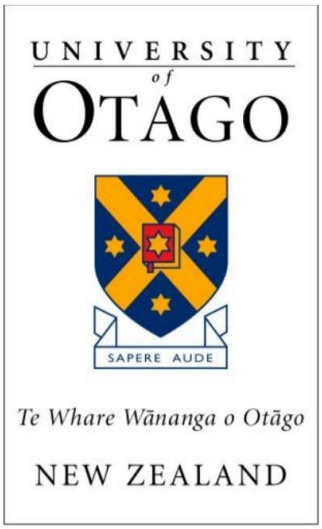
Functional Conservation Of The Male Germ-line Regulator DUO1 In Flowering Plants

Jonathan Casey



A Thesis Submitted For The Degree Of

Master Of Science

At The University Of Otago, Dunedin, New Zealand

Biochemistry Department

May, 2017

Acknowledgements

The work I have achieved in my project would not have been possible with the help I have received from a number of people. Firstly I would like to sincerely thank my supervisor Dr Lynette Brownfield, for all her scientific guidance, patience and support she has giving me and allowing me to undergo this project.

I would also personally like to thank Simon Cabout for building my foundations of laboratory operations with all the guidance he initially gave me in my work when I was fresh into the lab. Also I'd personally like to thank Ben Peters and Robyn Lee for helping out whenever I had questions regarding lab work. A massive thanks also goes to all other members of lab 308 past and present for making it such a great atmosphere.

Lastly a massive thanks also goes to my family and friends who have supported and motivated me in my university endeavours and encouraged me to challenge myself. I wouldn't have been able to do this without such a great bunch of people in my life.

Abstract

In flowering plants the development of the male germ-line in pollen to form two functional sperm cells is essential for reproduction. DUO1 an R2R3 MYB transcription factor plays a key role in regulating genes required for germ-line division and specification into mature sperm cells. Mutant *duo1* pollen contain a singular germ-like cell that cannot fertilize. Present studies of DUO1 have focused on the model organism *Arabidopsis thaliana* with regulation of *DUO1* being of interest. This study tests if regulatory mechanisms are evolutionarily conserved across diverse flowering plant species with a focus on the legume *Medicago truncatula*.

A DUO1 orthologue in *M. truncatula* was able to complement the *A. thaliana duo1* mutant through restoring mitotic function at pollen mitosis II. Specification and differentiation of the male germ-line was also restored as shown by successful male transmission from the complemented male germ-line. This indicated functional conservation between the two proteins.

In the *A. thaliana DUO1* promoter a *cis*-regulatory module, *Regulatory region of DUO1* (*ROD1)* is known to provide germ-line specific expression. Here analysis of the full length *M. truncatula DUO1* promoter using a fluorescent promoter reporter showed male germ-line specific GFP fluorescence in *A. thaliana* during pollen development similar to that of the *A. thaliana DUO1* promoter. Analysis of a 5′ deletion series of the *M. truncatula DUO1* promoter then identified key regions necessary for expression with no GFP seen in a construct lacking the *ROD1* module. The *ROD1* regulatory module from *M. truncatula* was then isolated and used in a promoter reporter to show that *M. truncatula ROD1* provides germ-line specific expression similar to *A. thaliana ROD1*. Collectively these results demonstrated that *ROD1* is an evolutionarily conserved *cis*-regulatory module that determines germ-line specific expression of *DUO1* critical for pollen development in dicotyledonous plants.

A. thaliana and *M. truncatula* DUO1 is thought to be able to auto-activate expression through DUO1 binding sites in the *DUO1* promoter. This was explored through dual luciferase assays, however technical difficulties lead to inconclusive results and no clear conclusions could be made.

Table of Contents

2. Methods

5. Auto-activation of DUO1

6. Discussion

List of Figures

List of Tables

List of abbreviations

°C	Degrees Celsius
BLAST	Basic Local Alignment Search Tool
ATG	Translational start site
A. thaliana	*Arabidopsis thaliana*
AthDUO1	*Athaliana truncatula DUO1*
AthROD1	*ROD1* region within AthDUO1 promoter
att	Gateway® attachment sites
A. tumerfaciens	*Agrobacterium tumefaciens*
bp	Base pair
CamV 35S	Cauliflower mosaic Virus *35S*
CYCB1;1	CyclinB1;1
DAPI	4,6-diamidino-2-phenylindole
DAT	DUO1 activated target
DAZ	DUO1-Activated Zinc Finger
DNA	2-Deoxyribonucleic Acid
dNTP	deoxynucleotide triphosphate
DUO1	DUO POLLEN GRAIN 1
E. coli	*Escherichia coli*
EMSA	electro-mobility shift assay
F1	First generation post cross
fmol	Femtomole
g	Grams
g	Gravity
gDNA	Genomic DNA

gem1	*gemini pollen* 1
GFP	Green fluorescent protein
GRSF	Germ-line restrictive silencing factor
H2B	Histone H2B
kb	Kilo-base pair
LB	Luria bertani
LCG1	Lily Generative Cell Specific 1
M	Molar
mm	Milli-meters
mM	Milli-molar
MGH3	Male gamete-specifc histone 3
mL	Milli-litre
MS	Murashige and Skoog
MtrDUO1	*Medicago truncatula* DUO1
MtrROD1	ROD1 region within MtrDUO1 promoter
Ng	Nanogram
PCR	Polymerase chain reaction
promDUO1	Promoter of DUO1
ROD1	Regulator of DUO 1
ROD2	Regulator of DUO 2
RFP	Red fluorescent protein
T1	First generation of transformed Arabidopsis
µl	microlitre
v/v	volume per volume
w/v	Weight per volume
WT	Wild-type

Chapter 1: Introduction

1.1 Plant sexual reproduction

The sexual reproduction of angiosperms (flowering plants) relies on an alternation of generations between a sporophyte (diploid; 2n) and a gametophyte (haploid; 1n) generation (Sargant, 1900). These generations differ significantly: morphologically, the sporophyte that is commonly seen consists of features like the roots, stem, leaves and flowers. The sporophyte is the dominant generation that ends in production of the spore via a meiotic division (Bhatt *et al.,* 2001). The spore develops into the reduced (1n) gametophyte generation that produces gametes. The male gametophyte is the pollen grain which contains one vegetative and two sperm cells at maturity (Figure 1.1). While the embryo sac containing the egg cell and central cell, is the female gametophyte.

During reproduction one sperm cell fuses to the egg cell and produces a diploid zygote, but for successful seed development to occur there must be a double fertilisation event (First described by Navashin, 1898) between the male and female gametes. The other sperm cell fuses to two polar nuclei in the female central cell to form a triploid structure which develops into the endosperm. Both fertilizations events are needed as the zygote forms the diploid sporophyte generation and the endosperm provides the nutrients for the zygote to grow successfully (Russell, 1992).

Understanding the genetics behind the double fertilisation in plants is of critical agricultural importance for a number of reasons; propagation of crop plants relies on reproductive success as without successful double fertilisation viable seeds are unable to be produced. Double fertilization is required for the formation of the highly nutritious endosperm. The endosperm is harvested as food in staple crops such as wheat, corn, barley and legumes. This means higher fertilization rates results in increased crop yields. My project focuses on the development of the two sperm cells inside a pollen grain.

1.2 Overview of pollen development

The male gametophyte develops in two sequential stages; microsporogenesis and microgametogenesis (Borg *et al*., 2009). Microsporogenesis is the first phase in which diploid sporogenesis cells in the anther of the plant differentiate into pollen mother cells. These pollen mother cells then undergo meiosis to create a tetrad of four haploid microspores (Owen & Makaroff, 1995). These microspores are held together by a thick callosic wall. In most species, pollen is shed as a monad (singular pollen grain) which relies on an enzymatic mixture of cell wall degrading enzymes secreted by a supportive layer of cells to degrade the callose cell wall and release the microspore completing microsporogenesis (Scott *et al*., 2004).

Microgametogenesis (Figure 1.1) is the maturation process of the single cell microspore into a mature pollen grain. The microspore enlarges and becomes polarised by fusing multiple vacuoles together pushing the nucleus against the microspore cell wall (Owen and Makaroff, 1995; Yamamoto *et al*., 2003). The polarised microspore then undergoes pollen mitosis I (PMI) which is an asymmetric mitotic division (Owen and Makaroff, 1995). This creates two unequally sized daughter cells; the generative cell (start of the male germ-line) and the vegetative cell. These two daughter cells are morphologically distinct and contain unique gene expression profiles that results in the different cell fates (Rutley and Twell, 2015).

The vegetative cell is the larger cell and exits the cell cycle after PMI and remains in G1 phase (Zarsky *et al*., 1992). The vegetative cell engulfs the generative cell to form a unique cell within a cell support structure. The role of the vegetative cell is then to protect the generative cell and accumulate carbohydrates and lipid reserves along with transcripts and proteins essential for rapid pollen tube growth (Pacini, 1996). The pollen tube forms from the vegetative cell after successful pollination and delivers the sperm cells to the embryo sac for double fertilization (Twell, 2011).

After PMI, the generative cell is the start of the male germ-line (Berger and Twell, 2011) and it maintains a condensed chromatin state (Twell, 1995). The generative cell continues the cell cycle through another mitotic division called pollen mitosis II (PMII)

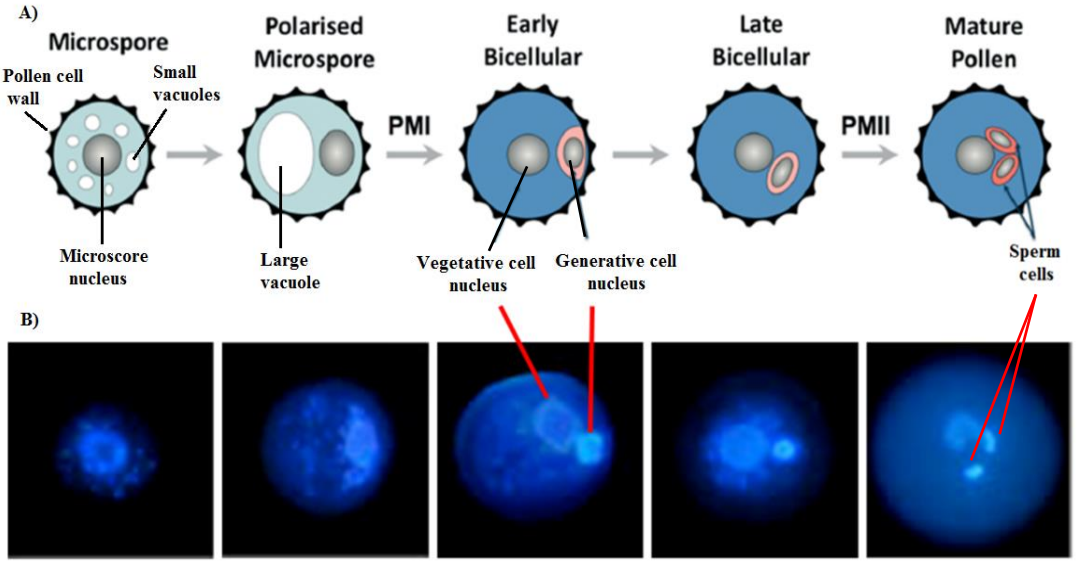


Figure 1.1: Schematic diagram of microgametogenesis in pollen of *Arabidopsis thaliana*. A) Schematic of microgametogenesis stages showing the development of a pollen grain: The microspore becomes polarized through the formation of a large vacuole then undergoes the asymmetric division PMI producing a vegetative cell and a generative cell. The generative cell then undergoes PMII to produce two cells which differentiate forming sperm cells in a mature pollen grain. **B)** DAPI stained images of pollen showing nuclei content through development. The male generative cells have condensed nuclei that stain brighter than the dispersed vegetative nucleus. Smaller dots at the microspore stage show organelles such as mitochondria that contain DNA (Figure adapted from Borg *et al*., 2011).

(Figure 1.1) and differentiates forming two genetically identical sperm cells needed for double fertilisation (Tanaka, 1997). The model plant *Arabidopsis thaliana* sheds tri-cellular pollen with PMII taking place prior to anthesis (release of pollen), however this is not the case for all angiosperms with the majority of species shedding bi-cellular pollen with PMII taking place in the growing pollen tube.

1.3 Male germ-line development

1.3.1 Role of asymmetric division in germ-line specification

Successful specification and differentiation of the male-germline relies on a number of key events, the first of which is an asymmetric division at PMI. Asymmetric cell division is necessary in animals and plants for the generation of distinct daughter cell types with various fates (Petricka *et al.*, 2009). The difference in cell fate of daughter cells can be determined by differential segregation of intrinsic cell fate determinants

or through extrinsic factors such as secreted or transmembrane proteins resulting in different signals being received in otherwise alike cells (Horvitz and Herskowitz, 1992). In animals, examples of intrinsic determinants are asymmetric mRNA and protein localisation, and stability of proteins in different daughter cells. In plants, intrinsic factors such as transcription factors have been known to regulate the creation of distinct cell fates (Petricka *et al.*, 2009; MacAlister *et al.*, 2007), but the regulatory mechanisms controlling how the transcription factors are switched on is still unknown.

