamination in the PCR reaction. To this end, PCR was repeated using a new set of PCR reagents. Despite repeating the PCR, a band remained present in the wild-type control.

A second explanation for observing a positive result in the WT control is potential amplification of an endogenous Arabidopsis gene, possibly a member of the *PIF* family. Therefore, a second PCR was performed using primers TaPIF3_E2_F and TaPIF3_E4_R designed to amplify a 1055 bp fragment of *TaPIF3* (Data not shown). Despite testing a second combination of primers, a positive result was still observed in the WT control. A positive result was also detected when using a second independent WT cDNA sample therefore suggesting the DNA template was not a source of contamination. It is not clear why a band can be observed in the wild-type control. Contamination would likely cause a band in the no-template control, which in this case is not clearly observed. Due to amplification of a ~500 bp product in the WT control, the described PCR cannot be used as evidence of *TaPIF3* expression in the overexpression lines.

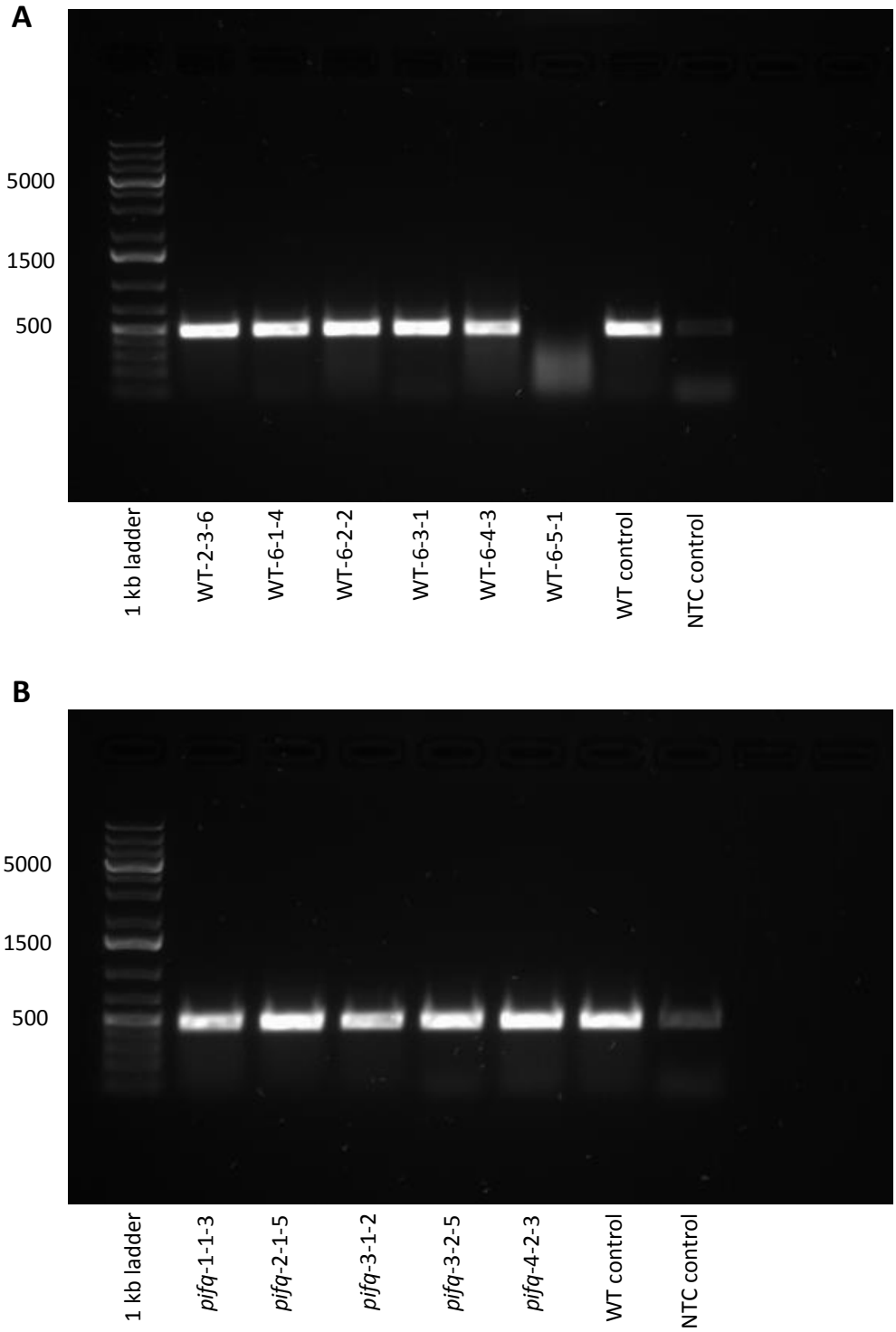


Figure 5.8 Confirmation of *TaPIF3* expression in T3 transgenic Arabidopsis overexpression lines. Confirmation of *TaPIF3* expression in T3 wild-type background **(A)** and *pifq* background **(B)** transgenic Arabidopsis plants. A 485 bp fragment of *TaPIF3* was amplified from individual cDNA samples. cDNA was synthesised from wild-type Arabidopsis as a control. A no-template control (NTC) was also used. Amplification of a ~500 bp product can be observed in the wild-type control. A 1 kb ladder is shown in lane 1. Band sizes are indicated in bp.

5.2.3 Overexpression of *TaPIF3* partially rescues the dark-grown phenotype of *pifq*

Expression of the *TaPIF3* transgene could not be confirmed in the described Arabidopsis overexpression lines due to a positive PCR result observed in the WT control (section 5.2.2). Despite this observation, the dark-grown phenotype of *TaPIF3* overexpression lines in the *pifq* mutant background was investigated. In Arabidopsis, the *pifq* mutant displays a constitutive photomorphogenic phenotype in the dark, including short hypocotyls and open cotyledons (Leivar et al, 2008, Shin et al, 2009). In addition, specific PIFs are known to be important repressors of chloroplast development (Stephenson et al 2009). As a result, the *pifq* mutant accumulates greater levels of the chlorophyll intermediate Pchlide in the dark relative to WT. Rescue of these phenotypes in *TaPIF3* overexpression lines would give insight into *TaPIF3* gene function and provide evidence that the *TaPIF3* transgene is expressed in these lines. Pchlide accumulation in the dark was therefore examined in WT, *pifq* and five independent T3 homozygote *TaPIF3* overexpression lines: *pifq*:35STaPIF3-1-1-3, *pifq*:35STaPIF3-2-1-5, *pifq*:35STaPIF3-3-1-2, *pifq*:35STaPIF3-3-2-5 and *pifq*:35STaPIF3-4-2-3 (figure 5.9). The data shows that the *pifq* mutant accumulated a significantly greater level of Pchlide compared to WT, consistent with previous findings (Stephenson et al, 2009). In addition, *TaPIF3* overexpression lines *pifq*:35STaPIF3-1-1-3, *pifq*:35STaPIF3-2-1-5, *pifq*:35STaPIF3-3-1-2 and *pifq*:35STaPIF3-3-2-5 accumulated significantly lower levels of Pchlide compared to *pifq*. Figure 5.9 clearly demonstrates that Pchlide accumulation is reduced to WT levels in *TaPIF3* overexpression lines *pifq*:35STaPIF3-1-1-3, *pifq*:35STaPIF3-2-1-5 and *pifq*:35STaPIF3-3-1-2. This finding suggests conservation in physiological function between *TaPIF3* and Arabidopsis *PIF1* and *PIF3*.

In addition, hypocotyl length of 5 d old dark-grown seedlings was quantified (figure 5.9). The hypocotyl length of *pifq* was significantly shorter than that of WT, consistent with the constitutive photomorphogenic phenotype of *pifq* under dark conditions. Hypocotyls of *TaPIF3* overexpression line *pifq*:35STaPIF3-2-1-5 were significantly elongated relative to *pifq* (figure 5.9). Despite this finding, there was no significant difference in hypocotyl length between *pifq* and the remaining four *TaPIF3* overexpression lines. The inability of *TaPIF3* to rescue the short hypocotyl phenotype of *pifq* suggests *TaPIF3* does not have a prominent role in promoting hypocotyl elongation as part of skotomorphogenic growth.

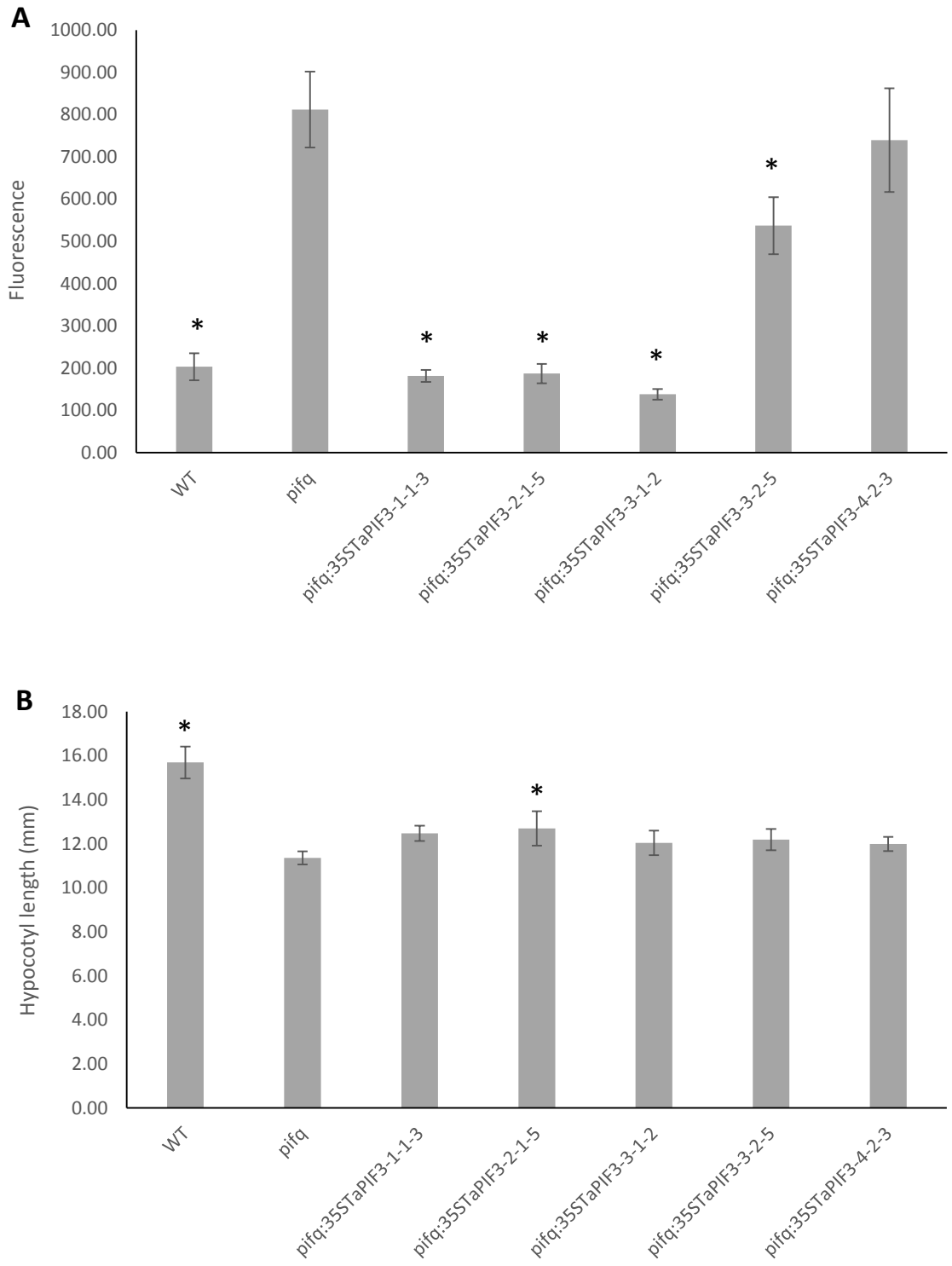


Figure 5.9 Overexpression of *TaPIF3* partially rescues dark-grown phenotype of the *pifq* mutant. (A) Protochlorophyllide accumulation in WT, *pifq*, and T3 *TaPIF3* overexpression lines. Protochlorophyllide was extracted from 5 d old dark-grown seedlings. (B) Hypocotyl growth of 5 d old dark-grown seedlings. Data presented in (A) and (B) are the mean ± 1 SD of 3 biological replicates. * indicates a significant difference ($p < 0.05$) when compared to *pifq* as calculated by a one-way ANOVA with a Dunnett's *post hoc* analysis.

5.3 Discussion

In Arabidopsis, the functional importance of PIFs has been characterised extensively through genetic and physiological studies. Since the founding member, PIF3, was first identified as a phytochrome-interacting factor in a yeast-two hybrid screen, a set of seven PIFs have been characterised in Arabidopsis (Ni et al, 1998). Characterisation of *pif* knockout mutants in Arabidopsis has revealed that PIFs function primarily as negative regulators of photomorphogenesis. Thus, the quadruple *pifq* (*pif1pif3pif4pif5*) mutant exhibits a constitutive photomorphogenic phenotype (Shin et al, 2009). However, despite their prominent role in regulating photomorphogenesis, PIFs have been shown to regulate many other pathways. Whereas the regulation of photomorphogenesis provides an example of overlapping functions of PIFs, certain PIFs function specifically to regulate a certain physiological response (Pham et al, 2018). For instance, PIF1 plays a major role in inhibiting light-dependent seed germination whereas PIF4 uniquely regulates hypocotyl elongation in response to light, temperature and diurnal conditions (Franklin et al, 2011, Lorrain et al, 2008, Oh et al, 2006). A summary of the specific functions of individual PIFs is provided in section 1.3.4.

The aim of this project is to identify and characterise the function of a *PIF3* orthologue in wheat. In Arabidopsis, PIF3 functions in coordination with other PIFs as a negative regulator of seedling deetiolation. In addition, PIF3 is an important regulator of chloroplast development and tetrapyrrole synthesis (Stephenson et al, 2009). A putative orthologue of *PIF3* in wheat, *TaPIF3,* was identified in chapter 3. TaPIF3 shares a highly conserved bHLH domain with Arabidopsis PIFs. In addition, TaPIF3 contains a conserved N-terminal motif that has previously been characterised as being necessary for interaction with phyB (Khanna et al, 2004). In a phylogenetic analysis of Arabidopsis, rice and wheat PIF/PIL sequences, TaPIF3 groups closely with PIF3 indicating the two genes are orthologous.

In contrast to Arabidopsis however, the role of PIFs in monocot species is not fully understood. Six putative PHYTOCHROME-INTERACTING FACTOR-LIKE (OsPIL11-OsPIL16) sequences have currently been identified in rice (Nakamura et al, 2007). All six rice PIL proteins contain a conserved APB motif, but only OsPIL15 contains an APA motif. In a similar manner to Arabidopsis PIFs, OsPILs have been show to regulate hypocotyl elongation in

transgenic Arabidopsis (Nakamura et al, 2007). In addition, overexpression of *OsPIL13* promotes internode elongation in rice (Todaka et al, 2012).