Several experiments have shown that asymmetric division appears to be critical for the formation of the male germ-line in flowering plants. In *Nicotiana tabacum* when the polarisation of the microspore is disrupted prior to PMI by the microtubule inhibitor colchicine, a symmetrical division occurs. Analysis of cell fate of these symmetric daughter cells by expression of cell-type specific markers show a vegetative cell fate with vegetative cell specific markers being expressed in both cells (Eady *et al.*, 1995). Additionally, in a male gametophyte mutant *gemini pollen1* (*gem1*) discovered through an EMS mutagenic screen in *A. thaliana*, the asymmetric division is also disrupted. This resulted in cell-type specific markers showing vegetative cell fate in both daughter cells after PMI (Park *et al*., 1998). This data indicates that the vegetative cell is the default cell fate pathway and the generative cell fate is actively specified through the asymmetric division. As asymmetric division also occurs in isolated microspores (Eady *et al*., 1995), it is likely an unequal inheritance of an intrinsic factor is necessary for specification of the male germ-line.

The difference in chromatin state of the vegetative and generative cell along with asymmetrically inherited intrinsic determinants such as transcription factors or factors that regulate transcription may result in differential gene expression. So asymmetric division is necessary for the specification of the male germ-line progenitor cell by specifically turning on genes required for generative cell division likely through an asymmetrically inherited factor such as a transcription factor.

1.3.2 Transcriptome analysis of male germ-line formation

The formation of the two distinct daughter cells after PMI has led to a closer look at the genes and potential regulatory networks that are involved in their respective developments. Genes involved in male germ-line formation have been explored through the development of new techniques that allow the profiling of the transcriptome of the whole pollen grain and also more recently transcriptomes of isolated sperm cells (Rutley and Twell, 2015). These pollen specific gene expression profiles have been conducted in *Zea maize;* Corn (Engel *et al*., 2003), *Lilium longiflorum;* Lily (Okada *et al*., 2006), *A. thaliana* (Twell, 2011), *A. thaliana* sperm cells (Borges *et al*., 2008) and *Oryza sativa* (Rice) sperm cells (Russell *et al.*, 2012).

Collectively, these transcriptome profiles along with forward and reverse genetic approaches have helped develop possible gene regulatory networks of genes needed for germ-line development by helping identify key genes involved in male germ-line development (Rutley and Twell, 2015). Gene ontology enrichment in the data sets show genes for processes such as DNA repair, ubiquitin-mediated proteolysis and cell cycle progression are differentially expressed in the male germ-line compared to the vegetative cell. The different transcriptomes of the two cell types suggests that distinct gene regulatory networks occur in the vegetative cell and the male germ-line (For a review, see Rutley and Twell, 2015). Thus, the regulation of the genetic framework that switches on the male germ-line transcriptional profile in the developing generative cell is of great interest to understand for improved plant breeding.

1.3.3 Early expression of DUO1 specifies the male germ-line

While the cell fate determinant to regulate specification of the male germ-line is still unknown, a critical early event in the male germ-line development is the expression of the transcription factor DUO POLLEN 1 (DUO1) (Rotman *et al*., 2005). DUO1 is the master switch in coordinating cell cycle progression and genes required for sperm cell function. Discovering the determinant leading to the activation of *DUO1* is of great interest and understanding the regulation of *DUO1* may lead to finding the key determinant for male germline formation. DUO1 will be discussed in detail below but

a review of transcriptional regulation will first show the importance that transcription factors have on gene regulation and control of genetic networks.

1.4 Importance of transcriptional regulation

The modulated expression of genes relies on transcriptional regulation mechanisms. These mechanisms are dependent on the binding of regulatory proteins called transcription factors. Transcription factors have the ability to specifically bind certain DNA sequences called *cis*-regulatory elements (CREs). Transcription factors can bind specific CREs depending on the structural motifs they possess such as helix-loop-helix, helix-turn-helix and zinc finger binding domains (Pabo and Sauer, 1992). The selective binding of transcription factors to various DNA sequences can result in different regulatory properties. CREs are found in promoters, enhancers and silencers which are typically non-coding DNA sequences that can control rates of transcription. Transcriptional regulation is undoubtedly an important aspect of male germ-line formation as plant male germ cells have a diverse transcriptome with expression of at least 65 specific genes in the male germ-line (Engel *et al.*, 2003; Borges *et al*., 2008).

In plants the promoter is the 5´ upstream region that regulates expression of the gene. Promoters typically consist of different regulatory domains such as; Core promoters which consist of sequence elements such as the TATA box and initiator sequence are where general transcription factors bind to recruit RNA polymerase II to the pre-initiation complex at the transcription initiation site which are proximal to the gene being transcribed. In more distal sites from the core promoter there are enhancers (Increase transcription) or repressor (decrease transcription) regulatory regions with various CRE's which modulate transcription independent of their position to the initiation site via binding of transcription factors that influence formation of the pre-initiation complex (Arnosti and Kulkarni, 2005; Cowell, 1994). The chromatin state of cells plays a major role in transcriptional regulation and histone modifying enzymes are able to remodel chromatin to either favour or inhibit the binding of transcription factors to CRE's (Berger, 2007).

Transcription regulation depends on a combination of all these regulatory mechanisms to influence the complex patterns of gene expression that plants will experience due to developmental and environmental cues. Multiple different transcription factors can specifically bind different CRE's of a promoter and influence gene expression. Higher order DNA-protein complexes at various genetic loci can be formed due to the combinatorial control of transcription factor binding. This will be reflected in the gene expression pattern of any given tissue at any given time (Singh, 1998). Thus in order to understand the regulation of a certain gene it can be beneficial to identify CRE's and transcription factors involved in the transcriptional activation of the gene.

In addition post-transcriptional gene silencing (PTGS) is a mechanism that may result in genetic silencing following transcription (Vance and Vaucheret, 2001). PTGS in plants occurs when foreign or overexpressed RNA molecules are cleaved to form 21-25 nucleotide RNA molecules that can bind complementary mRNA sequences to form double stranded RNA sequences. These double stranded RNA sequences are then bound and degraded via nucleases, reducing the amount of mRNA that is translated (Hohn and Vazquez, 2011). In plants this system acts as an adaptive immune system targeted against viruses (Ratcliff *et al*., 1999)

1.5 DUO1 and germ-line differentiation in *A. thaliana*

1.5.1 The *duo1* pollen mutant

The *duo1* line in *A. thaliana* is a male gametophyte mutant identified by Durbarry e*t al*., 2005 which phenotypically only possesses one generative-like cell when stained with DAPI giving a bi-cellular phenotype rather than tri-cellular as seen in WT (Figure 1.2 A and B). The singular generative-like cell suggests cell cycle defects in the mutant with cells remaining in G2 phase of the cell cycle during PMII (Durbarry *et al*., 2005).

These singular *duo1* generative-like cells cannot successfully fertilize either the egg or central cell (Rotman *et al*., 2005) suggesting that as well as cell cycle defects, key features of male gamete differentiation and function are impaired in *duo1*. As there is no male transmission the mutant can only be seen in heterozygote (*DUO1*/*duo1)* lines

carrying the mutation. Since pollen is post meiosis 50% of pollen from a *DUO1*/*duo1* line will be wildtype (Tri-cellular) (Figure 1.2A) and 50% will be *duo1* (Bi-cellular) (Figure 1.2B).

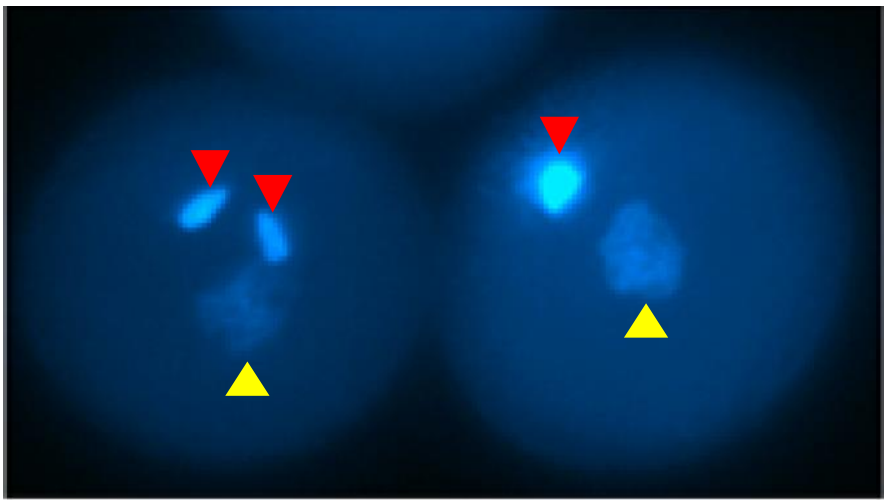


Figure 1.2: Microscope image of DAPI stained wild type and *duo1* mutant pollen of *A. thaliana*. Left) DAPI staining of tri-cellular wild type pollen grain with two sperm cell nuclei (Red arrow-head) visualized with condensed chromatin and vegetative cell nuclei (Yellow arrow-head) with dispersed chromatin. **Right)** DAPI staining of a bi-cellular *duo1* mutant showing a pollen grain with one generative like cell nuclei (Red arrow-head) and one vegetative cell nuclei (Yellow arrow-head) with dispersed chromatin. (Image adapted from Brownfield *et al*., 2009)

1.5.2 DUO1 is an R2R3 MYB transcription factor

The *DUO1* gene (At3g60460) was identified via map based cloning and sequencing by Rotman *et al*., 2005. The *DUO1* gene encodes a MYB transcription factor of the R2R3 subfamily. Currently 126 different R2R3 MYB transcription factors have been identified in *A. thaliana* with extensive involvement in regulation of gene expression during development (Yanhui *et al*., 2006). The expression of the DUO1 protein has been shown to accumulate specifically in the generative (expression is seen briefly before PMII) and sperm cell nuclei of pollen (Rotman *et al*., 2005).

The MYB family of proteins is characterised by a conserved DNA binding domain called the MYB domain (Saikumar *et al*., 1990). The R2R3 MYB subfamily DNA binding domain consists of two MYB repeats of about 52 amino acid residues each

forming three helices (Martin and Paz-Ares, 1997). The second and third helix repeat form a helix-turn-helix structure. A key characteristic of the MYB repeat is three regularly spaced tryptophan residues (Ogata *et al*., 1992), which form a tryptophan cluster in the three-dimensional helix-turn-helix structure creating a hydrophobic core that places adjacent amino acids in the appropriate spatial arrangement for interactions with DNA (Saikumar *et al*., 1990). The third helix contains the recognition domain that binds the DNA in its major groove (Ogata *et al*., 1994). DUO1 is distinguished from other R2R3 MYB transcription factors by the presence of a supernumerary Lysine residue (K66) which is not present in other R2R3 MYB transcription factors (Rotman *et al*., 2005). The K66 residue has been shown to be functionally important for correct transactivation of DUO1 targets in *A. thaliana* (Borg, 2010).

To confirm the role of DUO1 as a putative transcription factor and understand how the loss of DUO1 influences the development of the male germ-line in a *duo1* mutant, further research focused on the identification of genes transcriptionally regulated by DUO1. These DUO1 activated target (DAT) genes would help elucidate the role DUO1 played in the development of the germ-line cells.

Male germ-line specific reporter genes were tested to see if DUO1 activated them through promoter reporter constructs being transformed into *DUO1/duo1* plants (Brownfield *et al*., 2009). Promoter sequences of the germ-line specific genes drive expression of a nuclear localisation signal from histone H2B connected to the open reading frame of a fluorescent protein (GFP) for visualisation. Two genes *GCS1* (Mori *et al*., 2006), *GEX2* (Engel *et* al., 2005; Mori *et al*., 2014) which code for proteins essential for fertilization and another gene *MGH3* (Okada *et al*., 2005) a male gamete specific histone H3, were shown to be regulated by DUO1 with germ-line specific expression observed in all three wild type pollen grains, however there was an absence of GFP expression from *duo1* pollen grains, showing functional DUO1 is necessary for expression of these genes (Brownfield *et al*., 2009).

Work by Brownfield *et al*., (2009) also showed that the cell cycle regulator cyclinB1;1 (CYCB1;1) which is a key regulator of the G2/M transition in the cell cycle (Colón-Carmona *et al*., 1999) requires DUO1 for germ-line expression explaining why *duo1* generative-like cells fail to divide. Additionally when *CYCB1;1* is expressed under

the control of the *DUO1* promoter (*pDUO1:CYCB1;1*) *duo1* generative-like cells complete PMII. These rescued sperm cells however were unable to fertilise suggesting sperm cell function was affected in the *duo1* mutant. This was confirmed by verifying that DAT genes *GSC1* and *MGH3* which are normally expressed by DUO1 were not active in the *pDUO1:CYCB1;1* complemented *DUO1/duo1* line.

Following the discovery of these initial targets a microarray approach was used to identify further DUO1 target genes. This involved analysing data from ectopic expression of DUO1 in *A. thaliana* seedlings (Borg *et al.*, 2011). Microarray data from the seedlings were analysed for potential DAT genes by selectively filtering genes that are differential expressed in the ectopic DUO1 seedlings, detecting 63 putative DAT genes. Fifteen of these DAT gene promoters have been confirmed to have shown sperm cell specific expression in pollen through promoter reporters (Borg *et al*., 2011). The promoters of the DAT genes were also used to identify a potential DUO1 binding site. The DNA recognition helix of MYB's typically binds a core DNA motif of ACC (Tanikawa *et al*., 1993) and a core binding motif of AACCG was present in the DAT gene promoters (Borg *et al*., 2011). It was observed that the proximal distribution of MYB binding sites in DAT gene promoters showed a strong bias for within the first 250 bp upstream of the start codon (Borg *et al*., 2011).

The sequence recognition of DUO1 binding to AACCG was confirmed in electrophoretic mobility shift assays (EMSA) with recombinant DUO1 binding the MYB site of the promoter of the DAT gene *MGH3*. However when the MYB bind site was mutagenized to the sequence TCATGA, DUO1 was unable to bind, thus showing specificity to the AACCG MYB binding site (Borg *et al*., 2011).

DUO1 has also been shown to regulate two functionally redundant proteins; DUO1-Activated Zinc Finger I (DAZ1) and DUO1-Activated Zinc Finger 2 (DAZ2) (Borg, 2011). Ectopically expressed DUO1 in seedlings showed activated *DAZ1* and *DAZ2* (Borg *et al*., 2011). The *duo1* bi-cellular phenotype has been shown to be rescued by a construct containing the *DUO1* promoter driving expression of DAZ1 (*pDUO1:DAZ1*) showing a key role in *DAZ* genes mediating mitotic entry (Borg, 2011). Double knockouts of *DAZ1* and *DAZ2* (*daz1:daz2*) have also shown that they are essential for germ cell division producing bi-cellular pollen. It was also shown that

a construct containing the *DUO1* promoter driving the expression of the *DAZ* gene rescued the *daz1:daz2* phenotype and allowed regular cellular division (Borg *et al*., 2014). This elucidates the role DUO1 plays in controlling an essential gene regulatory network that is necessary for the regulatory hierarchy of mitotic entry of PMII in male germ-line cells (Borg *et al*., 2014).

DUO1 is thus a primary member in the gene regulatory network that controls the male generative cell development in pollen, however what regulates DUO1 itself is still unclear and is a key question in understanding male germ-line specification.

1.6. Regulation of DUO1

1.6.1 Proposed mechanisms of DUO1 regulation

While the importance of DUO1 in development of the male germ-line is clear, less is known about how the expression of *DUO1* is regulated. Analysis of a 1.2 kb upstream fragment of the *A. thaliana DUO1* promoter fused to cDNA of a fluorescent reporter protein showed sperm cell specific expression in pollen (Rotman *et al.*, 2005). Brownfield *et al.,* 2009 showed that the 1.2 kb *DUO1* promoter reporter construct had the same germ-line specific expression pattern as a translational fusion construct of promDUO1:DUO1-mRFP. This shows that *DUO1* is transcriptionally regulated as the promoter alone is sufficient to drive germline specific expression of a reporter construct. Analysis by confocal laser scanning microscopy (CLSM) also showed that expression of the 1.2 kb *DUO1* promoter reporter is first seen specifically in the male germ-line soon after the asymmetric division before the generative cell is engulfed (Brownfield *et al.,* 2009), showing *DUO1* is early in male germ-line development.

A proposed mechanism of *DUO1* regulation based on studies in Lily was first suggested by Haerizadeh *et al.,* 2006. This mechanism proposed that flowering plant male germ-line specific genes are maintained in a repressed state in non–male germ-line cells via negative transcriptional regulation mediated by a protein called “Germ-line Restrictive Silencing Factor” (GRSF). Haerizadeh *et al.,* 2006 identified that in Lily a male germ-line specific gene called “Lily Generative Cell 1” (LGC1) (Xu *et al*., 1999) was regulated through this silencing mechanism. Regulation of

LGC1 was analysed by a 0.8 kb promoter deletion series which showed a silencing mechanism is responsible for controlling its expression (Singh *et al*., 2003). GRSF was later identified to bind a CRE in the *LGC1* promoter resulting in silencing in non-germ-line cells (Haerizadeh *et al*., 2006). A sequence similar to the CRE to where GRSF bound was also identified to be present in the *DUO1* promoter in *A. thaliana* (Haerizadeh *et al*., 2006).

However, in *A. thaliana* removal of this sequence via site directed mutagenesis or truncations by a 5′ deletion series of the *DUO1* promoter does not result in deregulation with sperm cell specific expression still seen (Brownfield *et al.*, 2009). A 5′ deletion series of the *DUO1* promoter (Figure 1.3) has shown that *DUO1* promoter fragments with and without GRSF binding sites showed similar germ-line specific expression indicating *DUO1* expression is not dependent on a silencing mechanism such as *LGC1* in Lily. The deletion series also showed a minimal promoter fragment of 155 bp upstream of the ATG was needed to drive germ-line specific expression (Figure 1.3) demonstrating a positive transcriptional activation region within this fragment (Brownfield *et al*., 2009).

A recent proposal of a transcriptional regulation mechanism of DUO1 was by Zheng *et al*., (2014) in which an AT-Rich Interacting Domain protein (ARID1) was been shown to interact with the *DUO1* promoter by chromatin immunoprecipitation (Zheng *et al*., 2014). ARID1 was reported to bind -600 to -300 up stream of the ATG of *DUO1* and suggested to be a positive transcriptional factor. The disruption of ARID1 in *arid1-1* plants results in a reduced level of DUO1, however *DUO1* and its targets are still expressed. ARID1 is also expressed in microspores and vegetative cells (Zheng *et al.*, 2014) suggesting it is not the sole regulatory protein controlling the male germ-line formation. Rather the reduction of DUO1 in *arid1-1* may be due to the alterations in the H3K9ac mark at the *DUO1* promoter as ARID1 has been shown to associate with histone acetylation resulting in chromatin modification increasing expression of *DUO1*. This means ARID1 is unlikely to be the key regulator in expression of DUO1. Instead ARID1 may help enhance expression of *DUO1* through the active euchromatin marks of H3K9ac due to ARID1 binding to the *DUO1* promoter

rather than being the direct transcriptional activator binding to the proximal promoter of *DUO1*

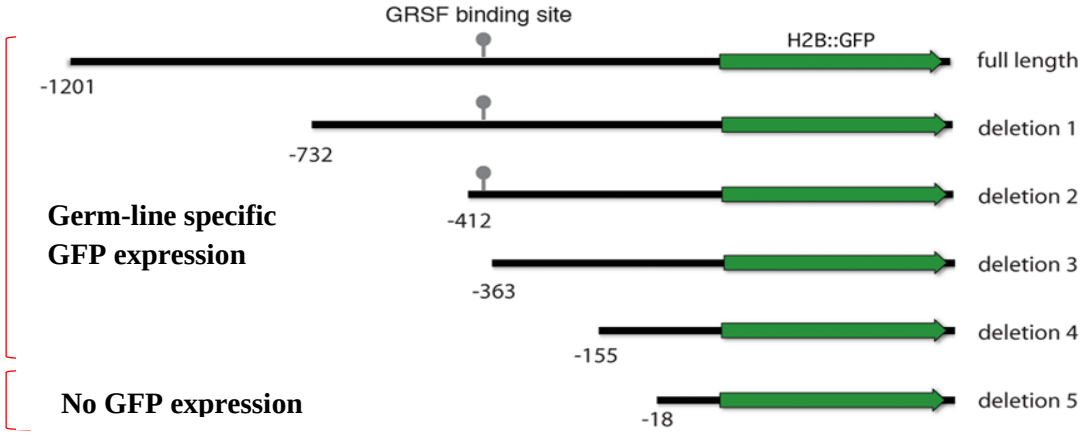


Figure 1.3: Schematic representation of the 5´ deletion series used to analyse *A.thaliana* DUO1 promoter activity. 6 deletions of the *DUO1* promoter were made; 3 which include the GRSF binding site and 3 shorter fragments without it. Deletions 1-4 showed germ-line specific expression of GFP which was similar to the full length construct. The minimal region needed to drive germ-line specific expression was the deletion 4 (-155 bp) reporter construct and no expression was seen in deletion 5. Figure adapted from Brownfield *et al*., 2009.

1.6.2 Conserved *cis*-regulatory elements determine DUO1 germ-line specific expression

Further analysis of the *DUO1* promoter in *A. thaliana* by Peters *et al*., 2016 has shown that various lengths of the 5´ deletions of the *DUO1* promoter linked to reporter constructs show reduced levels of expression of reporter helping identify regions/CREs in the promoter needed for expression of *DUO1*. Two sections of the promoter display independent activity in wild type and *duo1* pollen mutants and contribute to DUO1 expression. These sections have been termed Regulatory region of DUO1 1 (*ROD1*; position -153:-61 upstream from the ATG site) and Regulatory region of DUO1 2 (*ROD2*; position -426:-273 upstream from the ATG site). *ROD1* was established as the minimal region needed to drive male germ-line specific expression (Peters *et al*., 2016) through analysis of various fragments of *A. thaliana* DUO1 promoter reporter constructs.

A comparative alignment of *ROD1* within the *DUO1* promoter of various dicotyledonous plants revealed three short highly conserved sequences that could be potential CRE's (Figure 1.4). The putative conserved CRE motifs were an AG rich region GAGARAAA at the distal end, followed by three adjacent copies of a repeated motif with a GTGG core with some similar nucleotides surrounding the core giving a consensus sequence of DNGTGGV. Two copies of a YACCYGY motif then follow with at least one motif having the consensus sequence YAACCGY repeat (Figure 1.4). Mutating these sites in a 198 bp *ROD1* promoter reporter construct has identified that the GTGG core sites are essential for male germ-line specific DUO1 expression, with no expression seen when these sites are altered. Altering just the GAGA core sites resulted in male germ-line specific expression and altering the YAACCGY motifs resulted in a reduced germ-line specific expression level of the reporter (Zohrab, 2015).

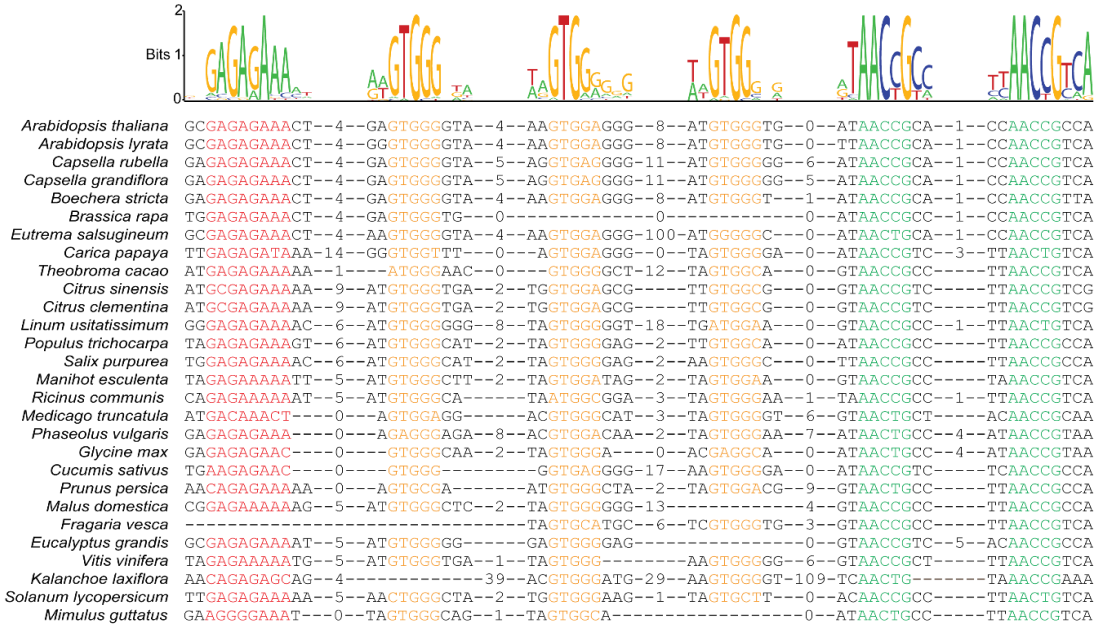


Figure 1.4: ROD1 comparative genomic alignment from dicotyledonous angiosperms. 85 bp region of *A. thaliana* ROD1 is aligned against 28 ROD1 regions from other dicotyledonous angiosperms with conserved motifs highlighted and PSSM for sequences about. GAGARAAA motif is shown in red. DNGTGGV motif is shown in yellow. YAACCYGY motif is shown in green. (Figure adapted from Peters *et al*., 2016)

1.6.3 Auto-activation of DUO1

The DUO1 promoter shows a potential mechanism of auto-activation by binding to at least one of the YAACYGY CRE's within *ROD1* (Figure 1.4) as the DUO1 protein has the consensus binding site AACCG (Borg *et al.,* 2011). The experiment described above with altered YAACCGY motifs showed reduced levels of reporter consistent with an auto-activation role for these sites. Further testing of promoter reporter constructs containing four YAACYGY repeat CRE's in total analysed by transformation of a heterozygous *duo1* line. From this wildtype lines with the reporter showed male germ-line specific fluorescence, while *duo1* pollen grains with singular generative-like cells rarely had expression. This shows that DUO1 is necessary for high activities of this construct likely due to auto-activation by binding to the YAACYGY CRE's. (Peters *et al.*, 2016). An interaction has also been shown in a yeast-one-hybrid assay with DUO1 being used as the activation domain and the DUO1 promoter used as the bait, there is an interaction (Peters and Brownfield, unpublished).