In maize, at least seven putative PHYTOCHROME-INTERACTING FACTORS (ZmPIFs) that contain a conserved APB motif have been identified (Kumar et al, 2016). An interaction between two ZmPIFs, ZmPIF3.1 and ZmPIF3.2, with the Pfr form of ZmphyB1 has been confirmed. However, neither ZmPIF3.1 or ZmPIF3.2 show an interaction with ZmphyB2 (Kumar et al, 2016). These results indicate that the phytochrome and PIF signalling system is somewhat conserved in maize. Furthermore, in a phylogenetic analysis, Shi et al (2018) identified putative orthologues of the Arabidopsis *PIF4* and *PIF5* genes. Expressing *ZmPIF4* and *ZmPIF5* ectopically in transgenic Arabidopsis plants suggested that ZmPIF4 and ZmPIF5 participate in photomorphogenesis and the shade avoidance response. Thereby implying a conserved role to PIF4 and PIF5 in the response to phytochrome signalling in plants (Shi et al, 2018).

Studies in rice and maize, therefore suggests some conservation of PIF function between Arabidopsis and monocot species. However, it remains uncertain as to whether TaPIF3 can be expected to show conservation of function to PIF3. To further investigate the function of TaPIF3, transgenic Arabidopsis lines that overexpress *TaPIF3* in a range of *pif* mutant backgrounds have been produced. To generate these lines, TaPIF3-5xMyc was cloned into the expression vector pGWB502Ω. The pGWB502Ω vector carries a modified CaMV 35S promoter for strong, constitutive expression of the transgene. The pGWB502Ω-TaPIF3 construct was subsequently transformed into *A. tumefaciens* for transformation of Arabidopsis. *Agrobacterium*-mediated transformation was performed on six plants of each of the following genotypes: WT, *pif1*, *pif3*, *pif1pif3* and *pifq.* Homozygote *TaPIF3* overexpressing lines were later established in a WT, *pif1*, *pif3* and *pifq* mutant background (tables 5.6-5.9).

To confirm that the transgene was expressed in homozygote lines, PCR amplification of *TaPIF3* was performed from individual T3 cDNA samples. A small amplicon of *TaPIF3* was selected due to previous issues amplifying large fragments of the *TaPIF3* gene. Homozygote overexpression lines in both WT and *pifq* backgrounds proved positive for the 485 bp amplicon. However, the WT control also proved positive for the transcript (figure 5.8). It remains uncertain as to why this band is present in the wild-type control. Likely explanations

include contamination in the PCR reaction or off-target amplification of an Arabidopsis gene. Expression of the *TaPIF3* transgene in these lines therefore remains to be confirmed. Additional attempts at confirming *TaPIF3* expression using a real-time PCR approach would have the benefit of providing valuable expression level data.

5.3.1 TaPIF3 represses protochlorophyllide synthesis in the dark

In Arabidopsis, members of the PIF family of transcription factors primarily act to repress photomorphogenesis in the dark. To investigate *TaPIF3* gene function, the dark-grown phenotype of *TaPIF3* overexpression lines in the *pifq* mutant background was examined. The *pifq* mutant lacks functional copies of *PIF1*, *PIF3*, *PIF4* and *PIF5* and as such demonstrates a constitutive photomorphogenic phenotype, including short hypocotyls and open cotyledons, when grown under dark conditions (Leivar et al, 2008, Shin et al, 2009). Furthermore, dark-grown *pifq* seedlings accumulate greater levels of the chlorophyll intermediate Pchlide, relative to WT. Data presented in this chapter demonstrates that expression of *TaPIF3* in the *pifq* mutant rescues the Pchlide phenotype (figure 5.9). Pchlide accumulation was restored to WT levels in three independent *TaPIF3* overexpression lines. This finding suggests that TaPIF3 acts to regulate the tetrapyrrole biosynthetic pathway in a similar manner to Arabidopsis PIF1 and PIF3. In Arabidopsis, PIF1 and PIF3 have been shown to regulate tetrapyrrole synthesis by repressing the expression of key chlorophyll synthesis genes (Stephenson et al, 2009). Chlorophyll biosynthetic genes under the regulation of PIF1 and PIF3 include *HEMA1* that encodes the rate-limiting enzyme glutamyl tRNA reductase. Following the expression of *HEMA1* in the described *TaPIF3* overexpression lines would be an important indicator to whether *TaPIF3* has a conserved function in regulating this process. Despite Pchlide levels being restored in three *TaPIF3* overexpression lines, one line *pifq*:35STaPIF3-4-2-3 accumulated a similar level of Pchlide to *pifq* (figure 5.9). Expression of *TaPIF3* in these lines are yet to be confirmed, it is therefore possible that *TaPIF3* is not expressed in this line or that it is expressed at a low level.

Ectopic expression of *TaPIF3* did not rescue the short hypocotyl phenotype of *pifq*. Figure 5.9 shows *TaPIF3* overexpression lines demonstrated a small increase in hypocotyl elongation in dark grown seedlings compared to *pifq*. However, this effect on hypocotyl

length was only significant in one of the *TaPIF3* overexpression lines. In Arabidopsis, PIFs function redundantly to repress photomorphogenic growth. Consequently, single mutants of *PIF1*, *PIF3*, *PIF4* and *PIF5* demonstrate a similar hypocotyl length to WT in the dark (Shin et al, 2009). Double and triple *PIF* mutants display only marginally shorter hypocotyls than WT. Therefore, if multiple PIF-like transcription factors regulate skotomorphogeneis in wheat, expression of *TaPIF3* alone may not be expected to fully rescue the short hypocotyl of *pifq*.

The transgenic lines generated in this chapter therefore provide a valuable tool for studying TaPIF3 function for ongoing studies. To investigate whether TaPIF3 shares conservation of PIF3 function, the ability of *TaPIF3* expression to complement PIF3 specific phenotypes could be examined. For example, PIF3 is known to be an important repressor of chloroplast development. Thus, *pif3* seedlings grown in the dark show partially developed chloroplasts. Etioplasts in *pif3* seedlings can be characterised by increased membrane development and PLBs that are reduced in size (Stephenson et al, 2009). Rescue of these phenotypes in *TaPIF3* overexpression lines would therefore indicate a degree of conserved function with PIF3.

Chapter 6. General Discussion

6.1 Identification of a *PIF3* orthologue in hexaploid wheat

The first aim of this research project was to identify Phytochrome Interacting Factor-like genes in wheat with a particular emphasis on identifying a putative *PIF3* orthologue. *PIF*3, the founding member of the *PIF* family in Arabidopsis was first identified in a yeast two-hybrid screen aimed at identifying primary phytochrome signalling partners (Ni et al, 1998). A family of seven PIF proteins, PIF1, PIF3, PIF4, PIF5, PIF6, PIF7 and PIF8 have since been identified in Arabidopsis that act redundantly to regulate light responses (Khanna et al, 2004). In addition to Arabidopsis, PIFs have been identified in a variety of plant lineages from bryophytes to angiosperms (Lee and Choi, 2017). In *Solanum lycopersicum* (tomato) a family of eight PIF proteins have been identified: *SlPIF1a, SlPIF1b, SlPIF3, SlPIF4, SlPIF7a, SlPIF7b, SlPIF8a, SlPIF8b,* although only SlPIF1a has been characterised and confirmed to function as a PIF (Llorente et al, 2016, Rosado et al, 2016). *PIF* genes have also been identified in the moss *Physcomitrella patens* (Possart et al, 2017). *P. patens* contains four putative *PIF* orthologues: Pp-PIF1, Pp-PIF2, PpPIF3 and PpPIF4. In contrast, the liverwort *Marchantia polymorpha* has only one *PIF* gene, *Mp-PIF,* which like Arabidopsis PIFs is able to regulate gene expression and is targeted for degradation by phytochrome in response to red light (Inoue et al, 2016). In monocot species, *PIF3*-like (PIL) genes have been identified in rice and maize. Rice has been shown to encode a family of six PIL proteins that are termed OsPIL11-16 (OsPIL13 also known as OsPIL1) (Nakamura et al, 2007). The findings of Nakamura et al (2007) indicated that there are three pairs of OsPILs among which OsPIL15 and OsPIL16 are homologous to PIF3 and OsPIL13 and OsPIL14 are homologous to PIF4 and PIF5.

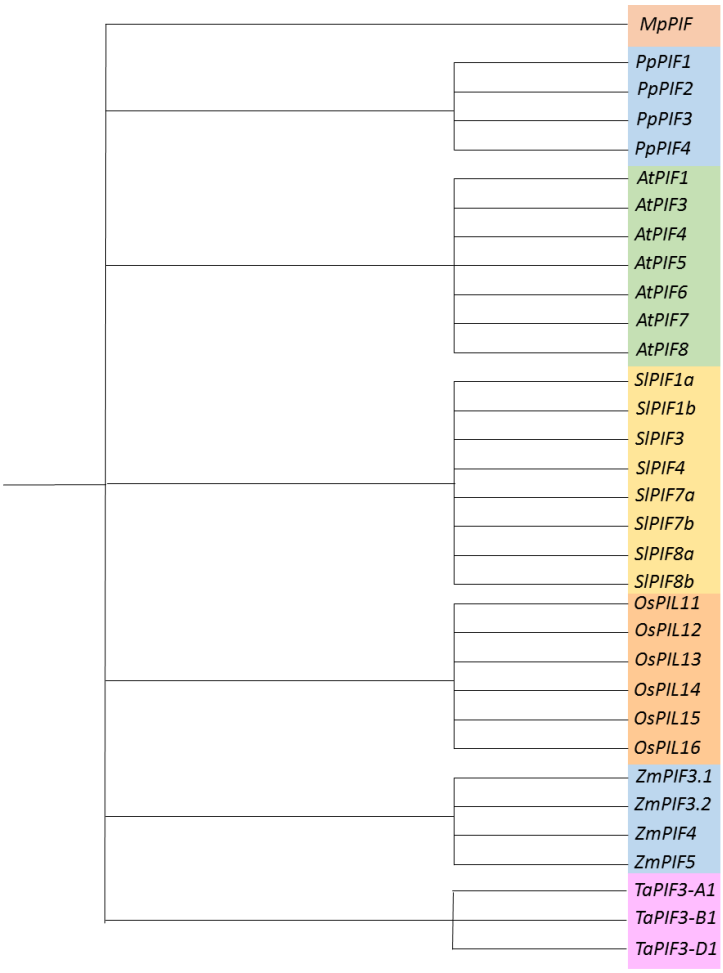


Figure 6.1 Schematic representation of *PIF* genes identified in plant species. Schematic phylogenetic tree of *PIF* genes identified in plant species. *PIFs* have currently been identified in *Marchantia polymorpha* (*Mp*), *Physcomitrella patens* (*Pp*), *Arabidopsis thaliana* (*At*), *Solanum lycopersicum* (*Sl*), *Oryza sativa* (*Os*), *Zea mays* (*Zm*) and *Triticum aestivum* (*Ta*).

The sequences of known Arabidopsis and rice *PIFs* were used to identify a putative *PIF3* orthologue in the bread wheat genome. Three homoeologous genes were identified termed *TaPIF3-A1*, *TaPIF3-B1* and *TaPIF3-D1.* Phylogenetic analysis of the Arabidopsis PIFs, rice PILs and identified wheat sequences further indicated that TaPIF3 was homologous to Arabidopsis PIF3 and to OsPIL15 and OsPIL16. In Arabidopsis, PIF3 functions predominately as a negative regulator of phytochrome-mediated seedling de-etiolation. As such, *pif3* mutant seedlings grown under continuous red light demonstrate a short hypocotyl and open cotyledons compared to WT (Kim et al, 2003, Leivar et al, 2008). The mechanism by

which PIF3 mediates the negative regulation of seedling de-etiolation is primarily by controlling phyB abundance (Leivar et al, 2008). PIF3 is also known to be an important repressor of chloroplast development in etiolated seedlings (Stephenson et al, 2009). For instance, PIF3 represses chlorophyll synthesis by regulating the expression of key chlorophyll synthesis genes (Stephenson et al, 2009). Consequently, dark grown *pif3* seedlings accumulate Pchlide to a greater level than WT. In addition, PIF3 interacts with the core clock component Timing of CAB expression 1 (TOC1) to optimise PIF-induced growth (Soy et al, 2016). Furthermore, PIF3 regulates ethylene-induced hypocotyl elongation in light and also modulates freezing tolerance by negatively regulating *CBF* genes (*C-REPEAT BINDING FACTOR*) genes (Jiang et al, 2017, Zhong et al, 2012). The TaPIF3 protein could therefore be hypothesized to regulate equivalent responses in wheat.

The findings of Nakamura et al (2007) indicated that rice encodes a family of six PIL proteins. In addition to identifying a putative PIF3 orthologue, additional PIF-like sequences were searched for in the wheat genome. To identify additional PIF-like genes in wheat, the amino acid sequences of Arabidopsis PIFs, OsPIL16 and TaPIF3 were entered into InterPro to identify characteristic features of the protein sequences. All sequences contained a Myc-type, basic helix-loop-helix domain as would be expected. In addition, the Arabidopsis, rice and wheat sequences were further classified into a STEROL REGULATORY ELEMENT-BINDING PROTEIN family (PTHR12565) according to the PANTHER Classification System. Genes encoding proteins containing the PTHR12565 signature were therefore extracted from the IWGSC CSS database and included in the phylogenetic analysis. Wheat sequences that closely grouped to other Arabidopsis PIFs such as PIF4 and PIF5, however, could not be identified from the analysis. It remains likely, that there are more wheat PIF-like sequences. In fact, three genes termed *TaPIL1*, *TaPIL2* and *TaPIL3* have recently been identified that show close homology to rice *OsPIL*s (Wallace, 2017). Specifically, a phylogenetic analysis revealed that TaPIL1 was homologous to OsPIL13 (OsPIL1), TaPIL2 was homologous to OsPIL11 and OsPIL12 and TaPIL3 was most homologous to OsPIL14. A comprehensive analysis of the new IWGSC RefSeq v1.0 assembly should be performed to search for additional PIF-like sequences in wheat.

PIF proteins in Arabidopsis contain a conserved C-terminal bHLH domain that provides dimerization (HLH domain) and DNA binding capacity (basic domain) (Toledo-Ortiz et al,

2003). In addition, all members of the Arabidopsis PIF family contain a conserved APB motif which facilitates interaction with phyB (Khanna et al, 2004). These two domains have also been identified in rice (Nakamura et al, 2007), tomato (Rosado et al, 2016), maize (Kumar et al, 2016) and *P. patens* (Possart et al, 2017) PIF protein sequences. To gain evidence that TaPIF3 functions as a PIF in wheat, these two domains were searched for in the predicted TaPIF3 protein sequence. Sequence alignments of the three homoeologues of TaPIF3 with Arabidopsis PIFs identified the presence of a conserved APB motif and a highly conserved bHLH domain. However, the identification of an APB motif does not necessarily confirm the ability of that protein to bind to phyB. For instance, an APB motif was identified in the amino acid sequence of OsPIL1, but an interaction between OsPIL1 and OsPhyB could not be detected in a yeast two-hybrid assay (Todaka et al, 2012). It would therefore be important to demonstrate an interaction between TaPIF3 and phytochrome experimentally.

6.2 Altering *TaPIF3* expression results in stem elongation phenotypes

The function of TaPIF3 in wheat has been investigated by exploiting a reverse genetics approach. A *TaPIF3* knockout mutant has been generated by crossing independent TILLING lines containing predicted loss-of-function mutations in each homoeologue of *TaPIF3*. In addition, *TaPIF3* overexpression lines have been generated. Initial comparisons between two *TaPIF3* homozygous triple mutants and four WT segregants indicated that knockout of *TaPIF3* leads to a decrease in plant height. In contrast, overexpression lines demonstrated a tall phenotype compared to controls. These findings indicate that TaPIF3 may be involved in the regulation of plant height. In rice, OsPIL1 has been implicated as a key regulator of internode elongation (Todaka et al, 2012). Todaka et al (2012) demonstrated that OsPIL1 overexpression lines show enhanced stem elongation, leading to a strikingly tall plant. An increase in internode elongation was established to be due to increased cell elongation rather than an enhanced cell division rate. This conclusion was supported by the finding that genes up-regulated in OsPIL1 overexpression lines include genes involved in cell wall development, organisation and biogenesis (Todaka et al, 2012).