Additionally, it has been shown by luciferase reporter assays that *A. thaliana* DUO1 is able to bind to its own promoter with a 91 bp promoter fragment containing just a single AACCG motif drives expression of luciferase in the presence of constitutively expressed DUO1 (Aidley, 2011). A construct with the AACCG CRE altered to CATGA showed reduced levels of luciferase. This shows the YAACYGY CRE may not be necessary for switching on germ-line specific expression, but rather enhancing the expression of DUO1. So a role in auto-activation of DUO1 is highly likely and shows the importance that transcriptional regulation has on the expression of *DUO1*.

1.7 DUO1 in other flowering plant species

DUO1 can be used as a useful tool to help understand the processes involved in male germ-line production. Much is already known on the regulation and downstream effects of *DUO1* in *A. thaliana* and conserved mechanisms may be present in flowering plants. Studies in the laboratory of Professor David Twell (University of Leicester) have identified DUO1 orthologue's in Lily, *Oryza sativa* (Rice), *Solanum*

lycopersicum (Tomato) and complementation analysis has shown a conserved germ-line regulatory function among all species (Twell *et al*., unpublished).

There has been limited work on the regulation of DUO1 from these species. Conserved germ-line specific expression has been seen in pollen development from the full length *DUO1* promoter of tomato (Twell *et al*., unpublished). But the *DUO1* promoter from rice has shown expression differences in early development with reporter construct expression seen in late microspores and in vegetative cells in early bi-cellular pollen along with expression in sperms of mature pollen. This shows there may be potential regulatory differences of *DUO1* between monocots (rice) and eudicots (*A. thaliana* and tomato).

However no work has focused on the conservation or function of *ROD1* from these species. The main focus of this thesis is investigating the function and regulation of *DUO1* in another evolutionarily diverse plant species, the model legume *Medicago truncatula* and investigate functional conservation of *ROD1* in *M. truncatula* and *O. sativa*.

1.7.1 *M. truncatula* as a model legume

The legume *M. truncatula* has emerged as a model organism for legume genetics and genomics (Cook, 1999). Legumes are known for their nitrogen fixation abilities and high protein content and comprise some of the most important agricultural crops worldwide with notable species such as soybean (*Glycine max*) and pea (*Pisum sativum*) included. However, the size and complexity of these genomes has slowed down the genetic characterisation of these various crops. *M. truncatula* however, has a small diploid genome of 5×10^8 bp, is relatively easy to transform, is self-fertile with abundant seed production and has a rapid generation time (Barker *et al.,* 1990) making *M. truncatula* an ideal model organism.

Legumes are evolutionarily distinct from the commonly used model plant species such as *A. thaliana* which belongs to the Brassicaceae family. The use of *M. truncatula* will thus help enable research to address difficult questions in other crop species with

complicated genomes along with looking at the potentially conserved mechanisms of plant biology within evolutionarily distinct species.

1.8 Aims and objectives

This project seeks to investigate the function and regulation of *DUO1* from the legume *M. truncatula* (*MtrDUO1*) and investigate conservation of *ROD1* from the *M. truncatula* and *O. sativa DUO1* promoters. Identifying these mechanisms in *M. truncatula* may help to broaden the scope of how germ-line specific genes such as *DUO1* are regulated and their role of controlling the cell cycle and differentiation into mature sperm cells, ultimately increasing our understanding of plant reproduction.

To do this, I will first confirm if *MtrDUO1* and *AthDUO1* are functionally equivalent by complementation testing of the *A. thaliana duo1* mutant with MtrDUO1. I then will verify that the *MtrDUO1* promoter provides germ-line specific expression and analyse a 5′ deletion series of shorter promoter fragments to determine important regions needed for expression. This will be done by stably transforming *A. thaliana* with the various length promoter regions of *MtrDUO1* driving expression of a GFP reporter. This allowed fluorescence microscopy to observe if expression is germ-line specific and quantification of GFP levels.

A reporter construct containing four repeats of the conserved *ROD1* domain in the *MtrDUO1* and *OsaDUO1* promoters will be analysed to see if *ROD1* promotes germ-line specific expression. This will be done using the same GFP technique as the 5′ deletion series.

Auto-regulation of *DUO1* will be analysed by site directed mutagenesis of the putative DUO1 binding sites within fragments of the *DUO1* promoter from *AthDUO1*, *MtrDUO1* and *OsaDUO1* and assessing the ability of DUO1 to activate expression in a luciferase assay. Site-directed mutagenesis of the MYB sites will test the interactions between the DUO1 protein and position of the MYB sites in the *DUO1* promoter by using the promoter to drive expression of luciferase to analyse by a transient dual-luciferase assay system.

Chapter 2: Materials and Methods

2.1 Plant Growth and Tissue Culture

2.1.1 Plant material

A. thaliana plants were from the Columbia-0 background and heterozygous for *duo1-4* (Borg *et al*., 2014). *M. truncatula* DNA used was extracted from the Jester ecotype. *Nicotiana benthamiana* plant leaf tissue was used to conduct dual luciferase assays.

2.1.2 Seed germination and plant growth conditions

A. thaliana seeds were germinated either directly on soil consisting of compost with added vermiculite or on ½ Murashige and Skoog salt agar media consisting of 0.44% (w/v) of MS salts (Duchefa Biochemie) in milliQ water. The pH was adjusted to 5.8 using 1 M NaOH prior to addition of 0.75% (w/v) bacto-agar. Media was then autoclaved at 120 °C and 15 psi for 20 minutes. Half MS media was then allowed to cool to 55 °C before the appropriate antibiotic (e.g. Kanamycin) was added before pouring into Petri dishes. Seeds were surface sterilized by vapour-phase sterilization before plating. 1 mL 36% hydrochloric acid (HCl) was added to 3 mL of bleach and placed in a sealable container with a rack holding open 1.5 mL micro-centrifuge tubes with 3-4 mm of *A.thaliana* seeds to be sterilized. Seeds were left to sterilize for 6-8 hours before plating on to ½ MS agar plates with kanamycin (50 µg/mL).

Seeds grown on soil and plates were stratified at 4°C for 72 hours to synchronise germination. Seedlings resistant to antibiotics on MS agar plates were then transferred to soil to allow further growth. Selection of transformed seedlings grown on soil was selected by spraying with the herbicide Basta® three times, 3 - 4 days apart. Selected seedlings/plants were then grown in long day conditions of 16 hours light and 8 hours dark at 20°C.

2.1.3 Transformation of *A. thaliana* with *Agrobacterium tumefaciens*

A. thaliana plants were transformed with a modified floral dip method (Martinez-Trujillo *et al*., 2004). Any open flowers and formed siliques on the plants were removed prior to transformation with *Agrobacterium tumefaciens*. Transformed *A.tumefaciens* (Section 2.3.6) was grown for 2-3 days on LB agar plates containing appropriate antibiotics. Colonies were then scraped off the plates and suspended in infiltration medium containing 0.04% (w/v) MS salts and 0.5% (w/v) sucrose and the OD_{600} was adjusted to 0.6. Directly before floral dropping of infiltration medium containing 2.5% (v/v) silwet L-77 (Lehle seeds) was added to the suspended *A.tumefaciens* immediately prior to floral dip. The infiltration solution was pipetted in small drops onto inflorescences containing unopened floral buds. This process was then repeated twice to increase transformation efficiency; first 3-4 days later, then 3-4 days after the second floral dip. Following floral dipping plants were left to set seed and siliques were harvested when dried. Seeds were then grown as per section 2.1.2.

2.1.4 Small scale plant genomic DNA isolation for PCR

Genomic DNA was isolated using a protocol based on Edwards method for preparation of plant genomic DNA for PCR analysis (Edward *et al.*, 1991). One or two (1 cm^2) rosette or cauline leaves were placed in microcentrifuge tubes containing approximately 200 acid washed beads (425-600 µm) and flash frozen in liquid nitrogen. The leaves were macerated with a bead beater at 4°C. 500 µL of Edwards buffer (1 M Tris-HCl pH 7.5 20% (v/v), 5 M NaCl 5% (v/v), 0.5 M EDTA pH 8 5% (v/v), 10% SDS 5% (v/v) and 65% (v/v) milliQ water) was added and vortexed to homogenize. Samples were then centrifuged at 17,000 *g* for 5 minutes and 400 µL of supernatant was transferred to a new tube, 400 µL of isopropanol was added and vortexed briefly (3 sec) before being centrifuged at 17,000 *g* for 5 minutes. The supernatant was removed and the pellet washed with 500 µL of 70% ethanol. The pellet was allowed to dry completely before being dissolved in 100 µL TE buffer (10 mM Tri-HCl pH 7.5, 1 mM EDTA) and left to sit for 10 minutes at room temperature then vortexed briefly. Samples were used immediately for PCR or stored at -20°C.

2.2 Multi-site Gateway® Cloning

Multi-site Gateway® Cloning (Thermo Fisher Scientific) technology was used to generate expression constructs. Gateway® technology is based on the ability of lambda (λ) bacteriophage to site specifically integrate into *Escherichia coli* (*E.coli*) (Landy, 1989). The system uses recombination between site specific attachment (*att*) sites; *att*B on *E.coli* DNA and *att*P on the Lambda chromosome to produce *att*L and *att*R sites flanking the DNA insertion (Landy, 1989). Multi-site Gateway® Cloning uses these *att* sites and isolated enzymes to recombine DNA *in vitro* (Hartley *et al.*, 2000). PCR products or DNA with *att*B sites can be recombined into entry vectors with corresponding *att*P sites via BP reaction to form an entry clone. BP reactions form *att*L sites around the inserted PCR or DNA, *att*L sites can then be recombined with *att*R sites of a destination vector via LR reaction to form an expression vector. Specific *att* sites can only recombine with each other and using multiple *att* sites allows Gateway© Cloning to recombine up to three vectors (Entry clone vectors and a destination vector) in a predefined order, orientation and reading frame (Cheo *et al.*, 2004).

2.2.1 PCR amplification of DNA for Gateway® Cloning

Primer pairs were designed to flank a genomic DNA or gBlock sequence of interest to be cloned and are shown in Table 2.1. The primers are designed with *att*B or TOPO sites (A-over hang) at the 5′ and 3′ ends of the PCR product (Table 2.1). The *att*B sites on PCR products allow recombination of fragments via BP reaction's into a Donor vector. PCR reactions used the high fidelity polymerases Phusion (New England Biolabs) (Table 2.2) or Kapa HiFi (Kapa Biosystems) (Table 2.2) to eliminate the incorporation of incorrect nucleotides. Extension time for different PCR's was based on PCR fragment length with a rate of 30 seconds/kb. Touchdown PCR conditions (Table 2.3) were used with reactions that had primers with low annealing temperatures due to high A/T content.

Table 2.1: Primers used in Gateway® and traditional restriction enzyme cloning. *att* sites are in italics. Restriction enzyme sites are colour coded; Pst1, HindIII, BamHI, SacI

Primer description	Primer Sequence
*Ath promDUO*1 F *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*ACGTCC GAAGTTTCCCTCTTGG
Ath promDUO1 R *att*B1R	*GGGGACTGCTTTTTTGTACAAACTTG*CGCTAA TCGATCTCTCTCTCG
Mtr promDUO1 Reverse *att*B1R	*GGGGACTGCTTTTTTGTACAAACTTG*TCTTTTT CTACCTATAATAATATTATTCA
Mtr DUO1 cDNA F *att*B1F	*GGGGACAAGTTTGTACAAAAAAGC*AGGCTCA ATGCATGGGAAAAAAGATC
Mtr DUO1 cDNA R stop *att*B2R	*GGGGACCACTTTGTACAAGAAAGCTGGGT*TCA GAGCTTAGAGAAATTTGAAGG
Mtr DUO1 cDNA R ns *att*B2R	*GGGGACCACTTTGTACAAGAAAGCTGGGT*AG AGCTTAGAGAAATTTGAAGGTG
Mtr 726 promDUO1 F *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*CGTAAC GGAATATAAGTATGAGTG
Mtr 511promDUO1 Foward *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*CAAGGT GAATAAGTGGGTGTCC
Mtr 243promDUO1 F *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*TGACAA ACTAGTGGAGGACGTG
Mtr 167promDUO1 F *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*CATGGA ATAATGTATGTGTGC
Mtr promDUO1 repeat1 *Att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*TGACAA ACTAGTGGAGGACGTG
Mtr promDUO1 repeat1 PstI R	GCTGGCCTGCAGGCACACATACATTATTCC ATG
Mtr promDUO1 repeat2 PstI F	GCTGGCCTGCAGTGACAAACTAGTGGAGGA CGTG
Mtr promDUO1 repeat2 HindIII R	GCTGGCAAGCTTGCACACATACATTATTCC ATG
Mtr promDUO1 repeat3 HindIII F	GCTGGCAAGCTTTGACAAACTAGTGGAGGA CGTG
Mtr promDUO1 repeat3 BamHI R	GCTGGCGGATCCGCACACATACATTATTCC ATG
Mtr promDUO1 repeat4 BamHI F	GCTGGCGGATCCTGACAAACTAGTGGAGGA CGTG
Mtr promDUO1 repeat4 *att*B1R	*GGGGACTGCTTTTTTGTACAAACTTG*GCACAC ATACATTATTCCATG
Osa tetramer F1 repeat 1 *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*GAGCGT ACGCGCTAGAATTC
Osa tetramer R4 repeat1 *att*B1R	*GGGGACTGCTTTTTTGTACAAACTTG* GACGGTCGGGTTTGATTCGC
Osa tetramer R1 repeat 2 Pst1	GCTGGCCTGCAGGACGGTCGGGTTTGATTC GC

Osa tetramer F2 repeat 2 Pst1	GCTGGCCTGCAGGAGCGTACGCGCTAGAATTC
Osa tetramer R2 repeat 3 HindIII	GCTGGCAAGCTTGACGGTCGGGTTTGATTCGC
Osa tetramer F3 repeat 3 HindIII	GCTGGCAAGCTTGAGCGTACGCGCTAGAATTC
Osa tetramer R3 repeat 4 BamHI	GCTGGCGGATCCGACGGTCGGGTTTGATTCGC
Osa tetramer F4 repeat 4 BamHI	GCTGGCGGATCCGAGCGTACGCGCTAGAATTC
Osa tetramer ATG F1 repeat 1 *att*B4F	*GGGGACAACTTTGTATAGAAAAGTTG*CCCTGTGGCTGGGCTTTTG
Osa tetramer ATG R4 repeat1 *att*B1R	*GGGGACTGCTTTTTTGTACAAACTTG*TTGCCCTCACAACCGCC
Osa tetramer ATG R1 repeat 2 Pst1	GCTGGCCTGCAGTTGCCCTCACAACCGCC
Osa tetramer ATG F2 repeat 2 Pst1	GCTGGCCTGCAGCCCTGTGGCTGGGCTTTTG
Osa tetramer ATG R2 repeat 3 HindIII	GCTGGCAAGCTTTTGCCCTCACAACCGCC
Osa tetramer ATG F3 repeat 3 HindIII	GCTGGCAAGCTTCCCTGTGGCTGGGCTTTTG
Osa tetramer ATG R3 repeat 4 BamHI	GCTGGCGGATCCTTGCCCTCACAACCGCC
Osa tetramer ATG F4 repeat 4 BamHI	GCTGGCGGATCCCCCTGTGGCTGGGCTTTTG
Ath ROD1 Forward TOPO	AGAGAAACTTGGGGAGTGGGG
Ath ROD1 Reverse TOPO	CGCTAATCGATCTCTCTCTCG
Mtr ROD1 Forward TOPO	TGACAAACTAGTGGAGGACGTG
Mtr ROD1 Reverse TOPO	TCTTTTTCTACCTATAATAATATTATTCA
Osa ROD1 Forward TOPO	GAGCGTACGCGCTAGAATTC
Osa ROD1 Reverse TOPO	ATTGCCCTCACAACCGCC
Prom Ath DUO1 full length F HindIII	GCTAAGCTTCGCTAATCGATCTCTCTCTCG
Prom Ath DUO1 full length R SacII	GCTCCGCGGACGTCCGAAGTTTCCCTCTT
Ath DUO1 4x ROD1 F SacII	GCTCCGCGGGGGCCCGAGCGGATAAC
Ath DUO1 4x ROD1 R SacII	GCTCCGCGGCGACTAGCTTCAGCGTGTCC
MGH3 F HindII	GCTAAGCTTTTGAGCAGATCGGAAGAGACG
MGH3 R SacII	GCTCCGCGGTTCTTCGAGAGAACGATGATG

Table 2.2: PCR reaction recipes for Phusion (Left) and KAPA HiFi (Right) for use with Gateway® Cloning

Phusion reaction component	**Volume (µL)**	**KAPA HiFi reaction component**	**Volume (µL)**
Water	30.4	Water	33.5
5x Phusion buffer	10	5x KAPA buffer	10
dNTP (10 mM)	1	dNTP (10 mM)	1.5
MgCl (5 mM)	2		
Forward Primer (10 mM)	2	Forward Primer (10 mM)	1.5
Reverse Primer (10 mM)	2	Reverse Primer (10 mM)	1.5
DNA Template	2	DNA template	1
Phusion Polymerase	0.6	KAPA HiFI Polymerase	1
Total	**50 µl**	**Total**	**50 µl**

Table 2.3: PCR Conditions for Phusion and KAPA HiFi for use in Gateway® cloning. Extension time of 30 seconds per kb was used depending on fragment length. T* refers to starting annealing temperature which was dropped 1°C for 15 cycles in touchdown PCR

Phusion		**KAPA HiFi Touchdown**	
Temperature (°C)	**Time (Seconds)**	**Temperature (°C)**	**Time (Seconds)**
98	120	95	120
98	20	94	20
55	20 **x30**	T*	15 **x15**
72	30 / kb	72	30 / kb
72	30 / kb x2	94	20 **x15**
		58	15
		72	30 / kb
		72	30 / kb x2

PCR products were checked for size on 1% or 2% agarose gels. The PCR product was then excised from the gel and purified with a High Pure PCR Product Purification Kit (Roche) according to the manufacturer's instructions.

2.2.2 BP reaction to generate Gateway® Entry clones

PCR products with *att*B sites were used in a BP reaction. These were carried out using the enzyme BP Clonase II (Life technologies) to insert the DNA fragments into suitable DONR vectors to form Entry vectors. The BP reaction (Table 2.4) was incubated overnight at 22°C and 0.5 µL of proteinase K (Life technologies) was added and incubated at 37°C to terminate the reaction. DONR vectors used were pDONRP4-1R (Life Technologies) for first part Entry Clones (5′ elements e.g. promoters), pDONR221 for second part clones (e.g. gene of interest) and pDONRP2R-P3 for third part clones (3′ elements e.g. fluorescent tags) (Table 2.5). Three µL of the BP reaction mix are then transformed into DH5α *E.coli* as per section 2.4.4

Table 2.4: BP reaction recipe for creating Gateway® Entry Clones

BP reaction	**Volume (µL)**
PCR product	3.5
pDONR vector	0.5
BP Clonase II	1
Total	**5**

Table 2.5: Entry vectors used to create Entry Clones via BP reaction

Entry Vector	**Resistance***	**Forward *att* Site**	**Reverse *att* Site**
pDONRP4-P1R	Kan50	attB4	attB1r
pDONR221	Kan50	attB1	attB2
pDONRP2R-P3	Kan50	attB2r	attB3
pCR8/GW/TOPO	Spec100	N/A	N/A

*Kan50 = Kanamycin concentration 50 µg/mL. Spec100 = Streptomycin 100 µg/mL

2.2.3 Construction of promoter sequence tetramers

Four copies of a promoter sequence were amplified individually to have specific restriction sites on the ends of the amplicons. The amplicons then underwent restriction digests with specific enzymes (BamH1, HindIII, Pst1) and purified to remove enzymes and unwanted cut DNA. The four different amplicons were then ligated together using T4 ligase (Thermofisher scientific) and amplified again using primers flanking the ligated amplicon with *attB* primers. The amplified DNA sequence was then ready to be used in a BP reaction.

2.2.4. gBlock gene fragments for TOPO® Entry Clones

gBlock gene fragments are double stranded synthetic DNA strands produced by Integrated DNA Technologies. These were designed and purchased to incorporate site directed changes in sequences of interest. gBlock gene fragments were re-suspended in milliQ water to a final concentration of 20 ng/µL.

The gBlock Gene fragments used contained no *att* sites so TOPO® TA cloning was used to recombine the fragments into an Entry Clone that could be used in Gateway® cloning. An A-overhang was added to the gBlock gene fragments using *Taq* polymerase (Table 2.6) before being used in a TOPO® reaction (Table 2.7) to generate Entry Clones (Table 2.8).

Table 2.6: A-Overhang reaction used on gBlock gene fragments

A – Overhang reaction	Volume (µl)
10x PCR Buffer	1
MgCl (5 mM)	0.3
dNTP (10 mM)	0.5
Water	2.9
gBlock gene fragment (20ng/µl)	5
Taq Polymerase	0.3
Total	**10**

2.2.5 TOPO® TA Cloning to produce Entry Clones for Gateway®

TOPO® cloning is based on A-overhang recombination of a DNA fragment into a linearized entry vector with T-overhangs through the enzyme topoisomerase I. The entry vector pCR8/GW/TOPO (Table 2.5) has suitable *att* sites for further Gateway® Cloning to create expression constructs through Multisite Gateway® LR reactions. The TOPO reaction (Table 2.7) was incubated for 5 minutes then transformed into DH5α *E.coli* as per section 2.4.4.

Table 2.7: TOPO® reaction recipe

TOPO® reaction	Volume (µL)
TOPO vector	0.5
PCR/DNA product	1
Salt solution (1.2 M NaCl 0.06 M $MgCl_2$)	1
Water	2.5
Total	**5**

Table 2.8: Entry clones used in Multisite Gateway® cloning. First column is the DNA fragment used to clone into the Entry vector which is either into first, second or third part entry vectors. Entry vector is the vector cloned into. Antibiotic is the resistance of the entry clone. Origin refers to who cloned the construct: Johnny Casey; JC, Megan Styles; MFS, Ben Peters; BP, Lynette Brownfield, LB.