Furthermore, in Arabidopsis, PIFs have also been shown to regulate stem elongation. PIF4, for example, regulates hypocotyl elongation in response to a range of environmental

stimuli. Under light limiting conditions, PIF4 promotes growth by causing increased expression of shade-induced genes such as *ATHB-2* and *PIL1* as part of the shade-avoidance response (Lorrain et al, 2008). PIF4 also promotes hypocotyl elongation in response to high temperature by regulating levels of auxin and the expression of key auxin biosynthesis genes (Franklin et al, 2011). PIF4 is therefore widely associated as having the most specific role in elongation responses amongst the Arabidopsis PIF family. However, PIF3 has been implicated in regulating ethylene-induced hypocotyl elongation in response to light (Zhong et al, 2012). Ethylene is a phytohormone that promotes hypocotyl elongation in the light through downstream signalling molecules ETHYLENE-INSENSITIVE 3 (EIN3) and EIN3-like 1 (EIL1). Zhong et al (2012) demonstrated that EIN3 activates *PIF3* gene expression to trigger a PIF3-dependent growth promotion pathway.

Observations that TaPIF3 may play a role in regulating elongation responses in wheat are therefore not surprising. To investigate the phenotype of plants with altered *TaPIF3* expression levels further, a comparative phenotype experiment was set up that included the *TaPIF3* triple mutant and *TaPIF3* overexpression lines with increased replication. Despite previous findings, the triple mutant did not show a reduced height phenotype compared to the WT segregant control. The triple mutant however did demonstrate a significant reduced plant height phenotype compared to Cadenza wild-type. A reduction in plant height was due to a significant reduction in the emergence above the flag leaf, peduncle length and lengths of internodes 2 and 3.

Analysis of *TaPIF3* overexpression lines also failed to replicate previous findings. Overexpression line B3602 R6P1 demonstrated a small increase in tiller length compared to the control line (figure 4.10). However, this increase in height was not significant at the 5% level of significance. In addition, overexpression line B3602 R7P1 showed no increase in plant height compared to the control. The phenotypes of the triple mutant and overexpression lines are therefore not consistent enough to associate *TaPIF3* as having a role in promoting stem elongation in wheat. The uncertain role of TaPIF3 in elongation responses could be attributed to a number of reasons. Firstly, variable environmental conditions experienced when growing these plants in the glasshouse could account for phenotypic differences. Triple mutant lines and WT segregants photographed in figure 4.6, when a difference in plant height was first observed, were grown during the winter months.

This is in contrast to the comparative phenotype experiment, which was conducted during the summer. Differences in light intensity, temperature and day length may be expected to influence the regulation and activity of PIF proteins affecting the phenotypes of these plants. Limitations of the described TILLING lines and overexpression lines could also account for not observing a consistent phenotype. For example, the *TaPIF3* triple mutant generated in this project has not been confirmed as a knockout mutant. The three mutant alleles of *TaPIF3* each have a predicted deleterious effect, however, the impact of each mutation needs to be confirmed *in vivo*. In addition, the *TaPIF3* triple mutant and the WT segregant lines would ideally have been backcrossed prior to this analysis to prevent off-target mutations causing or influencing a phenotype. Another possibility is that PIF proteins in wheat act redundantly to promote stem elongation. In this case, knockout of just one PIF may not be enough to observe a strong phenotype.

6.3 Alteration of *TaPIF3* expression levels indicate a role of TaPIF3 in the promotion of ear elongation

Results of the comparative phenotype experiment indicated a potential role of TaPIF3 in the regulation of ear length. The *TaPIF3* triple mutant displayed a significant decrease in ear length compared to the WT segregant control. In contrast, both overexpression lines B3602 R6P1 and B3602 R7P1 demonstrated a significant increase in ear length compared to controls. It is possible that altered ear length in these lines is due to changes to the elongation of the rachis internode, which then develops into the ear later in development. The effect on ear length could therefore be a by-product of TaPIF3 dependent promotion of internode elongation. It would be of interest to determine whether an increase in ear length in the overexpression lines results in an increase in number or weight of grain as this could provide an opportunity to increase wheat yields.

6.4 TaPIF3 does not interact with RHT-1 in a yeast two-hybrid assay

In Arabidopsis, PIF3 and PIF4 have been shown to interact with all five Arabidopsis DELLA proteins (Feng et al, 2008, de Lucas et al, 2008). Deletion studies have indicated that this interaction involves the bHLH DNA recognition domain of the PIFs and the leucine heptad repeat (LHR) region of the DELLA GRAS domain (de Lucas et al, 2008). Through this interaction, DELLAs have been shown to physically interact with PIF3 and PIF4 and sequester them from binding to their target promoters (Feng et al 2008, de Lucas et al, 2008). More recently, an interaction between DELLAs and PIFs has also been shown to promote the degradation of PIF proteins through the ubiquitin-proteasome system (Li et al, 2016). Interestingly, Li et al (2016) implicated DELLAs as negative regulators of hypocotyl elongation by causing degradation of PIF3. DELLAs have also been shown to protect against photo-oxidative stress by regulating chlorophyll and carotenoid biosynthesis, at least in part through the regulation of PIF activity (Cheminant et al, 2011).

These findings in Arabidopsis led to the hypothesis that the wheat DELLA protein RHT-1 interacts with PIF proteins in wheat to regulate growth and oxidative stress responses. The ability of RHT-1 to interact with TaPIF3 was therefore examined. A C-terminal truncated RHT-1 product that contains the LHR region was used as bait in a specific yeast two-hybrid assay. Results of the assay indicated that RHT-1 was not able to interact with TaPIF3. It remains possible that TaPIF3 requires a region of RHT-1 that was missing in the truncated bait product in order to form a physical interaction. In addition, an interaction between RHT-1 and TaPIF3 has only be tested using one method. Therefore, further studies are required before a firm conclusion can be drawn.

6.5 Conclusions and future research directions.

A putative *PIF3* orthologue, termed *TaPIF3*, has been identified in the wheat genome. The gene sequence of *TaPIF3* has been confirmed experimentally by amplification of *TaPIF3* genomic and cDNA sequences from biological material. Analysis of both *TaPIF3* knockout lines and *TaPIF3* overexpression line have indicated a possible role of TaPIF3 in the regulation of stem elongation. Results obtained from a subsequent comparative phenotype experiment, however, are not consistent enough to support this finding. It is possible that

functional redundancy exists between TaPIF3 and additional wheat PIF proteins. If this is the case, it may be necessary to knockout multiple *PIF* genes in wheat in order to observe a robust phenotype. Functional redundancy between PIFs is well characterised in Arabidopsis, for instance in the regulation of skotomorphogenesis. As such, the *pi1 pif3 pif4 pif5* quadruple mutant (*pifq*) displays a *constitutively photomorphogenic* (*cop*)-like phenotype, which is observed to a far lesser degree in the monogenic mutants (Shin et al, 2009, Stephenson et al, 2009).

In the immediate future, the *TaPIF3* TILLING lines should be backcrossed to remove any background mutations, which could be masking or exaggerating the true phenotype of *TaPIF3* knockout. In addition, the described overexpression lines need to be characterised for *TaPIF3* expression levels. Identification of transgenic lines with extremely high levels of *TaPIF3* expression maybe needed to detect a phenotypic abnormality. Once backcrossed TILLING lines have been generated and overexpression lines characterised, the role of TaPIF3 in regulating stem elongation could be scrutinised in a comparative phenotype experiment grown under controlled environment conditions to reduce the environmental variation that exists in the glasshouse.

In addition to stem elongation, the role of TaPIF3 in regulating other known PIF3 responses could be investigated. PIF3 is an important repressor of chloroplast development. Consequently, the *pif3* mutant displays partially formed chloroplasts in the dark (Stephenson et al, 2009). Transmission electron microscopy could therefore be used to study plastid ultrastructure in dark grown *Tapif3* knockout seedlings. Furthermore, PIF3 represses chlorophyll synthesis by regulating the expression of key chlorophyll synthesis genes (Stephenson et al, 2009). Accordingly, *pif3* mutants accumulate the chlorophyll precursor Pchlide in the dark. Accumulation of Pchlide could be studied in dark-grown seedlings of both the *TaPIF3* triple mutant and the overexpressors.

In parallel with these experiments, the ability of TaPIF3 to function as a PIF in Arabidopsis could be determined using the *TaPIF3* overexpression lines generated in Chapter 5. *TaPIF3* has been transformed into *pif1*, *pif3*, *pif1pif3* and *pifq* Arabidopsis mutant backgrounds. The ability of *TaPIF3* to complement *pif1* and *pif3* mutant phenotypes would indicate whether TaPIF3 shares conservation of function to a specific PIF in Arabidopsis.

In conclusion, this project has generated valuable genetic tools that will facilitate a comprehensive analysis of TaPIF3 function. Analysis of plants with predicted altered expression levels of TaPIF3 has indicated that TaPIF3 acts to regulate stem elongation. In wheat, mutants of the DELLA protein RHT-1, were the basis of the semi-dwarf lines so crucial to the Green Revolution. These mutants, however, display adverse pleiotropic effects such as reduced grain size and poor seedling establishment. TaPIF3 is therefore a promising target for the manipulation of stem elongation to increase wheat yields whilst avoiding these pleiotropic effects.

Chapter 7: Appendix: Supplementary Data

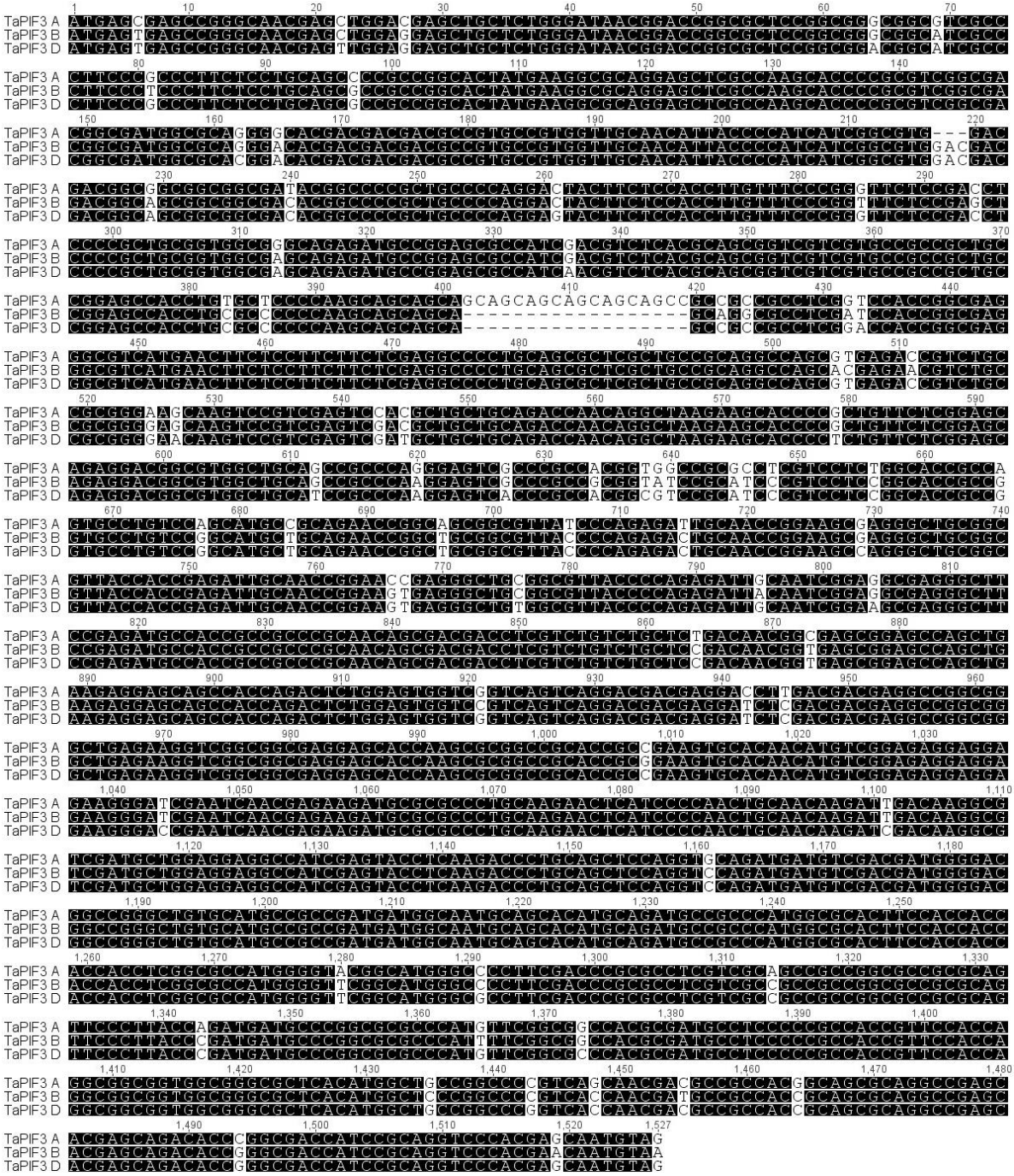


Figure S.1 Alignment of the cDNA sequences of the A, B and D homoeologues of *TaPIF3.* The CDS of the *TaPIF3* homoeologues were aligned using MUSCLE alignment. Conserved nucleotides are shown in reverse.