First part PCR or DNA	**Entry vector**	**Antibiotic**	**Origin**
promAth1244 bp	pDONRP4-P1R	Kan50	MFS
promMtr171 bp	pDONRP4-P1R	Kan50	JC
promMtr247 bp	pDONRP4-P1R	Kan50	JC
promMtr517 bp	pDONRP4-P1R	Kan50	MFS
promMtr726 bp	pDONRP4-P1R	Kan50	MFS
promMtrROD1 tetramer	pDONRP4-P1R	Kan50	JC
promOsa ROD1 Tetramer ATG	pDONRP4-P1R	Kan50	JC
promOsaROD1 Tetramer	pDONRP4-P1R	Kan50	JC
promAthDUO1 198 bp	pCR8/GW/TOPO	Spec100	JC
promAthDUO1-MYB1	pCR8/GW/TOPO	Spec100	JC
promAthDUO1-MYB2	pCR8/GW/TOPO	Spec100	JC
promAthDUO1-MYB1-2	pCR8/GW/TOPO	Spec100	JC
promMtrDUO1	pCR8/GW/TOPO	Spec100	JC
promMtrDUO1-MYB1	pCR8/GW/TOPO	Spec100	JC
promMtrDUO1-MYB2	pCR8/GW/TOPO	Spec100	JC
promMtrDUO1-MYB1-2	pCR8/GW/TOPO	Spec100	JC
promOsaDUO1	pCR8/GW/TOPO	Spec100	JC
promOsaDUO1-MYB1	pCR8/GW/TOPO	Spec100	JC
promOsaDUO1-MYB2	pCR8/GW/TOPO	Spec100	JC
promOsaDUO1-MYB1-2	pCR8/GW/TOPO	Spec100	JC
Second part PCR or DNA inserts	**Entry Vector**	**Antibiotic**	**Origin**
AthDUO1 cDNA	pDONR221	Kan50	BP
Ath(m)DUO1 cDNA	pDONR221	Kan50	JC
Mtr DUO1 STOP gDNA	pDONR221	Kan50	MFS
Mtr DUO1 No Stop gDNA	pDONR221	Kan50	MFS
H2B	pDONR221	Kan50	BP

Min35S:H2B	pDONR221	Kan50	BP
NLS:GFP	pENTR221	Kan50	LB
Third part PCR or DNA inserts	**Entry Vector**	**Antibiotic**	**Origin**
eGFP	pENTRP2RP3	Kan50	LB

2.2.6 Generation of Gateway® expression vectors by LR reaction

Entry Clones were recombined with a destination vector to create an expression vector through the LR reaction with LR clonase II (Life technologies). The LR reaction can recombine one (single-site), two (two-site) or three (three-site) Entry Clones with a destination vector. Two-site LR and three-site reactions allow the recombination of up to three Entry Clone DNA fragments flanked by different *att* sites that allow recombination into a Destination vector in a predefined order and orientation to create an expression vector (Cheo *et al.*, 2004). The reaction mixture's for the various LR reactions are outlined below in Table 2.9. The reactions were incubated at 22°C overnight then 1.0 µL of proteinase K (Life technologies) was added and incubated at 37°C for 10 minutes to stop the reaction. Table 2.10 shows all expression vectors created by Gateway Cloning.

Table 2.9: LR reaction recipes for making single-site, two-site and three-site Gateway® expression vectors

Single-site LR reaction	**Volume/concentration**
pDONR221	125 ng
Destination vector	75 ng
LR Clonase II	1.0 µL
dH_2O	Up to 5.0 µL
Total	**5 µL**
Two-site LR reaction	
pDONRP4-P1R entry clone	5 fmol
pDONR221 entry clone	5 fmol
Destination vector	10 fmol
LR Clonase II plus	1.0 µL
dH_2O	Up to 5 µL
Total	**5 µL**
Three-site LR reaction	
pDONRP4-P1R entry clone	5 fmol
pDONR221 entry clone	5 fmol
pDONRP2R-P3 entry clone	5 fmol
Destination vector	10 fmol
LR clonase II plus	1.0
dH_2O	Up to 5 µl
Total	**5 µl**

To calculate the ng of plasmid DNA needed the following formula was used to achieve the desired f mols: ng needed = desired f mol x size of vector (bp) x (660 x 10-6).

Table 2.10: Expression vectors created by Gateway ® cloning and Restriction enzyme cloning

Destination vector	**First part**	**Second part**	**Third part**
pK7m24GW	*promAthDUO1-1244*	Mtr DUO1 stop	
pB7m34GW	*promAthDUO1-1244*	*Mtr DUO1 no stop*	GFP
pK7m24GW	*promMtrDUO1-247*	*NLS-GFP*	
pK7m24GW	*promMtrDUO1-171*	*NLS-GFP*	
pB7m34GW	*promMtrROD1x4*	*min35sH2B*	GFP
pB7m34GW	*promMtrDUO1-171*	H2B	GFP
pB7m34GW	*promMtrDUO1-247*	H2B	GFP
pB7m34GW	*promMtrDUO1-517*	H2B	GFP
pB7m34GW	*promMtrDUO1 726*	H2B	GFP
pB7m34GW	*promAthDUO1*-1244	H2B	GFP

pB7m34GW	*promAthROD1x4*	min35sH2B	GFP
pB7m34GW	*promOsaROD1x4*	min35sH2B	GFP
pB7m34GW	*promOsaROD1x4 ATG*	min35sH2B	GFP
pB7m24GW	*promAthDUO1-1244*	*Ath DUO1*	
pB7m24GW	*promAthDUO1-1244*	*Ath (m)DUO1*	
pB7m24GW	*promUB14*	*Ath (m)DUO1*	
pGREEN-0800-GW-LUC	*promAthDUO1-198*		
pGREEN-0800-GW-LUC	*promAthDUO1- MYB1*		
pGREEN-0800-GW-LUC	*promAthDUO1-MYB2*		
pGREEN-0800-GW-LUC	*promAthDUO1-Myb1-2*		
pGREEN-0800-GW-LUC	*promMtrDUO1*		
pGREEN-0800-GW-LUC	*promMtrDUO1-MYB-1*		
pGREEN-0800-GW-LUC	*promMtrDUO1-MYB-2*		
pGREEN-0800-GW-LUC	*promMtrDUO1-MYB1-2*		
pGREEN-0800-GW-LUC	*promOsaDUO1*		
pGREEN-0800-GW-LUC	*promOsaDUO1-MYB1*		
pGREEN-0800-GW-LUC	*promOsaDUO1-MYB2*		
pGREEN-0800-GW-LUC	*promOsaDUO1-MYB1-2*		
pGREEN 0800-5-LUC	*promAthDUO1-1244*		
pGREEN 0800-5-LUC	*promAthROD1x4*		
pGREEN 0800-5-LUC	*promMGH3*		

2.3 Bacterial culture and plasmid isolation

2.3.1 Luria Bertani (LB) medium

LB medium consisted of 1% (w/v) peptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl dissolved in MilliQ H_2O. For LB agar 1.5% (w/v) Agar was added. LB medium/LB agar was autoclaved for 20 minutes at 120 °C and 15 psi. LB agar was allowed to cool before appropriate antibiotics (Table 2.8 & 2.10) were added before pouring 25 ml into petri dishes and allowing to set.

2.3.2 Super optimal broth

Super optimal broth (SOB) was used for the culture of competent *E.coli*. SOB medium consisted of 2% (w/v) peptone, 0.5% (w/v) Yeast extract, 0.058% (w/v) NaCl, 0.019% (w/v) KCl, 0.25% (w/v) $MgSO_4$, 0.095% (w/v) $MgCl_2$. The pH was adjusted to 7 with 10 M NaOH then autoclaved at 120°C and 15 psi for 20 minutes.

2.3.3 Preparation of competent *E. coli*

"Ultra-competent" *E. coli* were prepared as described by Sambrook (2006) which is a modified version of the Inoue method for preparation and transformation of competent *E. coli* (Inoue, 1990). A single colony of DH5α *E.coli* was picked from a plate that had been grown for 16-20 hours at 37°C and transferred into 25 mL of LB medium in a 250 mL flask. The inoculated LB medium was then incubated for 6-8 hours at 37°C with vigorous shaking (250-300 RPM). Three 1 L flasks containing 250 mL SOB media were inoculated with three different volumes of starter culture; 25 µL, 50 µL, 1 mL. Inoculated flasks were incubated at room temperature (18-22°C) overnight with moderate shaking. The OD_{600} was then taken and cultures were monitored every 45 minutes until 0.55 was reached. Cultures were transferred to an ice water bath for 10 minutes. Cells were then harvested by centrifuging at 2500 *g* for 10 minutes at room temperature in 50 ml Falcon tubes. Media supernatant was removed via an aspirator and cells were then re-suspended in ice cold Inoue transformation buffer (55 mM $MnCl_2$, 15 mM $CaCl_2$ and 250 mM KCl) by swirling gently. Cells were then harvested by centrifugation at 2500 *g* for 10 minutes and the supernatant removed via aspiration. Cells were re-suspended gently in 20 ml of ice cold Inoue transformation buffer and placed on ice for 10 minutes. Fifty µl aliquots of suspended cells were aliquoted into pre-chilled sterile 1.5 ml microfuge tubes and flash frozen in liquid nitrogen. Cells were then stored at -80 °C until use.

2.3.4 Transformation of competent *E. coli*

Two strains of *E. coli* were used for transformations. TOPO reactions, BP reactions and one/two part LR reactions were transformed into lab made competent DH5α (Section 2.3.3). Lower efficiency transformations such as 3-part LR reactions used OneShot TOP10 competent *E.coli* (Invitrogen) cells. For both 50 µL aliquots of cells the DNA to be transformed into either aliquots was thawed on ice for 15 minutes. If the DNA was from a mini-prep the DNA was diluted 1/100 or 1/1000 depending on concentration. 1µl (125 ng) of plasmid or 2.5 µl of the TOPO® reactions, BP reaction or LR reactions was added to a 50 µl aliquot and incubated on ice for 30 minutes. Cells

were then heat shocked at 42°C for 45 seconds (DH5α cells) or 30 seconds (OneShot TOP10 cells). Five hundred µL of room temperature LB medium was added and cells were incubated at 37°C on a shaker at 220 RPM for 1 hour to recover. 50 µL was then plated onto LB agar containing appropriate antibiotics (Table 2.8 & 2.10). Remaining cells were centrifuged at 3000 *g* for 3 minutes and supernatant was removed and the pallet was re-suspended in 100 µL LB medium and plated on LB agar containing appropriate antibiotics. Plates were then incubated at 37°C overnight.

2.3.5 Preparation of competent *A. tumefaciens*

A single colony of *A. tumefaciens* strain GV3101 with or without the pSOUP helper plasmid was grown overnight in 5 mL LB medium containing 50 ng/mL Rifampicin and 50 ng/mL Gentamycin. One ml of the overnight culture was used as a starter culture for 50 mL LB medium in a 250 mL flask with the antibiotics and shaken vigorously (250 RPM) at 28 °C until the culture grew to OD_{600} of 0.5 to 1.0. The culture was chilled on ice and then centrifuged at 3000 *g* for 5 minutes at 4°C. Cells were re-suspended in 1 mL of ice-cold $CaCl_2$ solution (20 mM). Fifty µL aliquots of suspended cells were aliquoted into pre-chilled sterile 1.5 mL microfuge tubes and flash frozen in liquid nitrogen. Cells were then stored at -80°C until use.

2.3.6 Transformation of competent *A. tumefaciens*

A modified freeze-thaw method was used to transform *A. tumefaciens*. Plasmids were chilled on ice for 20 minutes. 1 - 3 µL or 0.5-1 µg of plasmid DNA was then immediately added to a frozen 50 µL aliquot of cells. Cells were immediately incubated in a 37°C water bath for 5 minutes. 0.5 mL of LB medium was then added before incubating at 28°C for 2-4 hours. Cells were then plated on LB agar plates containing appropriate antibiotics. Plated cells were then allowed to grow for 2-3 days at 28°C for colonies to form.

2.3.7. PCR colony screen for transformed bacteria

Individual colonies from transformed *E. coli* and *A. tumefaciens* were PCR screened for the plasmid of interest. Gateway® vectors have M13 primer sites that flank *att* or

TOPO® sites so either forward or reverse M13 primers (Table 2.11) and a screening/cloning primer (Table 2.1 & 2.11) that bind within the inserted DNA were used to confirm the presence of the correct plasmid. A 10 µL Platinum Taq (Invitrogen) PCR reaction was used to screen individual colonies (Table 2.12). Colonies were patched on to LB agar containing appropriate antibiotics with a 100 µL pipette tip and then the pipette tip is transferred to the 10 µL PCR reaction mix to remove remaining cells providing the DNA to be screened.

Table 2.11. Primers designed for PCR colony screening.

GFP R screening primer - end 251 bp from ATG	AAGTCGTGCTGCTTCATGTG
M13F screening primer	GTAAAACGACGGCCAG
M13R screening primer	CAGGAAACAGCTATGAC
Ath promDUO1 F screening - 154 bp from cloning site	CGAGAGAAACTTGGGGAGTG
Mtr promDUO1 F screening- 153 bp from cloning site	TGTGTGCGTGTGTAGCAAAG

Table 2.12: Platinum *Taq* Polymerase PCR colony screen reaction mixture and PCR conditions

Platnium *Taq* reaction components	**Volume (µL)**	**Temperature (°C)**	**Time (seconds)**
10x PCR Buffer (-$MgCl_2$)	1	98	600
$MgCl_2$ (5 mM)	0.3	98	30
dNTP (10 mM)	0.2	55	30 **- x30**
Forward primer (10 mM)	0.2	72	60 / kb
Reverse primer (10 mM)	0.2	72	60 / kb x2
Platnium *Taq* polymerase	0.04		
Water	8.06		
Total	**10 (µl)**		

2.3.8 Plasmid DNA isolation from bacteria

Plasmids from transformed *E. coli* were extracted by small scale mini plasmid DNA preparations using a PureLink™ HQ Mini Plasmid Purification Kit (Invitrogen); all reagents used were part of the kit and prepared as instructed. Five mL of LB medium with appropriate antibiotics was inoculated with transformed *E. coli* cells and grown overnight at 37°C at 220 RPM. Cells were then harvested by centrifuging 3 mL of culture in two separate 1.5 mL microfuge tubes. The supernatant was removed and the pellet from one tube was re-suspended in 240 µL RNase A solution (0.1 mg/mL), this was then added to the second microfuge tube used to re-suspend the remaining cells and 240 µL of lysis buffer was then added and mixed gently by inverting the tube 4-8 times. This was then incubating at room temperature for 3-5 minutes before 340 µL of Neutralization/Binding buffer was added and inverted 4-8 times, then centrifuged for 10 minutes at maximum speed (17,000 *g*). The Supernatant was pipetted carefully into a PureLink™ spin column inside a 2 mL collection tube and centrifuged for 1 minute at 14,000 *g* and 650 µL of wash buffer was added to the column and centrifuged again for 1 minute at 14,000 *g*. The flow through was discarded and the column was then centrifuged at 17,000 *g* for 1-3 minutes to remove residual wash buffer. DNA was then eluted by adding 50 µL of elution buffer and incubating for 1 minute then centrifuged at 17,000 *g* for 1 minute. The elution tube contains plasmid DNA which was either used immediately or stored at -20°C.