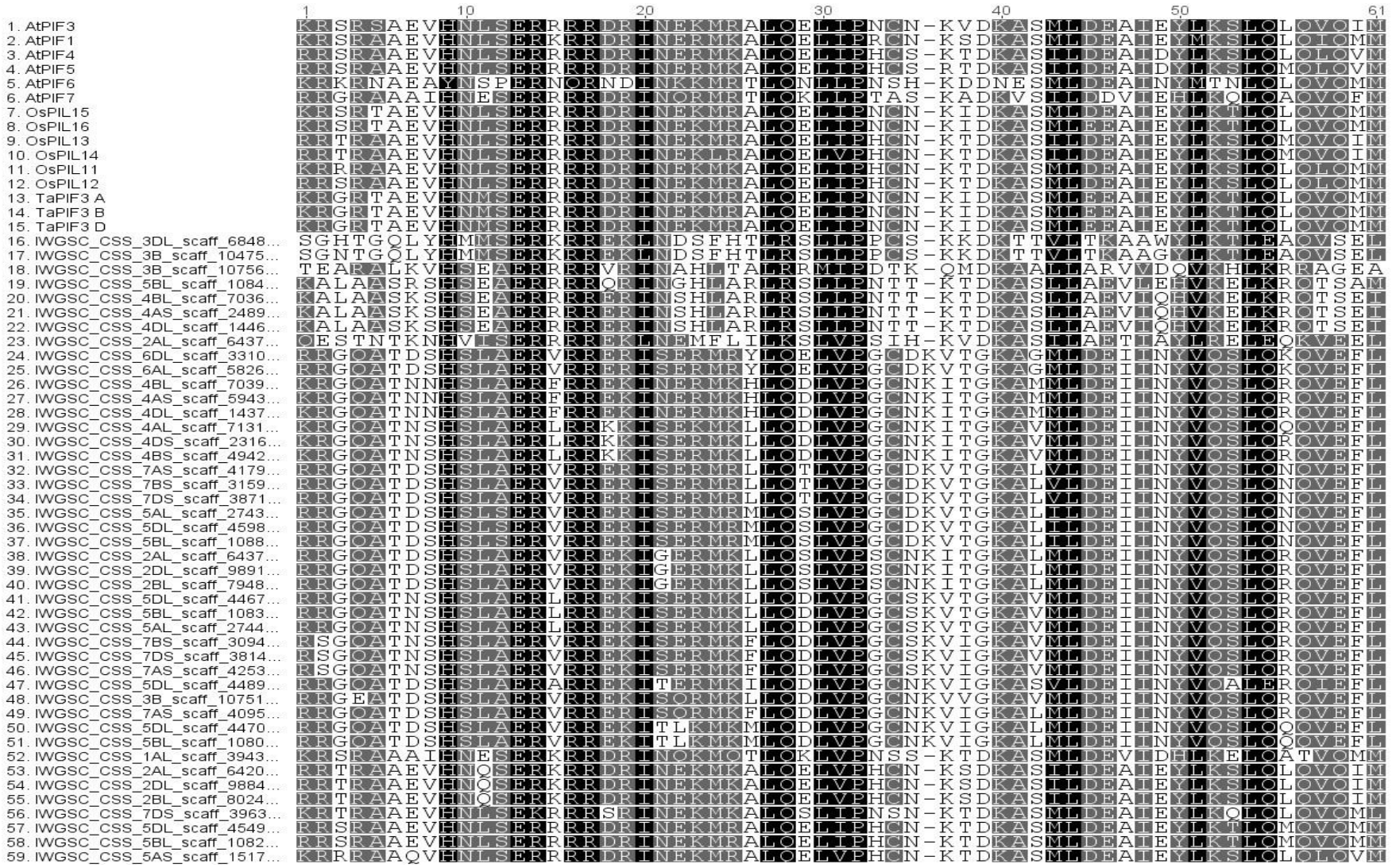


Figure S.2 Multiple sequence alignment of the bHLH domain of Arabidopsis PIFs, rice OsPILs and identified wheat sequences. Alignment was used to construct a phylogenetic tree. Wheat sequences are labelled with the corresponding genomic scaffolds.

EcoRI

GAATTCGGACGGCCATGAGTGAGCCGGGCAACGAGCTGGAGGAGCTGCTCTGGGATAACGGAC 63

CGGCGCTCCGGCGGGCGGCATCGCCCTTCCCTCCCTTCTCCTGCAGCGCCGCCGGCACTATGAAG 128

GCGCAGGAGCTCGCCAAGCACCCCGCGTCGGCGACGGCGATGGCGCAGGGACACGACGACGAC 191

GCCGTGCCGTGGTTGCAACATTACCCCATCATCGGCGTGGACGACGACGGCAGCGGCGGCGACA 255

CGGCCCCGCTGCCCCAGGACTACTTCTCCACCTTGTTTCCCGGTTTCTCCGAGCTCCCCGCTGCGG 321

TGGCGAGCAGAGATGCCGGAGCGCCATCGACGTCTCACGCAGCGGTCGTCGTGCCGCCGCTGCC 385

GGAGCCACCTGCGCCCCCCAAGCAGCAGCAGCAGGCGCCTCGATCCACCGGCGAGGGCGTCAT 448

GAACTTCTCCTTCTTCTCGAGGCCCCTGCAGCGCTCGCTGCCGCAGGCCAGCACGAGAACGTCTG 513

CCGCGGGGAGCAAGTCCGTCGAGTCGACGCTGCTGCAGACCAACAGGCTAAGAAGCACCCCGCT 577

GTTCTCGGAGCAGAGGACGGCGTGGCTGCAGCCGCCCAAGGAGTCGCCCGCCGCGGTATCCGC 640

ATCCCGTCCTCCGGCACCGCCGGTGCCTGTCCGGCATGCTGCAGAACCGGCTGCGGCGTTACCCC 705

AGAGACTGCAACCGGAAGCGAGGGCTGCGGCGTTACCACCGAGATTGCAACCGGAAGTGAGGG 768

CTGCGGCGTTACCCCAGAGATTACAATCGGAGGCGAGGGCTTCCGAGATGCCACCGCCGCCCGC 832

AACAGCGACGACCTCGTCTGTCTGCTCCGACAACGGTGAGCGGAGCCAGCTGAAGAGGAGCAGC 896

CACCAGACTCTGGAGTGGTCCGTCAGTCAGGACGACGAGGATCTCGACGACGAGGCCGGCGGG 959

CTGAGAAGGTCGGCGGCGAGGAGCACCAAGCGCGGCCGCACCGCGGAAGTGCACAACATGTCG 1022

GAGAGGAGGAGAAGGGATCGAATCAACGAGAAGATGCGCGCCCTGCAAGAACTCATCCCCAAC 1085

TGCAACAAGATTGACAAGGCGTCGATGCTGGAGGAGGCCATCGAGTACCTCAAGACCCTGCAGC 1149

TCCAGGTCCAGATGATGTCGACGATGGGGACGGCCGGGCTGTGCATGCCGCCGATGATGGCAAT 1213

GCAGCACATGCAGATGCCGCCCATGGCGCACTTCCACCACCACCACCTCGGCGCCATGGGGTTCG 1278

GCATGGGCCCCTTCGACCCGCGCCTCGTCGCCGCCGCCGGCGCCGCGCAGTTCCCTTACCCGATG 1343

ATGCCCGGCGCGCCCATTTTCGGCGGCCACGCGATGCCTCCCCCGCCACCGTTCCACCAGGCGGC 1408

GGTGGCGGGCGCTCACATGGCTCCCGGCCCCGTCACCAACGATGCCGCCACCGCAGCGCAGGCC 1472

GAGCACGAGCAGACACCGGGCGACCATCCGCAGGTCCCACGAACAATGAACGGTGAACAAAAA 1535

CTTATTTCTGAAGAAGATCTGAACGGTGAACAAAAACTTATTTCTGAAGAAGATCTGAACGGACT 1600

CGACGGTGAACAAAAACTTATTTCTGAAGAAGATCTGAACGGTGAACAAAAACTTATTTCTGAAG 1665

AAGATCTGAACGGTGAACAAAAACTTATTTCTGAAGAAGATCTGAACGGTAGCGCTTAAGATATC 1730

Myc **EcoRV**

Figure S.3 The *TaPIF3* construct sequence synthesized for overexpression studies. The full-length *TaPIF3* CDS was flanked with an EcoRI restriction site at the 5' end and an EcoRV restriction site at the 3' end. The translational ATG start codon and the TAA stop codon are indicated in bold. A 5x myc tag was added to the 3' end. The described restriction sites are underlined in bold and the myc tags are indicated by double underline.

Table S.1 Primer sequences used for PCR. The 5'-3' sequences of primers used for PCR are shown. Letters in bracket denote the homoeologue that primer was designed for**.**

Primer name	Primer sequence (5' to 3')	Length (bp)
TaPIF3_F1 (A)	ATGAGCGAGCCGGGCAAC	18
TaPIF3_F1 (B,D)	ATGAGTGAGCCGGGCAAC	18
TaPIF3_E2_F	AGGAGCTGCTCTGGGATAACG	21
TaPIF3_ME2_F	ATCGACGTCTCACGCAGC	18
TaPIF3_ME2_R	TGCGTGAGACGTCGATGG	18
TaPIF3_E2_R	CTGACTGACGGACCACTCCA	20
TaPIF3_E3_F	GAAGTGCACAACATGTCGGAG	21
TaPIF3_E3_R	CCTCTCCGACATGTTGTGCA	20
TaPIF3_E4_F	ATCGAATCAACGAGAAGATGCG	22
TaPIF3_E4_R	CTTGTTGCAGTTGGGGATGAGT	22
TaPIF3_3'_F	GCCATCGAGTACCTCAAGACC	21
TaPIF3_R1 (A)	CTACATTGCTCGTGGGACCTG	21
TaPIF3_R1 (B)	TTACATTGTTCGTGGGACCTGC	22
TaPIF3_R1 (D)	CTACATTGCTCGTGGGACCTG	21
TaPIF3_1BS_upstr1	ACACCTGTCGCCGCAAGA	18
TaPIF3_1BS_upstr2	CTCGTGACGGGAAGAGGGA	19
TaPIF3_1BS/1DS_upstr3	ACGGCTCTGCTTCGAGTTG	19
TaPIF3_5'_R	CCGTTATCCCAGAGCAGCT	19
pGWB_F	GGATGACGCACAATCCCACTATC	23
TaPIF3_attB1_F1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTGAGCCGGGCAAC	49
TaPIF3_attB2_R1	GGGGACCACTTTGTACAAGAAAGCTGGGTTTACATTGTTCGTGGGACCTGC	51

Table S.2 Primer names and sequences used for 5' RACE. The 5' to 3' sequences of primers used to perform 5' RACE are shown.

Primer name	Primer sequence (5' to 3')	Length (bp)
TaPIF3_5'RACE_1	ATGGGGTAATGTTGCAACCACGGCA	25
TaPIF3_5'RACE_2	CCGGGAAACAAGGTGGAGAAGTAGTC	26
TaPIF3_5'RACE_3	GGAGAAGTAGTCCTGGGGCAGCGG	24
TaPIF3_5'RACE_4	CCGGTCCGTTATCCCAGAGCAGCT	24
TaPIF3_5'RACE_5	GATGATGGGGTAATGTTGCAACCACGGC	28
TaPIF3_5'RACE_6	CCCAGAGCAGCTCCTCCAACTCGTT	25
TaPIF3_5'RACE_7	AGCTCCTCCAACTCGTTGCCCGG	23

Table S.3 Names and sequences of homoeologues specific primers designed to genotype three TILLING lines.

Primer name	Primer sequence (5′ to 3′)	Length (bp)
1AS_forward_till_1	TGCGTAGGAGCGAGCCGG	18
1AS_forward_till_2	CGGGCGCGGGTACTGATGG	19
1AS_forward_till_4	CTTCTCCTGCAGCCCCGC	18
1AS_forward_till_5	ACATTACCCCATCATCGGCGTG	22
1AS_reverse_till_1	CTTGGGGAGCACAGGTGGCT	20
1AS_reverse_till_2	GTGGACTCGACGGACTTGCTTC	22
1AS_reverse_till_3	CGGTGCCAGAGGACGAGGC	19
1AS_reverse_till_4	CGCCCTCGCCGGTGGACC	18
1BS_forward_till_1	AGGTAAAGCGCTGGACACATGG	22
1BS_forward_till_3	TCAGGACGACGAGGTAAAGCG	21
1BS_reverse_till_1	CGGGATCGCTACTAGTACATACC	23
1BS_reverse_till_3	TGGATGTCACGCTGTGCTCAT	21
1DS_forward_till_2	GGAACAAGTCCGTCGAGTCGAT	22
1DS_forward_till_3	AAGCAGCAGCAGCCGCC	17
1DS_forward_till_4	GAGACTGCAACCGGAAGCC	19
1DS_forward_till_5	CAACCGGAAGTGAGGGCTGT	20
1DS_reverse_till_1	GGTGTGTCCAGTGCTCACCTC	21
1DS_reverse_till_3	CTGCAACAGCAATTTCATACAACAT	25
1DS_reverse_till_4	GTACTGCGTTCGTGGTGTGT	20

List of References

Abdel-Ghany, S.E. (2009) Contribution of plastocyanin isoforms to photosynthesis and copper homeostasis in *Arabidopsis thaliana* grown at different copper regimes. *Planta,* 229, 767-779.

Abdel-Ghany, S.E., Burkhead, J.L., Gogolin, K.A., Andrés-Colás, N., Bodecker, J.R., Puig, S., Peñarrubia, L. and Pilon, M. (2005) AtCCS is a functional homolog of the yeast copper chaperone Ccs1/Lys7. *FEBS Letters,* 579 (11), 2307-2312.

Achard, P., Cheng, H., De Grauwe, L., Decat, J., Schoutteten, H., Moritz, T., Van Der Straeten, D., Peng, J. and Harberd, N.P. (2006) Integration of Plant Responses to Environmentally Activated Phytohormonal Signals. *Science,* 311 (5757), 91-94.

Achard, P., Renou, J.-P., Berthomé, R., Harberd, N.P. and Genschik, P. (2008) Plant DELLAs Restrain Growth and Promote Survival of Adversity by Reducing the Levels of Reactive Oxygen Species. *Current Biology,* 18 (9), 656-660.

Ahn, S., Anderson, J.A., Sorrells, M.E. and Tanksley, S.D. (1993) Homoeologous relationships of rice, wheat and maize chromosomes. *Molecular and General Genetics,* 241, 483-490.

Alabadí, D., Gil, J., Blázquez, M.A. and García-Martínez, J.L. (2004) Gibberellins Repress Photomorphogenesis in Darkness. *Plant Physiology,* 134 (3), 1050-1057.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. (2003) Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana*. *Science,* 301 (5633), 653-657.

Al-Sady, B., Ni, W., Kircher, S., Schafer, E. and Quail, P.H. (2006) Photoactivated Phytochrome Induces Rapid PIF3 Phosporylation Prior to Proteasome-Mediated Degradation. *Molecular Cell,* 23, 439-446.

Apel, K. and Hirt, H. (2004) REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. *Annual Review of Plant Biology,* 55 (1), 373-399.

Atchley, W.R., Terhalle, W. and Dress, A. (1999) Positional Dependence, Cliques, and Predictive Motifs in the bHLH Protein domain. *Journal of Molecular Evolution,* 48, 501-516.

Barnes, C. and Bugbee, B. (1991) Morphological Responses of Wheat to Changes in Phytochrome Photoequilibrium. *Plant Physiology,* 97, 359-365.

Bauer, D., Viczián, A., Kircher, S., Nobis, T., Nitschke, R., Kunkel, T., Panigrahi, K.C.S., Ádám, É., Fejes, E., Schäfer, E. and Nagy, F. (2004) Constitutive Photomorphogenesis 1 and Multiple Photoreceptors Control Degradation of Phytochrome Interacting Factor 3, a Transcription Factor Required for Light Signaling in Arabidopsis. *The Plant Cell,* 16 (6), 1433-1445.

Beale, S.I. (1999) Enzymes of chlorophyll biosynthesis. *Photosynthesis Research,* 60 (1), 43-73.

Beevers, L., Loveys, B., Pearson, J.A. and Wareing, P.F. (1970) Phytochrome and Hormonal Control of Expansion and Greening of Etiolated Wheat Leaves. *Planta,* 90, 286-294.

Bent, A. (2006*) Arabidopsis thaliana* Floral Dip Transformation Method IN: Wang, K. (ed.) *Agrobacterium Protocols*. Totowa, NJ: Humana Press, 87-104.

Bentley, A., Maclennan, B., Calvo, J. and Dearolf, C.R. (2000) Targeted Recovery of Mutations in Drosophila. *Genetics,* 156 (3), 1169-1173.

Bhattacharjee, S. (2005) Reactive oxygen species and oxidative burst: roles in stress, senescence and signal. *Current Science,* 89, 1113-1121.

Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., Marshall, W.S. and Khvorova, A. (2006) 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods,* 3, 199.

Boettcher, M. and Mcmanus, Michael t. (2015) Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Molecular Cell,* 58 (4), 575-585.

Borner, A., Plaschke, J., Korzun, V. and Worland, A.J. (1996) The relationships between the dwarfing genes of wheat and rye. *Euphytica,* 89, 69-75.