2.3.9 Sequence verification of DNA

To verify DNA sequence of entry clones DNA samples were sequenced by the Genetic Analysis Services (GAS) of the University of Otago, in accordance with GAS procedural instructions.

Expression vectors were also verified by restriction enzyme digests with restriction enzymes obtained from New England Biolabs (U.S.A.) and Roche (Germany). Digests were conducted as instructed by the manufacturer's guidelines.

2.4 Microscopy

2.4.1 DAPI staining of pollen nuclei

Nuclear content of pollen was visualised by staining with 4',6-diamidino-2 phenylindole dihydrochloride (DAPI) as described by Park *et al.,* (1998). DAPI staining solution consisted of 0.2 M Na_2HPO_4 30.5% (v/v), 0.2 M NaH_2PO_4 18.5% (v/v), 0.5 M EDTA 0.2% (v/v), Triton X 0.1% (v/v), 0.5 mg/mL DAPI stock solution 0.8% (v/v) (Sigma). Three or four open flowers were placed in a microfuge tube containing 180 µL of DAPI solution. Pollen grains were released by being pulse vortexed 2-3 times then collected by centrifugation at 4000 *g* for 2-3 seconds to form a pollen pellet. Eight µl of the pollen pellet was transferred to a microscope slide and allowed to settle for 30 seconds before a 22 mm x 22 mm coverslip was mounted. If slides were not visualised immediately nail polish was used to seal the edges of the coverslip to prevent drying out of the pollen. Samples were visualised under a fluorescent microscope (Section 2.5.3).

2.4.2 Developmental dissections of inflorescence buds

To determine when promoter reporter constructs were first visualised during developmental, dissections of inflorescence buds at various growth stages were analysed. Buds were submersed in milliQ H_2O on a microscope slide and detached and sequentially arranged from +1 being the first open flower stage (mature tricellular pollen), -1 being the unopened bud next to the open flower and successive buds were arranged through to the -10 bud developmental stage (early microspores) (Lalanne and Twell, 2002). Individual buds were removed for dissection from the milliQ H_2O and placed on another microscope slide and submersed in DAPI staining solution (Fluorescence microscopy) or 0.3 M mannitol (confocal microscopy). Using dissection needles anthers were isolated and broken open to release developing pollen. Anther cellular debris was removed and 22 mm x 22 mm coverslips were mounted and sealed with nail polish. Samples were then visualised under a fluorescence microscope (Section 2.5.3) or by confocal laser scanning microscopy (CLSM section 2.4.4).

2.4.3 Fluorescent imaging of pollen

All fluorescent images of pollen grains were captured using an Olympus^{tm} IX71 inverted microscope with mercury lamp for fluorescence excitation. The Olympus DP controller and DP manager imaging software programs were used to preview, capture and save images. DP controller allowed exposure settings to be changed depending on the wavelength that was being visualised with the different emission filters of the microscope. An exposure time of 143.33 milliseconds was used for the GFP filter and 5 milliseconds was used for DAPI imaging. An ISO sensitivity of 1600 was used for all images. Once captured all images opened automatically in DP Manager and were saved in TIFF format

2.4.4 Confocal scanning laser microscopy

The different stages of pollen development from inflorescent dissections mounted in mannitol were visualised by confocal laser scanning microscopy (CLSM). An Olympus FluoView FV1000 microscope with Olympus FluoView software and Olympus BX61 camera with a 60X oil objective was used. Dwell time was set to 20 μs and the pinhole 100 nm. A 473 nm laser was used for GFP excitation with detection between 485 and 494 nm with the sensitivity of the photon multiplier detector altered between 500 and 650 mV depending on signal strength. Transmitted light was also visualised and the detector was set at 140 mV. All images are an average of four scans (Kalman set to 4). Lynette Brownfield assisted with the operation of the microscope and capture of images.

2.4.5 Fluorescence quantification

ImageJ 1.49v imaging software was used to measure the nuclear GFP fluorescence intensity of mature pollen sperm cells from transgenic plants. One hundred sperm cell nuclei from individual pollen grains were sampled for each plant. The background noise of the photos was reduced by first subtracting for a standard correction value of 500.0 pixels. Fluorescence intensity of GFP expressed in the nucleus of sperm cells was then measured by drawing a circle around the sperm cell nucleus and measuring GFP fluorescence. Then using the same size circle area of the nucleus a measurement

was taken of the vegetative cell cytosol as a background measurement to correct for natural pollen cell wall auto-fluorescence. The background measurement was then subtracted from the sperm cell nuclear measurement to give the fluorescent intensity.

2.5 Transient dual luciferase assay in *Nicotiana benthamiana*

2.5.1 *A.tumefaciens* mediated infiltration transformation of *N. benthamiana* leaves

N. benthamiana plants were grown in long day conditions of 16 hours light at 20°C till 4-6 weeks old. 5 mL of LB medium containing appropriate antibiotics was inoculated with freshly grown *A. tumefaciens*, then cultured overnight at 28°C shaking at 200 rpm and 1.5 mL of culture was removed into a micro-centrifuge tube and centrifuged at 1,000 *g* for 10 minutes to pellet. Supernatant was removed and pelleted cells were re-suspended in 1mL infiltration buffer (500 mM MES, 20 mM $Na_3PO_4.12H_2O$, 1 M acetosyringone 0.0001% (v/v), D-glucose 0.05% (w/v) and made up in MilliQ water). The previous wash step was repeated and cells pelleted at 1000 *g* for 10 mins. The supernatant was removed and cells were re-suspend in 1 mL infiltration buffer to remove any trace of antibiotics. The OD_{600} of a 1 in 10 dilution of suspended cells was measured and then used to calculate titres of the amount of re-suspended cells needed to bring cells to a final OD_{600} of 0.1. 1 ml of a 10:1 dilution of bait to prey *A. tumefaciens* strains were set up in 1.5 ml micro-centrifuge tubes. 1 hour before infiltrating *N. benthamiana* were placed under white light and watered to allow for easier infiltration. One mL of *A. tumefaciens* dilutions to be infiltrated were taken up in a 1 ml syringe with no needle. Two larger leaves that aren't the cotyledons were then chosen to be infiltrated. The syringe tip was placed against the underside of the leaf and *A. tumefaciens* was gently infiltrated. Either side of the midrib region of the 2 leaves were infiltrated and marked with permanent marker or pin pricked to identify infiltrated leaves. Plants were allowed to grow for 2-5 days before analysing through dual luciferase assays.

2.5.2 Dual luciferase assays of leaf tissue extract

Passive lysis buffer (Promega) was diluted from a 5x stock solution and 50 µL was aliquoted into individual wells of a white 96 well flat bottom microreader plate.

Infiltrated *N. benthamiana* leaf tissue was then excised using a 0.7 mm diameter hole press and placed into wells with passive lysis buffer. Excised leaves were then ground until homogenous and contained no visible large leaf debris. DLAR (Luciferase substrate) and RLAR (Renilla substrate) Promega luciferase reagents were defrosted and brought to room temperature and Renilla stop reagent was added to the RLAR buffer. A pre-set program designed to measure the luminescence of the Firefly luciferase and Renilla was chosen on the Clariostar. The injectors of the Clariostar microplate reader were primed with 500 µL of the DLAR reagent in pump 1 and RLAR reagent in pump 2. The plate layout was set on the protocol on the Clariostar to match the samples in the 96 well plate and injector volumes were set to 50 µL. Once measurements were taken the Clariostar software was closed and data opened in the Mars analysis Software. The data from the Clariostar was then available to be viewed and extract to a excel spread sheet for further analysis.

2.6 Bioinformatics

2.6.1 DUO1 orthologue identification and alignment

A TBLASTN search of the *M. truncatula* genome was undertaken on Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) (Goldstein *et al*., 2012) using the *A. thaliana* DUO1 amino acid. Selected DUO1 orthologues (AthDUO1 and MtrDUO1) were then aligned using Geneious 6.1 (http://www.geneious.com; Kearse et al., 2012) Blosum62 pairwise alignment program.

2.6.2 *ROD1* Multiple nucleotide alignment

ROD1 alignments of *AthROD1*, *MtrROD1* and *OsaROD1* were done by using the multiple nucleotide alignment MaliN available from SoftBerry (http://www.softberry.com/) with standard alignment settings (Gap Continuation penalty = 0.5, Gap Initiation penalty = 2, Match score = 5, Mismatch penalty = -4, Pre-defined matrix = BLAST_MatrixX).

Chapter 3

Complementation of the *A. thaliana duo1* mutant with *M. truncatula* DUO1

3.1 Introduction

The *A. thaliana* DUO1 protein (AthDUO1) encodes a 297 amino acid R2R3 MYB transcription factor that is first expressed after the asymmetric division in the developing male germ-line (Rotman *et al*., 2005). In *A. thaliana* it has since been established that DUO1 has an important role in regulating a plethora of key genes which are necessary for the generative cell cycle progression at PMII and differentiation (Brownfield *et al*., 2009, Borg *et al*., 2011).

The key feature that defines AthDUO1 from other R2R3 MYB transcription factors is the presence of a supernumerary lysine residue (K66) at the beginning of the R3 domain (Rotman *et al*., 2005). The K66 residue has been shown to be functionally important for correct transactivation of DUO1 targets in *A. thaliana* (Borg, 2010). The K66 residue has been widely conserved across divergent monocotyledonous and dicotyledonous angiosperm orthologues of DUO1 (Rotman *et al*., 2005; Borg, 2010) and it is feasible that this conservation is important functionally and has been evolutionarily retained.

Little is known if DUO1 orthologues have the same key male germ-line regulator function across the various angiosperm species. While sequence similarity suggests there is a conserved DUO1 gene among angiosperms, the function of these orthologues has not yet been widely tested. Although functional orthologues have been identified in rice, lily and tomato (Twell, unpublished) which are all from evolutionarily distinct flowering plant families to *A. thaliana* (Figure 3.1). This chapter thus explores whether DUO1 from the evolutionarily distinct legume *M. truncatula* (Fabaceae family; Figure 3.1) is functionally conserved to that of AthDUO1 in *A. thaliana*.

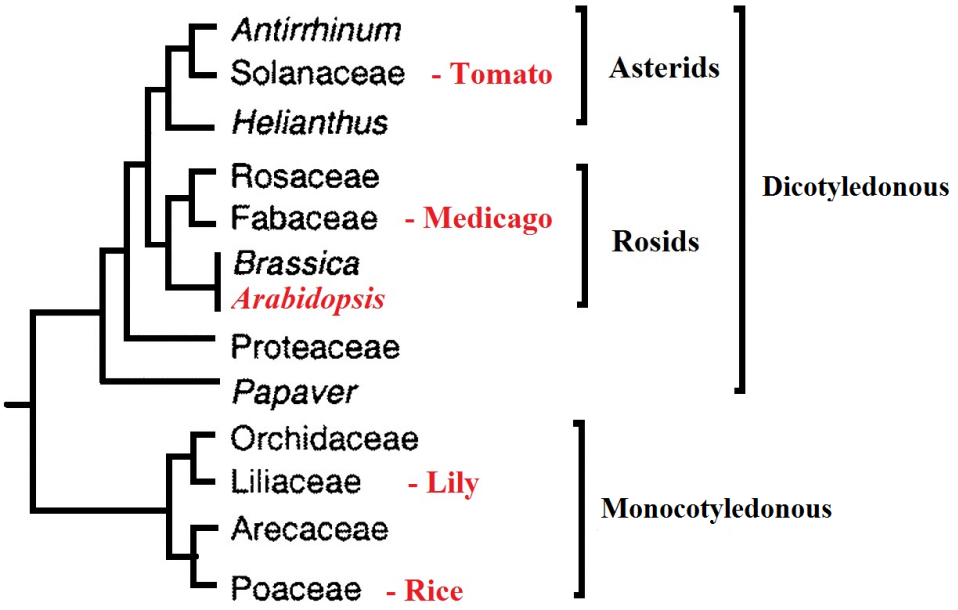


Figure 3.1. Phylogenetic tree of dicotyledonous and monocotyledonous flowering plant families. Phylogenetic tree contains key families from dicot and monocot plant species showing evolutionary diversity among the various families. Common model organisms from various key plant families are shown in red; Tomato, *Medicago*, *Arabidopsis*, lily, rice. (Adapted from Ma *et al*., 2000)

3.2. Results

3.2.1 Identification of *M. truncatula* DUO1

To identify a DUO1 orthologue in *M. truncatula* a TBLASTN search of the *M. truncatula* genome was undertaken on Phytozome (Goldstein *et al*., 2012) using the *A. thaliana* DUO1 amino acid (a.a) sequence. This identified a MYB transcription factor (Medtr8g006470) that had an E value of 8.2^{-26}. The amino acid sequences from Medtr8g006470 and AthDUO1 were aligned and it was shown there is 50% identical residues with the highest similarity in the R2R3 domain (Figure 3.2). Importantly the lysine at position K66 of AthDUO1 was conserved consistent with MYB homology of DUO1 (Figure 3.2). No other hits from the TBLASTN search contained the conserved K66 residue so it was concluded Medtr8g006470 was the only orthologue of AthDUO1. Medtr8g006470 is here after referred to as MtrDUO1.