Borthwick, H.A., Hendricks, S.B., Parker, M.W., Toole, E.H. and Toole, V.K. (1952) A Reversible Photoreaction Controlling Seed Germination. *Proceedings of the National Academy of Sciences,* 38 (8), 662-666.

Botticella, E., Sestili, F., Hernandez-Lopez, A., Phillips, A. and Lafiandra, D. (2011) High resolution melting analysis for the detection of EMS induced mutations in wheat Sbe*lla* genes. *BMC Plant Biology,* 11 (156).

Boylan, M.T. and Quail, P.H. (1991) Phytochrome A overexpression inhibits hypocotyl elongation in transgenic Arabidopsis. *Proceedings of the National Academy of Sciences,* 88, 10806-10810.

Briat, J.-F., Curie, C. and Gaymard, F. (2007) Iron utilization and metabolism in plants. *Current Opinion in Plant Biology,* 10, 276-282.

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X. and Huard, F. (2010) Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Research,* 119, 201-212.

Cao, D., Cheng, H., Wu, W., Soo, H.M. and Peng, J. (2006) Gibberellin Mobilizes Distinct DELLA-Dependent Transcriptomes to Regulate Seed Germination and Floral Development in Arabidopsis. *Plant Physiology,* 142 (2), 509-525.

Casal, J.J., Candia, A.N. and Sellaro, R. (2014) Light perception and signalling by phytochrome A. *Journal of Experimental Botany,* 65 (11), 2835-2845.

Casal, J.J., Deregibus, V.A. and Sanchez, R.A. (1985) Variations in Tiller Dynamics and Morphology in *Lolium multiflorum* Lam. Vegetative and Reproductive Plants as affected by Differences in Red/Far-Red Irradiation. *Annals of Botany,* 56, 553-559.

Castelfranco, P.A. and Jones, O.T.G. (1975) Protoheme Turnover and Chlorophyll Synthesis in Greening Barley Tissue. *Plant Physiology,* 55 (3), 485-490.

Chandler, P.M., Marion-Poll, A., Ellis, M. and Gubler, F. (2002) Mutants at the *Slender1* Locus of Barley cv Himalaya. Molecular and Physiological Characterization. *Plant Physiology,* 129 (1), 181-190.

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M.-F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M.,

Bernard, M., Leroy, P. and Chalhoub, B. (2005) Molecular Basis of Evolutionary Events That Shaped the Hardness Locus in Diploid and Polyploid Wheat Species (Triticum and Aegilops). *The Plant Cell,* 17 (4), 1033-1045.

Chapman, J.A., Mascher, M., Buluc, A., Barry, K., Georganas, E., Session, A., Strnadova, V., Jenkins, J., Sehgal, S., Oliker, L., Schmutz, J., Yelick, K.A., Scholz, U., Waugh, R., Poland, J.A., Muehlbauer, G.J., Stein, N. and Rokhsar, D.S. (2015) A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biology,* 16 (26).

Cheminant, S., Wild, M., Bouvier, F., Pelletier, S., Renou, J.-P., Erhardt, M., Hayes, S., Terry, M.J., Genschik, P. and Achard, P. (2011) DELLAs Regulate Chlorophyll and Carotenoid Biosynthesis to Prevent Photooxidative Damage during Seedling Deetiolation in Arabidopsis. *Plant Cell,* 23 (5), 1849-1860.

Chen, A., Li, C., Hu, W., Lau, M.Y., Lin, H., Rockwell, N.C., Martin, S.S., Jernstedt, J.A., Lagarias, J.C. and Dubcovsky, J. (2014) PHYTOCHROME C plays a major role in the acceleration of wheat flowering under long-day photoperiod. *Proceedings of the National Academy of Sciences,* 111 (28), 10037-10044.

Cheng, M., Fry, J.E., Pang, S., Zhou, H., Hironaka, C.M., Duncan, D.R., Conner, T.W. and Wan, Y. (1997) Genetic Transformation of Wheat Mediated by *Agrobacterium tumefaciens*. *Plant Physiology,* 115 (3), 971-980.

Chiba, Y., Shimizu, T., Miyakawa, S., Kanno, Y., Koshiba, T., Kamiya, Y. and Seo, M. (2015) Identification of *Arabidopsis thaliana* NRT1/PTR FAMILY (NPF) proteins capable of transporting plant hormones. *Journal of Plant Research,* 128 (4), 679-686.

Christensen, A.H., Sharrock, R.A. and Quail, P.H. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplast by electroporation. . *Plant Molecular Biology,* 18, 675-689.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Ruden, D.M. and Lu, X. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w^{1118};iso-2;iso-3. *Fly,* 6 (2), 80-92.

Clavijo, B.J., Venturini, L., Schudoma, C., Accinelli, G.G., Kaithakottil, G., Wright, J., Borrill, P., Kettleborough, G., Heavens, D., Chapman, H., Lipscombe, J., Barker, T., Lu, F.-H., Mckenzie, N., Raats, D., Ramirez-Gonzalez, R.H., Coince, A., Peel, N., Percival-Alwyn, L., Duncan, O., Trösch, J., Yu, G., Bolser, D.M., Namaati, G., Kerhornou, A., Spannagl, M., Gundlach, H., Haberer, G., Davey, R.P., Fosker, C., Palma, F.D., Phillips, A.L., Millar, A.H., Kersey, P.J., Uauy, C., Krasileva, K.V., Swarbreck, D., Bevan, M.W. and Clark, M.D. (2017) An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Research,* 27 (5), 885-896.

Clemens, S. (2001) Molecular mechanisms of plant metal tolerance and homeostasis. *Planta,* 212 (4), 475-486.

Cordell, D., Drangert, J.-O. and White, S. (2009) The story of phosphorus: Global food security and food for thought. *Global Environmental Change* 19, 292-305.

Cornah JE, Terry MJ, & Smith AG (2003) Green or red: what stops the traffic in the tetrapyrrole pathway? *Trends in plant science* 8(5):224-230.

Cornah, J.E., Roper, J.M., Pal Singh, D. and Smith, A.G. (2002) Measurement of ferrochelatase activity using a novel assay suggests that plastids are the major site of haem biosynthesis in both photosynthetic and non-photosynthetic cells of pea (Pisum sativum L.). *Biochemical Journal,* 362 (Pt 2), 423-432.

Davière, J.-M. and Achard, P. (2013) Gibberellin signaling in plants. *Development,* 140 (6), 1147-1151.

De Lucas, M., Daviere, J.-M., Rodriguez-Falcon, M., Pontin, M., Iglesias-Pedraz, J.M., Lorrain, S., Fankhauser, C., Blazquez, M.A., Titarenko, E. and Prat, S. (2008) A molecular framework for light and gibberellin control of cell elongation. *Nature,* 451 (7177), 480-484.

Demarsy, E. and Fankhauser, C. (2009) Higher plants use LOV to perceive blue light. *Current Opinion in Plant Biology,* 12 (1), 69-74.

Deregibus, V.A., Sanchez, R.A. and Casal, J.J. (1983) Effects of Light Quality on Tiller Production in *Lolium* spp. *Plant Physiology,* 72, 900-902.

Dill, A. and Sun, T.-P. (2001) Synergistic Derepression of Gibberellin Signaling by Removing RGA and GAI Function in *Arabidopsis thaliana*. *Genetics,* 159 (2), 777-785.

Dill, A., Jung, H.-S. and Sun, T.-P. (2001) The DELLA motif is essential for gibberellin-induced degradation of RGA. *Proceedings of the National Academy of Sciences,* 98 (24), 14162-14167.

Dill, A., Thomas, S.G., Hu, J., Steber, C.M. and Sun, T.-P. (2004) The Arabidopsis F-Box Protein SLEEPY1 Targets Gibberellin Signaling Repressors for Gibberellin-Induced Degradation. *The Plant Cell,* 16 (6), 1392-1405.

Duek, P.D. and Fankhauser, C. (2005) bHLH class transcription factors take centre stage in phytochrome signaling. *Trends in Plant Science,* 10 (2), 51-54.

Duggan, J. and Gassman, M. (1974) Induction of Porphyrin Synthesis in Etiolated Bean Leaves by Chelators of Iron. *Plant Physiology,* 53 (2), 206-215.

Eckardt, N. (2007) Thylakoid Development from Biogenesis to Senescence, and Ruminations on Regulation. *The Plant Cell,* 19, 1135-1138.

Eckardt, N.A. (2006) Three Arabidopsis GID1 Genes Encode Gibberellin Receptors with Overlapping Functions. *The Plant Cell,* 18 (12), 3353-3353.

Edwards, D., Batley, J. and Snowdon, R.J. (2013) Accessing complex crop genomes with next-generation sequencing. *Theoretical and Applied Genetics,* 126, 1-11.

Eriksson, S., Bohlenius, H., Moritz, T. and Nilsson, O. (2006) GA_4 Is the Active Gibberellin in the Regulation of *Leafy* Transcription and Arabidopsis Floral Initiation *The Plant Cell,* 18, 2172-2181.

Feldman M., Bonjean AP. and Angus WJ. (2001) 'Origin of cultivated wheat', *The world wheat book: a history of wheat breeding,* Paris, France Lavoisier Publishing pg. 3-56.

Feng, S., Martinez, C., Gusmaroli, G., Wang, Y., Zhou, J., Wang, F., Chen, L., Yu, L., Iglesias-Pedraz, J.M., Kircher, S., Schafer, E., Fu, X., Fan, L.-M. and Deng, X.W. (2008) Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins. *Nature,* 451 (7177), 475-479.

Flintham, J.E. and Gale, M.D. (1983) The 'Tom Thumb' dwarfing gene, the *Rht3* in wheat. Effects on height, yield and grain quality. *Theoretical and Applied Genetics,* 66, 249-256.

Flintham, J.E., Börner, A., Worland, A.J. and Gale, M.D. (1997) Optimizing wheat grain yield: effects of Rht (gibberellin-insensitive) dwarfing genes. *The Journal of Agricultural Science,* 128 (1), 11-25.

Foo, E., Platten, J.D., Weller, J.L. and Reid, J.B. (2006) PhyA and cry1 act redundantly to regulate gibberellin levels during de-etiolation in blue light. *Physiologia Plantarum,* 127, 149-156.

Food and Agriculture Organization of the United Nations (FAO). 1996. *Rome Declaration on World Food Security and World Food Summit Plan of Action.* Rome.

Food and Agriculture Organization of the United Nations (FAO). 2009. *State of Food Insecurity in the World.* Rome.

Food and Agriculture Organization of the United Nations (FAO). 2002. *BREAD WHEAT Improvement and Production*. Rome.

Food and Agriculture Organization of the United Nations (FAO). 2017. *The Future of Food and Agriculture – Trends and Challenges.* Rome.

Food and Agriculture Organization of the United Nations (FAO). 2011*. The state of the world's land and water resources for food and agriculture (SOLAW) – Managing systems at risk*. Rome and Earthscan, London.

Foster, C.A. (1977) Slender: an accelerated extension growth mutant of barley. *Barly Genetics Newsletter,* 7, 24-27.

Foyer, C.H., Descourvières, P. and Kunert, K.J. (1994) Protection against oxygen radicals: an important defence mechanism studied in transgenic plants. *Plant, Cell & Environment,* 17 (5), 507-523.

Franck, F., Sperling, U., Frick, G., Pochert, B., Van Cleve, B., Apel, K. and Armstrong, G.A. (2000) Regulation of Etioplast Pigment-Protein Complexes, Inner Membrane Architecture, and Protochlorophyllide a Chemical Heterogeneity by Light-Dependent NADPH:Protochlorophyllide Oxidoreductases A and B. *Plant Physiology,* 124 (4), 1678-1696.

Franklin, K.A. and Quail, P.H. (2010) Phytochrome functions in Arabidopsis development. *Journal of Experimental Botany,* 61 (1), 11-24.

Franklin, K.A., Larner, V.S. and Whitelam, G.C. (2005) The signal transducing photoreceptors of plants. *The International Journal of Deveopmental Biology.,* 49, 653-664.

Franklin, K.A., Lee, S.H., Patel, D., Kumar, S.V., Spartz, A.K., Gu, C., Ye, S., Yu, P., Breen, G., Cohen, J.D., Wigge, P.A. and Gray, W.M. (2011) PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) regulates auxin biosynthesis at high temperature. *Proceedings of the National Academy of Sciences,* 108 (50), 20231-20235.

Frick, G., Su, Q., Apel, K. and Armstrong, G.A. (2003) An Arabidopsis porB porC double mutant lacking light-dependent NADPH:protochlorophyllide oxidoreductases B and C is highly chlorophyll-deficient and developmentally arrested. *The Plant Journal,* 35 (2), 141-153.

Frigerio, M., Alabadí, D., Pérez-Gómez, J., García-Cárcel, L., Phillips, A.L., Hedden, P. and Blázquez, M.A. (2006) Transcriptional Regulation of Gibberellin Metabolism Genes by Auxin Signaling in Arabidopsis. *Plant Physiology,* 142 (2), 553-563.

Fu, D., Uauy, C., Blechl, A. and Dubcovsky, J. (2007) RNA interference for wheat functional gene analysis. *Transgenic Research,* 16 (6), 689-701.

Fu, X., Richards, D.E., Ait-Ali, T., Hynes, L.W., Ougham, H., Peng, J. and Harberd, N.P. (2002) Gibberellin-Mediated Proteasome-Dependent Degradation of the Barley DELLA Protein SLN1 Repressor. *The Plant Cell,* 14 (12), 3191-3200.

Gao, X.-H., Xiao, S.-L., Yao, Q.-F., Wang, Y.-J. and Fu, X.-D. (2011) An updated GA signaling 'relief of repression' regulatory model. *Molecular Plant,* 4 (4), 601-606.

Gao, Y., Jiang, W., Dai, Y., Xiao, N., Zhang, C., Li, H., Lu, Y., Wu, M., Tao, X., Deng, D. and Chen, J. (2015) A maize phytochrome-interacting factor 3 improves drought and salt stress tolerance in rice. *Plant Molecular Biology,* 87 (4), 413-428.

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

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M. and Toulmin, C. (2010) Food Security: The Challenge of Feeding 9 Billion People. *Science,* 327 (5967), 812-818.

Goff, S.A. (1999) Rice as a model for cereal genomics. *Current opinion in plant biology,* 2 (2), 86-89.

Green, B.R. (2011) Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal,* 66, 34-44.

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L. and Henikoff, S. (2003) Spectrum of Chemically Induced Mutations From a Large-Scale Reverse-Genetic Screen in Arabidopsis. *Genetics,* 164 (2), 731-740.

Gupta, A.S., Heinen, J.L., Holaday, A.S., Burke, J.J. and Allen, R.D. (1993) Increased resistance to oxidative stress in transgenic plants that overexpress chloroplastic Cu/Zn superoxide dismutase. *Proceedings of the National Academy of Sciences,* 90, 1629-1633.

Hall, J.L. (2002) Cellular mechanisms for heavy metal detoxification and tolerance. *Journal of Experimental Botany,* 53 (366), 1-11.

Hall, J.L. and Williams, L.E. (2003) Transition metal transporters in plants. *Journal of Experimental Botany,* 54 (393), 2601-2613.

Halliwell, B. (2006) Reactive Species and Antioxidants. Redox Biology Is a Fundamental Theme of Aerobic Life. *Plant Physiology,* 141 (2), 312-322.

Hanaoka M, Kanamaru K, Fujiwara M, Takahashi H, & Tanaka K (2005) Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO reports* 6(6):545-550.

Hedden, P. (2003) The genes of the Green Revolution. *Trends in Genetics,* 19 (1), 5-9.

Hedden, P. (2012) Gibberellin Biosynthesis *eLS*. John Wiley & Sons, Ltd.

Hedden, P. and Thomas, S G. (2012) Gibberellin biosynthesis and its regulation. *Biochemical Journal,* 444 (1), 11-25.

Henry, I.M., Nagalakshmi, U., Lieberman, M., Ngo, K.J., Krasileva, K.V., Vasquez-Gross, H., Akhunova, A., Akhunov, E., Dubcovsky, J., Tai, T.H. and Comai, L. (2014) Efficient Genome-Wide Detection and

Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. *The Plant Cell,* 26, 1382-1397.

Hensel, G., Himmelbach, A., Chen, W., Douchkov, D.K. and Kumlehn, J. (2011) Transgene expression systems in the Triticeae cereals. *Journal of Plant Physiology,* 168, 30-44.

Hirano, K., Kouketu, E., Katoh, H., Aya, K., Ueguchi-Tanaka, M. and Matsuoka, M. (2012) The suppressive function of the rice DELLA protein SLR1 is dependent on its transcriptional activation activity. *The Plant Journal,* 71 (3), 443-453.

Hu, J., Mitchum, M.G., Barnaby, N., Ayele, B.T., Ogawa, M., Nam, E., Lai, W.-C., Hanada, A., Alonso, J.M., Ecker, J.R., Swain, S.M., Yamaguchi, S., Kamiya, Y. and Sun, T.-P. (2008) Potential Sites of Bioactive Gibberellin Production during Reproductive Growth in Arabidopsis. *The Plant Cell,* 20 (2), 320-336.

Huang, D.-D. and Wang, W.-Y. (1986) Genetic control of chlorophyll biosynthesis: Regulation of delta-aminolevulinate synthesis in Chlamydomonas. *Molecular and General Genetics MGG,* 205 (2), 217-220.

Huq, E., Al-Sady, B., Hudson, M., Kim, C., Apel, K. and Quail, P.H. (2004) PHYTOCHROME-INTERACTING FACTOR 1 Is a Critical bHLH Regulator of Chlorophyll Biosynthesis. *Science,* 305 (5692), 1937-1941.

Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M. and Yamaguchi, J. (2001) *slender* Rice, a Constitutive Gibberellin Response Mutant, Is Caused by a Null Mutation of the *SLR1* Gene, an Ortholog of the Height-Regulating Gene *GAI/RGA/RHT/D8*. *The Plant Cell,* 13 (5), 999-1010.

Inoue, K., Nishihama, R., Kataoka, H., Hosaka, M., Manabe, R., Nomoto, M., Tada, Y., Ishizaki, K. and Kohchi, T. (2016) Phytochrome Signaling Is Mediated by PHYTOCHROME INTERACTING FACTOR in the Liverwort *Marchantia polymorpha*. *The Plant Cell,* 28 (6), 1406-1421.

International Wheat Genome Sequencing Consortium. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science,* 361 (6403).

Ishijima, S., Uchibori, A., Takagi, H., Maki, R. and Ohnishi, M. (2003) Light-induced increase in free Mg2+ concentration in spinach chloroplasts: Measurement of free Mg2+ by using a fluorescent probe and necessity of stromal alkalinization. *Archives of Biochemistry and Biophysics,* 412 (1), 126-132.

Jackson, M.A., Anderson, D.J. and Birch, R.G. (2013) Comparison of Agrobacterium and particle bombardment using whole plasmid or minimal cassette for production of high-expressing, low-copy transgenic plants. *Transgenic Research,* 22 (1), 143-151.

Jiang, B., Shi, Y., Zhang, X., Xin, X., Qi, L., Guo, H., Li, J. and Yang, S. (2017) PIF3 is a negative regulator of the *CBF* pathway and freezing tolerance in Arabidopsis. *Proceedings of the National Academy of Sciences,* 114 (32), E6695-E6702.

Jiang, C. and Fu, X. (2007) GA action: turning on de-DELLA repressing signaling. *Current Opinion in Plant Biology,* 10 (5), 461-465.

Jung, H.-S. and Chory, J. (2010) Signaling between Chloroplasts and the Nucleus: Can a Systems Biology Approach Bring Clarity to a Complex and Highly Regulated Pathway? *Plant Physiology,* 152 (2), 453-459.

Kaiser, W.M. (1979) Reversible inhibition of the Calvin cycle and activation of oxidative pentose phosphate cycle in isolated intact chloroplasts by hydrogen peroxide. *Planta,* 145 (4), 377-382.

Kami, C., Lorrain, S., Hornitschek, P. and Fankhauser, C. (2010) Chapter Two - Light-Regulated Plant Growth and Development IN: Timmermans, M.C.P. (ed.) *Current Topics in Developmental Biology*. Academic Press, 29-66.

Kaneko, M., Itoh, H., Inukai, Y., Sakamoto, T., Ueguchi-Tanaka, M., Ashikari, M. and Matsuoka, M. (2003) Where do gibberellin biosynthesis and gibberellin signaling occur in rice plants. *The Plant Journal,* 35, 104-115.

Kawahara, Y., Bastide, M., Hamilton, J.P., Kanamori, H., Mccombie, W.R., Ouyang, S., Schwartz, D.C., Tanaka, T., Wu, J., Zhou, S., Childs, K.L., Davidson, R.M., Lin, H., Quesada-Ocampo, L., Vaillancourt, B., Sakai, H., Shin Lee, S., Kim, J., Numa, H., Itoh, T., Buell, C.R. and Matsumoto, T. (2013) Improvements of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice,* 6 (4).

Khanna, R., Huq, E., Kikis, E.A., Al-Sady, B., Lanzatella, C. and Quail, P.H. (2004) A Novel Molecular Recognition Motif Necessary for Targeting Photoactivated Phytochrome Signaling to Specific Basic Helix-Loop-Helix Transcription Factors. *The Plant Cell,* 16 (11), 3033-3044.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology,* 14 (4), R36.

King, R., Bird, N., Ramirez-Gonzalez, R., Coghill, J.A., Patil, A., Hassani-Pak, K., Uauy, C. and Phillips, A.L. (2015) Mutation Scanning in Wheat by Exon Capture and Next-Generation Sequencing. *PLoS One,* 10 (9).

King, R.W., Moritz, T., Evans, L.T., Junttila, O. and Herlt, A.J. (2001) Long-Day Induction of Flowering in Lolium temulentum Involves Sequential Increases in Specific Gibberellins at the Shoot Apex. *Plant Physiology,* 127 (2), 624-632.

Kipreos, E.T. and Pagano, M. (2000) The F-box protein family. *Genome Biology,* 1 (5), 3002.3001-3002.3007.

Kleine T, Maier UG, & Leister D (2009) DNA transfer from organelles to the Nucleus: The Idiosyncratic Genetics of Endosymbiosis. *Annual Review of Plant Biology* 60:115-138.

Kondou, Y., Higuchi, M. and Matsui, M. (2010) High-Throughput Characterization of Plant Gene Functions by Using Gain-of-Function Technology. *Annual Review of Plant Biology,* 61 (1), 373-393.

Kong, S.-G. and Okajima, K. (2016) Diverse photoreceptors and light responses in plants. *Journal of Plant Research,* 129 (2), 111-114.

Koornneef, M. and Van Der Veen, J.H. (1980) Induction and Analysis of Gibberellin Sensitive Mutants in *Arabidopsis thaliana* (L.) Heynh. *Theoretical and Applied Genetics,* 58, 257-263.

Koornneef, M., Elgersma, A., Hanhart, C.J., Van Loenen-Martinet, E.P., Van Rijn, L. and Zeevaart, J.a.D. (1985) A gibberellin insensitive mutant of *Arabidopsis thaliana*. *Physiologia Plantarum,* 65 (1), 33-39.

Krasileva, K.V., Vasquez-Gross, H.A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R.H., Wang, X., Borrill, P., Fosker, C., Ayling, S., Phillips, A.L., Uauy, C. and

Dubcovsky, J. (2017) Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences,* 114 (6), E913-E921.

Krieger-Liszkay, A. (2005) Singlet oxygen production in photosynthesis. *Journal of Experimental Botany,* 56 (411), 337-346.

Kumar, I., Swaminathan, K., Hudson, K. and Hudson, M.E. (2016) Evolutionary divergence of phytochrome protein function in *Zea mays* PIF3 signalling. *Journal of Experimental Botany,* 67 (14), 4231-4240.

Kumar, S.V., Lucyshyn, D., Jaeger, K.E., Alós, E., Alvey, E., Harberd, N.P. and Wigge, P.A. (2012) Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature,* 484, 242-245.

Kumaran, M.K., Bowman, J.L. and Sundaresan, V. (2002) *YABBY* Polarity Genes Mediate the Repression of *KNOX* Homeobox Genes in Arabidopsis. *The Plant Cell,* 14 (11), 2761-2770.

Ladha, J.K., Dawe, D., Pathak, H., Padre, A.T., Yadav, R.L., Singh, B., Singh, Y., Singh, Y., Singh, P., Kundu, A.L., Sakal, R., Ram, N., Regmi, A.P., Gami, S.K., Bhandari, A.L., Amin, R., Yadav, C.R., Bhattarai, E.M., Das, S., Aggarwal, H.P., Gupta, R.K. and R, H.P. (2003) How extensive are yield declines in long-term rice-wheat experiments in Asia? *Field Crops Research,* 81, 159-180.

Lanahan, M.B. and Ho, T.-H.D. (1988) Slender barley: A constitutive gibberellin-response mutant. *Planta,* 175 (1), 107-114.

Larkin, R.M., Alonso, J.M., Ecker, J.R. and Chory, J. (2003) GUN4, a Regulator of Chlorophyll Synthesis and Intracellular Signaling. *Science,* 299 (5608), 902-906.

Lawit, S.J., Wych, H.M., Xu, D., Kundu, S. and Tomes, D.T. (2010) Maize DELLA Proteins dwarf plant8 and dwarf plant9 as Modulators of Plant Development. *Plant and Cell Physiology,* 51 (11), 1854-1868.

Leach, L.J., Belfield, E.J., Jiang, C., Brown, C., Mithani, A. and Harberd, N.P. (2014) Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat. *BMC Genomics,* 15 (276).

Lee, S., Cheng, H., King, K.E., Wang, W., He, Y., Hussain, A., Lo, J., Harberd, N.P. and Peng, J. (2002) Gibberellin regulates Arabidopsis seed germination via *RGL2*, a *GAI/RGA*-like gene whose expression is up-regulated following imbibition. *Genes & Development,* 16, 646-658.

Leivar, P. and Monte, E. (2014) PIFs: Systems Integrators in Plant Development. *The Plant Cell,* 26, 56-78.

Leivar, P. and Quail, P.H. (2011) PIFs: pivotal components in a cellular signaling hub. *Trends in Plant Science,* 16 (1), 19-28.

Leivar, P., Monte, E., Al-Sady, B., Carle, C., Storer, A., Alonso, J.M., Ecker, J.R. and Quail, P.H. (2008) The Arabidopsis Phytochrome-Interacting Factor PIF7, Together with PIF3 and PIF4, Regulates Responses to Prolonged Red Light by Modulating phyB Levels. *The Plant Cell,* 20 (2), 337-352.

Leivar, P., Monte, E., Oka, Y., Liu, T., Carle, C., Castillon, A., Huq, E. and Quail, P.H. (2008) Multiple Phytochrome-Interacting bHLH Transcription Factors Repress Premature Seedling Photomorphogenesis in Darkness. *Current Biology,* 18 (23), 1815-1823.

Li, F.-W., Melkonian, M., Rothfels, C.J., Villarreal, J.C., Stevenson, D.W., Graham, S.W., Wong, G.K.-S., Pryer, K.M. and Mathews, S. (2015) Phytochrome diversity in green plants and the origin of canonical plant phytochromes. *Nature Communications,* 6, 7852.

Li, J., Li, G., Wang, H. and Wang Deng, X. (2011) Phytochrome Signalling Mechanisms. *The Arabidopsis Book / American Society of Plant Biologists,* 9, e0148.

Li, K., Yu, R., Fan, L.-M., Wei, N., Chen, H. and Deng, X.W. (2016) DELLA-mediated PIF degradation contributes to coordination of light and gibberellin signalling in Arabidopsis. *Nature Communications,* 7, 11868.

Li, X., Duan, X., Jiang, H., Sun, Y., Tang, Y., Yuan, Z., Guo, J., Liang, W., Chen, L., Yin, J., Ma, H., Wang, J. and Zhang, D. (2006) Genome-Wide Analysis of Basic/Helix-Loop-Helix Transcription Factor Family in Rice and Arabidopsis. *Plant Physiology,* 141, 1167-1184

Li, X.-P., Lan, S.-Q., Liu, Y.-P., Gale, M.D. and Worland, T.J. (2006) Effects of different *Rht-B1b*, *Rht-D1b* and *Rht-B1c* dwarfing genes on agronomic characteristics in wheat. *Cereal Research Communications,* 34 (2), 919-924.

Lin, K. and Zhang, D.-Y. (2005) The excess of 5' introns in eukaryotic genomes. *Nucleic Acids Research,* 33 (20), 6522-6527.

Liu, Z., Zhang, Y., Wang, J., Li, P., Zhao, C., Chen, Y. and Bi, Y. (2015) Phytochrome-interacting factors PIF4 and PIF5 negatively regulate anthocyanin biosynthesis under red light in Arabidopsis seedlings. *Plant Science,* 238, 64-72.

Llorente, B., D'andrea, L., Ruiz-Sola, M.A., Botterweg, E., Pulido, P., Andilla, J., Loza-Alvarez, P. and Rodriguez-Concepcion, M. (2016) Tomato fruit carotenoid biosynthesis is adjusted to actual ripening progression by a light-dependent mechanism. *The Plant Journal,* 85 (1), 107-119.

Lobell, D.B., Schlenker, W. and Costa-Roberts, J. (2011) Climate Trends and Global Crop Production Since 1980. *Science,* 333 (6042), 616-620.

Lorrain, S., Allen, T., Duek, P.D., Whitelam, G.C. and Fankhauser, C. (2008) Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors. *The Plant Journal,* 53, 312-323.

Martínez-García, J.F., Huq, E. and Quail, P.H. (2000) Direct Targeting of Light Signals to a Promoter Element-Bound Transcription Factor. *Science,* 288 (5467), 859-863.

Masuda, T. and Fujita, Y. (2008) Regulation and evolution of chlorophyll metabolism. *Photochemical & Photobiological Sciences,* 7 (10), 1131-1149.

Matthews, S. and Sharrock, R. (1996) The phytochrome gene family in grasses (Poaceae): a phylogeny and evidence that grasses have a subset of the loci found in dicot angiosperms. *Molecular Biology and Evolution,* 13, 1141-1150.

Mccallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (20000) Targeted screening for induced mutations. *Nature Biotechnology,* 18, 455-457.

Mcginnis, K.M. (2010) RNAi for functional genomics in plants. *Briefings in Functional Genomics,* 9 (2), 111-117.

Mcginnis, K.M., Thomas, S.G., Soule, J.D., Strader, L.C., Zale, J.M., Sun, T.-P. and Steber, C.M. (2003) The Arabidopsis SLEEPY1 Gene Encodes a Putative F-Box Subunit of an SCF E3 Ubiquitin Ligase. *The Plant Cell,* 15 (5), 1120-1130.

Meister, G. and Tuschi, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature,* 431, 343-439.

Meskauskiene, R. and Apel, K. (2002) Interaction of FLU, a negative regulator of tetrapyrrole biosynthesis, with the glutamyl-tRNA reductase requires the tetratricopeptide repeat domain of FLU. *FEBS Letters,* 532 (1–2), 27-30

Meskauskiene, R., Nater, M., Goslings, D., Kessler, F., Op Den Camp, R. and Apel, K. (2001) FLU: A negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences,* 98 (22), 12826-12831.

Met Office. 2018. *Was summer 2018 the hottest on record?*. [ONLINE] Available at: https://www.metoffice.gov.uk/news/releases/2018/end-of-summer-stats. [Accessed 6 September 2018]

Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., Mcanulla, C., Mcmenamin, C., Nuka, G., Pesseat, S., Sangrador-Vegas, A., Scheremetjew, M., Rato, C., Yong, S.Y., Bateman, A., Punta, M., Attwood, T.K., Sigrist, C.J., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P. and Letunic, I. (2015) The InterPro protein families database: the classification resource after 15 years. 43 (Database issue), D213-221.

Mochizuki, N., Brusslan, J.A., Larkin, R., Nagatani, A. and Chory, J. (2001) Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proceedings of the National Academy of Sciences,* 98 (4), 2053-2058.

Mochizuki, N., Tanaka, R., Grimm, B., Masuda, T., Moulin, M., Smith, A.G., Tanaka, A. and Terry, M.J. (2010) The cell biology of tetrapyrroles: a life and death struggle. *Trends in Plant Science,* 15 (9), 488-498.

Monte, E., Tepperman, J.M., Al-Sady, B., Kaczorowski, K.A., Alonso, J.M., Ecker, J.R., Li, X., Zhang, Y. and Quail, P.H. (2004) The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. *Proceedings of the National Academy of Sciences,* 101 (46), 16091-16098.

Montgomery, B.L. and Lagarias, J.C. (2002) Phytochrome ancestry: sensors of bilins and light. *Trends in Plant Science,* 7 (8), 357-366.

Mullet, J.E. (1988) Chloroplast Development and Gene Expression. *Annual Review of Plant Physiology and Plant Molecular Biology,* 39 (1), 475-502.

Murase, K., Hirano, Y., Sun, T.-P. and Hakoshima, T. (2008) Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature,* 456, 459-463.

Nagatani, A. (2004) Light-regulated nuclear localization of phytochromes. *Current Opinion in Plant Biology,* 7 (6), 708-711.

Nair, S.K. and Burley, S.K. (2000) Functional genomics: Recognizing DNA in the library. *Nature,* 404 (6779), 715-718.

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., Watanabe, Y., Nakamura, K., Kimura, T. and Ishiguro, S. (2007) Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. *Bioscience, Biotechnology, and Biochemistry,* 71 (8), 2095-2100.

Nakajima, M., Shimada, A., Takashi, Y., Kim, Y.C., Park, S.H., Ueguchi-Tanaka, M., Suzuki, H., Katoh, E., Iuchi, S., Kobayashi, M., Maeda, T., Matsuoka, M. and Yamaguchi, I. (2006) Identification and characterization of Arabidopsis gibberellin receptors. *The Plant Journal,* 46 (5), 880-889.

Nakamura Y, Kato T, Yamashino T, Murakami M, & Mizuno T (2007) Characterization of a set of phytochrome-interacting factor-like bHLH proteins in *Oryza sativa*. Bioscience, Biotechnology, and Biochemistry 71(5):1183-1191.

Neill, S., Desikan, R. and Hancock, J. (2002) Hydrogen peroxide signalling. *Current Opinion in Plant Biology,* 5 (5), 388-395.

Ni, M., Tepperman, J.M. and Quail, P.H. (1998) PIF3, a Phytochrome-Interacting Factor Necessary for Normal Photoinduced Signal Transduction, Is a Novel Basic Helix-Loop-Helix Protein. *Cell,* 95 (5), 657-667.

Ni, M., Tepperman, J.M. and Quail, P.H. (1999) Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature,* 400 (6746), 781-784.

Nicholls, P. (1978) Response of Barley Shoot Apices to Application of Gibberellic Acid and Abscisic Acid: Dependance on Tissue Sensitivity. *Australian Journal of Plant Physiology,* 5 (5), 581-588.

Nielsen, H. and Wernersson, R. (2006) An overabundance of phase 0 introns immediately after the start codon in eukaryotic genes. *BMC Genomics,* 7 (256).

Niyogi, K.K. (1999) PHOTOPROTECTION REVISITED: Genetic and Molecular Approaches. *Annual Review of Plant Physiology and Plant Molecular Biology,* 50 (1), 333-359.

Oh, E., Yamaguchi, S., Hu, J., Yusuke, J., Jung, B., Paik, I., Lee, H.-S., Sun, T.-P., Kamiya, Y. and Choi, G. (2007) PIL5, a Phytochrome-Interacting bHLH Protein, Regulates Gibberellin Responsiveness by Binding Directly to the GAI and RGA Promoters in Arabidopsis Seeds. *The Plant Cell,* 19 (4), 1192-1208.

Oh, E., Yamaguchi, S., Kamiya, Y., Bae, G., Chung, W.-I. and Choi, G. (2006) Light activates the degradation of PIL5 protein to promote seed germination through gibberellin in Arabidopsis. *The Plant Journal,* 47, 124-139.

O'Neill, D.P. and Ross, J.J. (2002) Auxin Regulation of the Gibberellin Pathway in Pea. *Plant Physiology,* 130 (4), 1974-1982.

Op Den Camp, R.G.L., Przybyla, D., Ochsenbein, C., Laloi, C., Kim, C., Danon, A., Wagner, D., Hideg, É., Göbel, C., Feussner, I., Nater, M. and Apel, K. (2003) Rapid Induction of Distinct Stress Responses after the Release of Singlet Oxygen in Arabidopsis. *The Plant Cell,* 15 (10), 2320-2332.

Park, E., Kim, J., Lee, Y., Shin, J., Oh, E., Chung, W.-I., Liu, J.R. and Choi, G. (2004) Degradation of Phytochrome Interacting Factor 3 in Phytochrome-Mediated Light Signaling. *Plant and Cell Physiology,* 45 (8), 968-975.

Park, H., Kreunen, S.S., Cuttriss, A.J., Dellapenna, D. and Pogson, B.J. (2002) Identification of the Carotenoid Isomerase Provides Insight into Carotenoid Biosynthesis, Prolamellar Body Formation, and Photomorphogenesis. *The Plant Cell,* 14 (2), 321-332.

Parks, B.M., Quail, P.H. and Hangarter, R.P. (1996) Phytochrome A Regulates Red-Light Induction of Phototropic Enhancement in Arabidopsis. *Plant Physiology,* 110 (1), 155-162.

Parry, M.a.J., Madgwick, P.J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., Labhilili, M. and Phillips, A.L. (2009) Mutation discovery for crop improvement. *Journal of Experimental Botany,* 60 (10), 2817-2825.

Pearce, S., Kippes, N., Chen, A., Debernardi, J.M. and Dubcovsky, J. (2016) RNA-seq studies using wheat PHYTOCHROME B and PHYTOCHROME C mutants reveal shared and specific functions in the regulation of flowering and shade-avoidance pathways. *BMC Plant Biology,* 16 (1), 141.

Pearce, S., Saville, R., Vaughan, S.P., Chandler, P.M., Wilhelm, E.P., Sparks, C.A., Al-Kaff, N., Korolev, A., Boulton, M.I., Phillips, A.L., Hedden, P., Nicholson, P. and Thomas, S.G. (2011) Molecular characterization of Rht-1 dwarfing genes in hexaploid wheat. *Plant Physiology,* 157 (4), 1820-1831.

Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J., Murphy, G.P. and Harberd, N.P. (1997) The Arabidopsis *GAI* gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes & Development,* 11 (23), 3194-3205.

Peng, J., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D. and Harberd, N.P. (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature,* 400 (6741), 256-261.

Peter, E. and Grimm, B. (2009) GUN4 Is Required for Posttranslational Control of Plant Tetrapyrrole Biosynthesis. *Molecular Plant,* 2 (6), 1198-1210.

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

Pham, V.N., Kathare, P.K. and Huq, E. (2018) Phytochromes and Phytochrome Interacting Factors. *Plant Physiology,* 176, 1025-1038.

Pingault, L., Choulet, F., Alberti, A., Glover, N., Wincker, P., Feuillet, C. and Paux, E. (2015) Deep transcriptome sequencing provides new insights into the structural and functional organization of the wheat genome. *Genome Biology,* 16 (29).

Pogson, B., Ganguly, D. and Albrecht-Borth (2015) Insights into chloroplast biogenesis and development. *Biochimica et Biophysica Acta,* 1847 (9), 1017-1024.

Pogson, B.J., Woo, N.S., Forster, B. and Small, I., D (2008) Plastid signalling to the nucleus and beyond. *Trends in Plant Science,* 13 (11), 602-609.

Possart, A., Xu, T., Paik, I., Hanke, S., Keim, S., Hermann, H.-M., Wolf, L., Hiß, M., Becker, C., Huq, E., Rensing, S.A. and Hiltbrunner, A. (2017) Characterization of Phytochrome Interacting Factors from the Moss *Physcomitrella patens* Illustrates Conservation of Phytochrome Signaling Modules in Land Plants. *The Plant Cell,* 29 (2), 310-330.

Pratt, A.J. and Macrae, I.J. (2009) The RNA-induced Silencing Complex: A versatile Gene-silencing Machine. *The Journal of Biological Chemistry* 284 (27), 17897-17901.

Proebsting, W.M., Hedden, P., Lewis, M.J., Croker, S.J. and Proebsting, L.N. (1992) Gibberellin Concentration and Transport in Genetic Lines of Pea. *Plant Physiology,* 100, 1354-1360.

Pyke, K.A. and Leech, R.M. (1994) A Genetic Analysis of Chloroplast Division and Expansion in *Arabidopsis thaliana*. *Plant Physiology,* 104, 201-207.

Rae, T.D., Schmidt, P.J., Pufahl, R.A., Culotta, V.C. and V. O'halloran, T. (1999) Undetectable Intracellular Free Copper: The Requirement of a Copper Chaperone for Superoxide Dismutase. *Science,* 284 (5415), 805-808.

Rakszegi, M., Kisgyorgy, B.N., Tearall, K., Shewry, P.R., Lang, L., Phillips, A. and Bedo, Z. (2010) Diversity of agronomic and morphological traits in a mutant population of bread wheat studied in the Healtgrain program. *Euphytica,* 174, 409-421.

Rebetzke, G.J., Richards, R.A., Fettell, N.A., Long, M., Condon, A.G., Forrester, R.I. and Botwright, T.L. (2007) Genotypic increases in coleoptile length improves stand establishment, vigour and grain yield of deep-sown wheat. *Field Crops Research,* 100 (1), 10-23.

Reed, J.W., Nagatani, A., Elich, T.D., Fagan, M. and Chory, J. (1994) Phytochrome A and Phytochrome B Have Overlapping but Distinct Functions in Arabidopsis Development. *Plant Physiology,* 104, 1139-1149.

Reed, J.W., Nagatani, A., Elich, T.D., Fagan, M. and Chory, J. (1994) Phytochrome A and Phytochrome B Have Overlapping but Distinct Functions in Arabidopsis Development. *Plant Physiology,* 104, 1139-1149.

Regnault, T., Davière, J.-M. and Achard, P. (2016) Long-distance transport of endogenous gibberellins in Arabidopsis. *Plant Signaling & Behavior,* 11 (1), e1110661.

Regnault, T., Davière, J.-M., Wild, M., Sakvarelidze-Achard, L., Heintz, D., Carrera Bergua, E., Lopez Diaz, I., Gong, F., Hedden, P. and Achard, P. (2015) The gibberellin precursor GA12 acts as a long-distance growth signal in Arabidopsis. *Nature Plants,* 1, 15073.

Reid, J.B., Botwright, N.A., Smith, J.J., O'neill, D.P. and Kerckhoffs, L.H.J. (2002) Control of Gibberellin Levels and Gene Expression during De-Etiolation in Pea. *Plant Physiology,* 128 (2), 734-741.

Richards DE, King KE, Ait-ali T, & Harberd NP (2001) How Gibberellin Regulates Plant Growth and Development: A Molecular Genetic Analysis of Gibberellin Signalling. *Annual Review of Plant Physiology and Plant Molecular Biology*. 52:67-88.

Rockwell, N.C., Su, Y.-S. and Lagarias, J.C. (2006) Phytochrome Structure and Signaling Mechanisms. *Annual Review of Plant Biology,* 57 (1), 837-858.

Rodríguez-Villalón, A., Gas, E. and Rodríguez-Concepción, M. (2009) Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown Arabidopsis seedlings. *The Plant Journal,* 60 (3), 424-435.

Rooke, L., Byrne, D. and Salgueiro, S. (2000) Marker gene expression driven by the maize ubiquitin promoter in transgenic wheat. *Annals of Applied Biology,* 136, 167-172.

Rosado, D., Gramegna, G., Cruz, A., Lira, B.S., Freschi, L., Setta, N.D. and Rossi, M. (2016) Phytochrome Interacting Factors (PIFs) in *Solanum lycopersicum*: Diversity, Evolutionary History and Expression Profiling during Different Developmental Processes. *PLoS One,* 11 (11).

Ross, J.J., Davidson, S.E., Wolbang, C.M., Bayly-Stark, E., Smith, J.J. and Reid, J.B. (2003) Developmental regulation of the gibberellin pathway in pea shoots. *Functional Plant Biology,* 30 (1), 83-89.

Russell, G.E. (2013) *Progress in Plant Breeding—1*. Elsevier Science.

Saito, H., Oikawa, T., Hamamoto, S., Ishimaru, Y., Kanamori-Sato, M., Sasaki-Sekimoto, Y., Utsumi, T., Chen, J., Kanno, Y., Masuda, S., Kamiya, Y., Seo, M., Uozumi, N., Ueda, M. and Ohta, H. (2015) The jasmonate-responsive GTR1 transporter is required for gibberellin-mediated stamen development in Arabidopsis. *Nature Communications,* 6, 6095.

Sakuraba, Y., Jeong, J., Kang, M.-Y., Kim, J., Paek, N.-C. and Choi, G. (2014) Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in Arabidopsis. *Nature Communications,* 5, 4636.

Sakuraba, Y., Kim, E.-Y., Han, S.-H., Piao, W., An, G., Todaka, D., Yamaguchi-Shinozaki, K. and Paek, N.-C. (2017) Rice Phytochrome-Interacting Factor-Like1 (OsPIL1) is involved in the promotion of chlorophyll biosynthesis through feed-forward regulatory loops. *Journal of Experimental Botany,* 68 (15), 4103-4114.

Sánchez-Fernández, R., Ardiles-Díaz, W., Montagu, M.V., Inzé, D. and May, M.J. (1998) Cloning of a novel *Arabidopsis thaliana* RGA-like gene, a putative member of the VHIID-domain transcription factor family. *Journal of Experimental Botany,* 49 (326), 1609-1610.

Sasaki, A., Itoh, H., Gomi, K., Ueguchi-Tanaka, M., Ishiyama, K., Kobayashi, M., Jeong, D.-H., An, G., Kitano, H., Ashikari, M. and Matsuoka, M. (2003) Accumulation of Phosphorylated Repressor for Gibberellin Signaling in an F-box Mutant. *Science,* 299 (5614), 1896-1898.

Sato S, Nakamura Y, Kaneko T, Asamizu E, & Tabata S (1999) Complete Structure of the Chloroplast Genome of *Arabidopsis thaliana*. *DNA Research* 6(5):283-290.

Schützendübel, A. and Polle, A. (2002) Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization. *Journal of Experimental Botany,* 53 (372), 1351-1365.

Schwechheimer, C. and Willige, B.C. (2009) Shedding light on gibberellic acid signalling. *Current Opinion in Plant Biology,* 12 (1), 57-62.

Seo, M., Hanada, A., Kuwahara, A., Endo, A., Okamoto, M., Yamauchi, Y., North, H., Marion-Poll, A., Sun, T.-P., Koshiba, T., Kamiya, Y., Yamaguchi, S. and Nambara, E. (2006) Regulation of hormone metabolism in Arabidopsis seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. *The Plant Journal,* 48 (3), 354-366.

Sharma, P., Jha, A.B., Dubey, R.S. and Pessarakli, M. (2012) Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. *Journal of Botany,* 2012, 26.

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

Shewry, P.R. (2009) Wheat. *Journal of Experimental Botany,* 60 (6), 1537-1553.

Shi, Q., Zhang, H., Song, X., Jiang, Y.E., Liang, R. and Li, G. (2018) Functional Characterization of the Maize Phytochrome-Interacting Factors PIF4 and PIF5. *Frontiers in Plant Science,* 8 (2273).

Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M. and Muricho, G. (2013) Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security,* 5 (3), 291-317.

Shin, A.-Y., Han, Y.-J., Baek, A., Ahn, T., Kim, S.Y., Nguyen, T.S., Son, M., Lee, K.W., Shen, Y., Song, P.-S. and Kim, J.-I. (2016) Evidence that phytochrome functions as a protein kinase in plant light signalling. *Nature Communications,* 7, 11545.

Shin, J., Kim, K., Kang, H., Zulfugarov, I.S., Bae, G., Lee, C.-H., Lee, D. and Choi, G. (2009) Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors. *Proceedings of the National Academy of Sciences,* 106 (18), 7660-7665.

Silverstone, A.L., Ciampaglio, C.N. and Sun, T.-P. (1998) The Arabidopsis *RGA* Gene Encodes a Transcriptional Regulator Repressing the Gibberellin Signal Transduction Pathway. *The Plant Cell,* 10 (2), 155-169.

Silverstone, A.L., Mak, P.Y.A., Martinez, E.C. and Sun, T.-P. (1997) The New *RGA* Locus Encodes a Negative Regulator of Gibberellin Response in *Arabidopsis thaliana*. *Genetics,* 146 (3), 1087-1099.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. and Facciotti, D. (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nature Biotechnology,* 23, 75-81.

Smith, H. and Jackson, G.M. (1987) Rapid Phytochrome Regulation of Wheat Seedling Extension *Plant Physiology,* 84, 1059-1062.

Sood, S., Gupta, V. and Tripathy, B.C. (2005) Photoregulation of the greening process of wheat seedlings grown in red light. *Plant Molecular Biology,* 59, 269-287.

Soy, J., Leivar, P., González-Schain, N., Martín, G., Diaz, C., Sentandreu, M., Al-Sady, B., Quail, P.H. and Monte, E. (2016) Molecular convergence of clock and photosensory pathways through PIF3–TOC1 interaction and co-occupancy of target promoters. *Proceedings of the National Academy of Sciences,* 113 (17), 4870-4875.

Sparks, C.A. and Jones, H.D. (2014) Genetic Transformation of Wheat via Particle Bombardment IN: Henry, R.J. and Furtado, A. (eds.) *Cereal Genomics: Methods and Protocols*. Totowa, NJ: Humana Press, 201-218.

Sponsel, V.M. and Hedden, P. (2004) Gibberellin Biosynthesis and Inactivation IN: Davies, P.J. (ed.) *Plant Hormones: Biosynthesis, Signal Transduction, Action!* Dordrecht: Springer Netherlands, 63-94.

Steber, C.M., Cooney, S.E. and Mccourt, P. (1998) Isolation of the GA-Response Mutant *sly1* as a Suppressor of *ABI1-1* in *Arabidopsis thaliana*. *Genetics,* 149 (2), 509-521.

Stephenson, P.G., Fankhauser, C. and Terry, M.J. (2009) PIF3 is a repressor of chloroplast development. *Proceedings of the National Academy of Sciences,* 106 (18), 7654-7659.

Susek, R.E., Ausubel, F.M. and Chory, J. (1993) Signal transduction mutants of arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell,* 74 (5), 787-799.

Talon, M., Koornneef, M. and Zeevaart, J.a.D. (1990) Accumulation of C_{19}-gibberellins in the gibberellin-insensitive dwarf mutant *gai* of *Arabidopsis thaliana* (L.) Heynh. *Planta,* 182, 501-505.

Tanaka, R. and Tanaka, A. (2007) Tetrapyrrole Biosynthesis in Higher Plants. *Annual Review of Plant Biology,* 58 (1), 321-346.

Tchounwou, P.B., Yedjou, C.G., Patlolla, A.K. and Sutton, D.J. (2012) Heavy Metals Toxicity and the Environment. *EXS,* (101), 133-164.

Tepperman, J., Hwang, Y.-S. and Quail, P.H. (2006) phyA dominates in transduction of red-light signals to rapidly responding genes at the initiation of Arabidopsis seedling de-etiolation. *the plant journal,* 48 (5), 728-742.

Tepperman, J., Tong, Z., Chang, H.-S., Wang, X. and Quail, P.H. (2001) Multiple transcription-factor genes are early targets of phytochrome A signaling. *Proceedings of the National Academy of Sciences,* 98 (16), 9437-9442.

Terry, M.J. and Kendrick, R.E. (1999) Feedback Inhibition of Chlorophyll Synthesis in the Phytochrome Chromophore-Deficient aurea andyellow-green-2 Mutants of Tomato. *Plant Physiology,* 119 (1), 143-152.

Terry, M.J. and Smith, A.G. (2013) A model for tetrapyrrole synthesis as the primary mechanism for plastid-to-nucleus signaling during chloroplast biogenesis. *Frontiers in Plant Science,* 4 (14), 1-14.

Tilman, D., Cassman, K.G., Matsons, P.A., Naylor, R. and Polasky, S. (2002) Agricultural sustainability and intensive production practices. *Nature,* 418, 671-677.

Todaka, D., Nakashima, K., Maruyama, K., Kidokoro, S., Osakabe, Y., Ito, Y., Matsukura, S., Fujita, Y., Yoshiwara, K., Ohme-Takagi, M., Kojima, M., Sakakibara, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2012) Rice phytochrome-interacting factor-like protein OsPIL1 functions as a key regulator of internode elongation and induces a morphological response to drought stress. *Proceedings of the National Academy of Sciences,* 109 (39), 15947-15952.

Toledo-Ortiz, G., Huq, E. and Quail, P.H. (2003) The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family. *The Plant Cell,* 15 (8), 1749-1770.

Toledo-Ortiz, G., Huq, E. and Rodríguez-Concepción, M. (2010) Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proceedings of the National Academy of Sciences,* 107 (25), 11626-11631.

Toyomasu, T., Kagahara, T., Hirose, Y., Usui, M., Abe, S., Okada, K., Koga, J., Mitsuhashi, W. and Yamane, H. (2009) Cloning and characterization of cDNAs encoding ent-copalyl diphosphate synthases in wheat: insight into the evolution of rice phytoalexin biosynthetic genes. *Bioscience, Biotechnology and Biochemistry,* 73 (3), 772-775.

Trebst, A., Depka, B. and Holländer-Czytko, H. (2002) A specific role for tocopherol and of chemical singlet oxygen quenchers in the maintenance of photosystem II structure and function in *Chlamydomonas reinhardtii*. *FEBS Letters,* 516 (1-3), 156-160.

Triantaphylidès, C., Krischke, M., Hoeberichts, F.A., Ksas, B., Gresser, G., Havaux, M., Van Breusegem, F. and Mueller, M.J. (2008) Singlet Oxygen Is the Major Reactive Oxygen Species Involved in Photooxidative Damage to Plants. *Plant Physiology,* 148 (2), 960-968.

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C., Tran, R.K., Khan, A.A., Filkov, V., Tai, T.H., Dubcovsky, J. and Comai, L. (2011) Discovery of Rare Mutations in Populations: TILLING by sequencing. *Plant Physiology,* 156, 1257-1268.

Tyler, L., Thomas, Stephen g., Hu, J., Dill, A., Alonso, J.M., Ecker, J.R. and Sun, T.-P. (2004) DELLA Proteins and Gibberellin-Regulated Seed Germination and Floral Development in Arabidopsis. *Plant Physiology,* 135 (2), 1008-1019.

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R.K., Tsai, H., Berardi, S., Comai, L. and Dubcovsky, J. (2009) A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biology,* 9 (115).

Ueguchi-Tanaka, M., Ashikari, M., Nakajima, M., Itoh, H., Katoh, E., Kobayashi, M., Chow, T.-Y., Hsing, Y.-I.C., Kitano, H., Yamaguchi, I. and Matsuoka, M. (2005) GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature,* 437, 693.

Ulijasz, A.T. and Vierstra, R.D. (2011) Phytochrome structure and photochemistry: recent advances toward a complete molecular picture. *Current Opinion in Plant Biology,* 14 (5), 498-506.

United Nations 2009. World population prospects: the 2008 revision population database. New York, UN Population Division.

Varbanova, M., Yamaguchi, S., Yang, Y., Mckelvey, K., Hanada, A., Borochov, R., Yu, F., Jikumaru, Y., Ross, J., Cortes, D., Ma, C.J., Noel, J.P., Mander, L., Shulaev, V., Kamiya, Y., Rodermel, S., Weiss, D. and Pichersky, E. (2007) Methylation of Gibberellins by Arabidopsis GAMT1 and GAMT2. *The Plant Cell,* 19 (1), 32-45.

Vavilin DV & Virmaas WFJ (2002) Regulation of the tetrapyrrole biosynthetic pathway leading to heme and chlorophyll in plants and cyanobacteria. PHYSIOLOGIA PLANTARUM 115:9–24.

Virgin, H.I. (1962) Light-induced Unfolding of the Grass Leaf. *Physiologia Plantarum,* 15, 380-389.

Walker, C., Yu, G.-H. and Weinstein, J. (1997) Comparative study of heme and Mg-protoporphyrin (monomethyl ester) biosynthesis in isolated pea chloroplasts: effects of ATP and metal ions. *Plant physiology and biochemistry,* 35 (3), 213-221.

Walker, C.J. and Willows, R.D. (1997) Mechanism and regulation of Mg-chelatase. *Biochemical Journal,* 327 (Pt 2), 321-333.

Wallace, B. (2017) The Interaction between Light and Gibberellin in the Regulation of Wheat Architecture. PhD thesis. University of Southampton.

Waters MT & Langdale JA (2009) The making of a chloroplast. *The EMBO Journal* 28(19):2861-2873.

Weeks, J.T., Anderson, O.D. and Blechl, A.E. (1993) Rapid Production of Multiple Independent Lines of Fertile Transgenic Wheat (*Triticum aestivum*). *Plant Physiology,* 102 (4), 1077-1084.

Wheatgenome.org.uk. (2015). *Triticeae Genomics For Sustainable Agriculture*. [online] Available at: http://www.wheatgenome.org.uk/ [Accessed 20 Jul. 2016].

Wilde, A., Churin, Y., Schubert, H. and Borner, T. (1997) Disruption of a *Synechocystis* sp. PCC 6803 gene with partial similarity to phytochrome genes alters growth under changing light qualities. *FEBS Letters,* 406, 89-92.

Wilde, A., Mikolajczyk, S., Alawady, A., Lokstein, H. and Grimm, B. (2004) The gun4 gene is essential for cyanobacterial porphyrin metabolism. *FEBS Letters,* 571 (1-3), 119-123.

Wildner, G.F. and Henkel, J. (1979) The effect of divalent metal ions on the activity of Mg++ depleted ribulose-1,5-bisphosphate oxygenase. *Planta,* 146 (2), 223-228.

Winkler, R.G. and Freeling, M. (1994) Physiological genetics of the dominant gibberellin-nonresponsive maize dwarfs, Dwarf8 and Dwarf9. *Planta,* 193 (3), 341-348.

Wolbang, C.M., Chandler, P.M., Smith, J.J. and Ross, J.J. (2004) Auxin from the Developing Inflorescence Is Required for the Biosynthesis of Active Gibberellins in Barley Stems. *Plant Physiology,* 134 (2), 769-776.

Xu, H., Liu, Q., Yao, T. and Fu, X. (2014) Shedding light on integrative GA signaling. *Current Opinion in Plant Biology,* 21, 89-95.

Yabuta, T. and Sumiki, Y. (1938) On the crystal of gibberellin, a substance to promote plant growth. *Journal of the Agricultural Chemical Society of Japan,* 14, 1526.

Yamaguchi, S. (2008) Gibberellin Metabolism and its Regulation. *Annual Review of Plant Biology,* 59 (1), 225-251.

Yamaguchi, S., Kamiya, Y. and Sun, T.-P. (2001) Distinct cell-specific expression patterns of early and late gibberellin biosynthetic genes during Arabidopsis seed germination. *The Plant Journal,* 28 (4), 443-453.

Yeh, K.-C. and Lagarias, J.C. (1998) Eukaryotic phytochromes: Light-regulated serine/threonine protein kinases with histidine kinase ancestry. *Proceedings of the National Academy of Sciences,* 95 (23), 13976-13981.

Yeh, K.-C., Wu, S.-H., Murphy, J.T. and Lagarias, J.C. (1997) A Cyanobacterial Phytochrome Two-Component Light Sensory System. *Science,* 277 (5331), 1505-1508.

Zhao, X.-Y., Yu, X.-H., Liu, X.-M. and Lin, C.-T. (2007) Light Regulation of Gibberellins Metabolism in Seedling Development. *Journal of Integrative Plant Biology,* 49 (1), 21-27.]

Zhong, S., Shi, H., Xue, C., Wang, L., Xi, Y., Li, J., Quail, Peter h., Deng, Xing w. and Guo, H. (2012) A Molecular Framework of Light-Controlled Phytohormone Action in Arabidopsis. *Current Biology,* 22 (16), 1530-1535.

Zhou, J., Liu, Q., Wang, Y., Zhang, S., Cheng, H., Yan, L., Li, L., Xie, X., Zhang, F. and Chen, F. (2014) Overexpression of OsPIL15, a phytochrome-interacting factor-like protein gene, represses etiolated seedling growth in rice. *Journal of Integrative Plant Biology,* 56 (4), 373-387.

Zhou, Q., Hare, P.D., Yang, S.W., Zeidler, M., Huang, L.-F. and Chua, N.-H. (2005) FHL is required for full phytochrome A signaling and shares overlapping functions with FHY1. *The Plant Journal,* 43 (3), 356-370.

Zhu, Y., Nomura, T., Xu, Y., Zhang, Y., Peng, Y., Mao, B., Hanada, A., Zhou, H., Wang, R., Li, P., Zhu, X., Mander, L.N., Kamiya, Y., Yamaguchi, S. and He, Z. (2006) *ELONGATED UPPERMOST INTERNODE* Encodes a Cytochrome P450 Monooxygenase That Epoxidizes Gibberellins in a Novel Deactivation Reaction in Rice. *The Plant Cell,* 18 (2), 442-456.

Zhu, Y., Tepperman, J.M., Fairchild, C.D. and Quail, P.H. (2000) Phytochrome B binds with greater apparent affinity than phytochrome A to the basic helix–loop–helix factor PIF3 in a reaction requiring the PAS domain of PIF3. *Proceedings of the National Academy of Sciences,* 97 (24), 13419-13424.