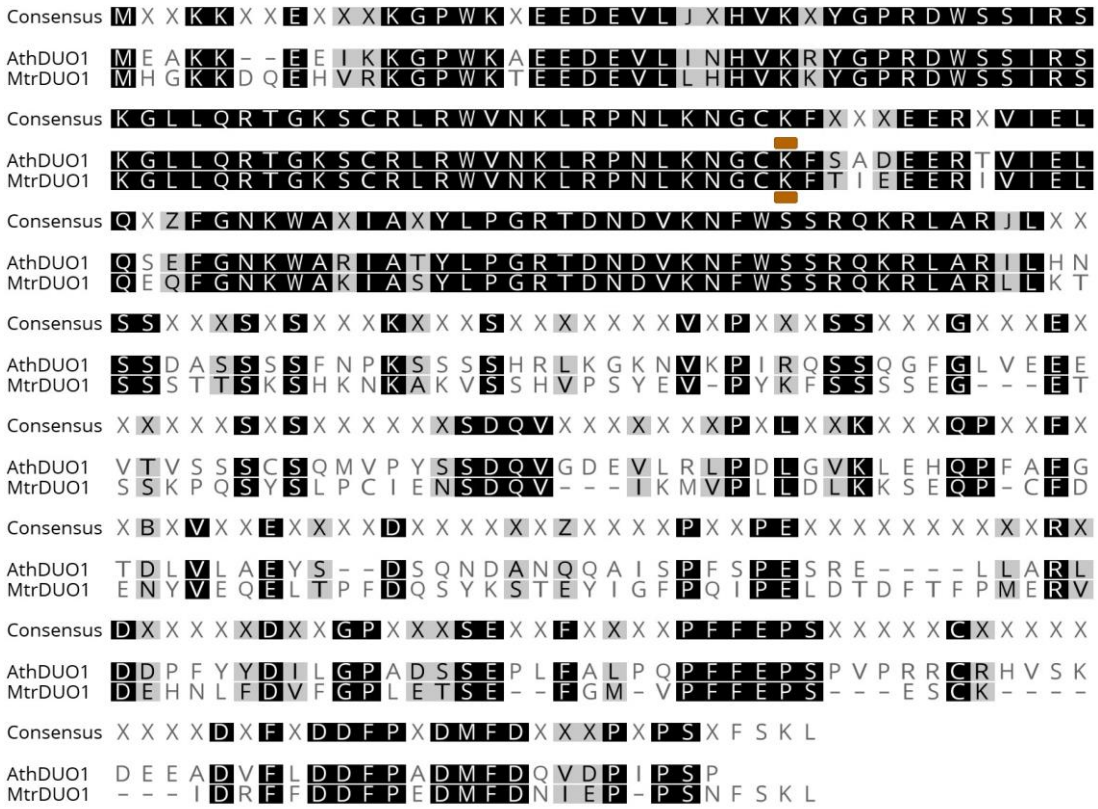


Figure 3.2. Amino acid alignment of AthDUO1 and MtrDUO1 protein sequences. The amino acid sequences of AthDUO1 and MtrDUO1 were aligned using a Geneious Blosum62 pairwise alignment program. Highlighted black sequences show identical residues with a large portion of these being in the first 122 amino acids (50% in total) containing the R2R3 MYB domain. Similar amino acids are highlighted grey. K66 of AthDUO1 is conserved in MtrDUO1 as shown by red boxes. Gaps in alignment are shown by - . Consensus sequence of the alignment is shown above the aligned residues; X represents undefined consensus sequence identities.

3.2.2 Generation of *DUO1/duo1* plants with MtrDUO1 complementation constructs

In order to test to see if MtrDUO1 could rescue cell cycle defects and sperm cell function in *duo1* pollen from *A. thaliana DUO1/duo1* plants, two expression vectors were created through Gateway® cloning. A 1244 bp sequence of the *A. thaliana DUO1* promoter (*promAthDUO1*) (Table 2.8) was used to drive expression of the genomic copy of *MtrDUO1* (Table 2.8). *promAthDUO1* was used as it has known male germ-line specific expression in *A. thaliana* (Brownfield *et al*., 2009) which would result in *MtrDUO1* being expressed specifically in the male germline and at the right time (early bi-cellular pollen) for correct function. *MtrDUO1* then either had a stop codon (Figure 3.3A) in one expression vector or the stop codon removed and the sequence for GFP added to it in the correct open reading frame (Figure 3.2B).

Expression vectors were verified through restriction digests and used to transform *A. thaliana DUO1/duo1* plants via floral dip (Section 2.1.3) and seeds collected. Seeds were sown and transgenic plants were selected (Section 2.1.2) and resulting pollen phenotypes analysed.

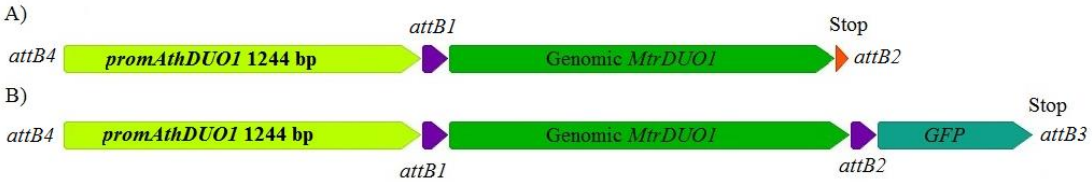


Figure 3.3. Schematic diagram of genomic *MtrDUO1* complementation expression constructs. *AthDUO1* promoter 1244 bp fragment drives expression of genomic *MtrDUO1* with either a natural stop codon in A) or a removed stop codon in B) followed with *GFP* in frame. *att* sequences between the *AthDUO1* promoter, genomic *MtrDUO1* and *GFP* are a result of Gateway® recombination cloning which provide site specific positioning of DNA fragments.

3.2.3 MtrDUO1 rescues the cell cycle phenotype of *duo1* pollen

Complementation of the cell cycle defect of *duo1* in *A. thaliana* would restore mitotic function at PMII resulting in a tri-cellular *duo1* pollen grain. If complementation occurs approximately 75% of the pollen grains from a transformed *DUO1*/*duo1* plant are expected to be a tri-cellular, resulting in a ratio of 3:1 tri-cellular to bi-cellular pollen. This is due to *DUO1/duo1* plants being heterozygous and pollen being the haploid generation, meaning 50% of the pollen has the DUO1 allele and is therefore tri-cellular. The remaining 50% of pollen will have the *duo1* allele and half of these will have the complementation construct. As following meiosis there's a 50% chance a *duo1* pollen grain will either have the complementation construct or not (assuming a single T-DNA insertion). Thus if complementation occurs half of the *duo1* pollen (25% of total pollen) will be tri-cellular and the other half (25% of total pollen) will remain *duo1* with a bi-cellular phenotype (Table 3.1). If the complementation constructs are non-functional we would expect to see a 1:1 ratio of tri-cellular to bi-cellular pollen grains as normally observed in *DUO1/duo1* plants.

Table 3.1: Expected pollen phenotypes from complement *DUO1/duo1* pollen

Genotype of pollen from complemented *DUO*/*duo1* plants	**Probability of occurrence post-meiosis (%)**	**Phenotype**
Wt	25	Tri-cellular
Wt + complementation construct	25	Tri-cellular
duo1 + complementation construct	25	Tri-cellular
duo1	25	Bi-cellular

To determine if the *duo1* phenotype had been complemented in individual primary transformant lines (T1 Lines), *DUO1/duo1* plants first had genomic DNA extracted and were genotyped by PCR screening with specific primer pairs to confirm insertion of complementation construct. PCR Amplicons from two plants of each transformed construct were sequenced to confirm the presence of the correct complementation sequence. Thirty two individual primary transformant lines (T1 lines) harbouring the *MtrDUO1* complementation construct with the stop codon and 28 T1 lines harbouring

the *MtrDUO1:GFP* complementation construct were generated. Two hundred pollen grains from each individual line were then examined by fluorescence microscopy. The pollen from individual T1 lines was stained with DAPI to score the number of generative/sperm cell nuclei to determine the number of pollen containing tri-cellular and bi-cellular pollen.

Of the thirty two transformed *MtrDUO1* stop T1 lines eighteen were likely to be DUO1/DUO1 as they contained over 95% tri-cellular pollen. Five *MtrDUO1* stop T1 lines contained pollen with an average of 76.9% tri-cellular pollen (3:1 ratio) (Figure 3.4 lines 1-5) consistent with complementation of the mutant mitotic phenotype in *duo1* pollen. Statistical analysis of the deviation of the ratio of bi-cellular pollen to tri-cellular pollen was carried out using a chi-square analysis on lines 1-5 showing a significant difference between complemented *DUO/duo1* lines (3:1 ratio) and *DUO1*/*duo1* lines (1:1 ratio) ($\chi^2 = 113.9$, $\rho<0.001$).

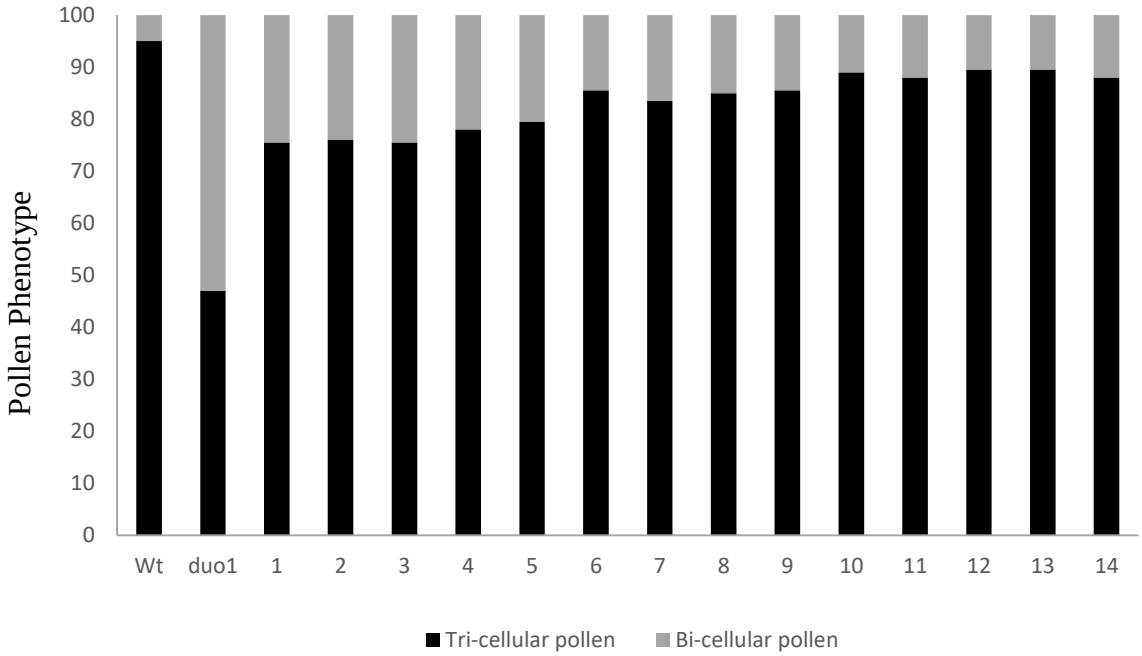


Figure 3.4: Percentage of tri-cellular and bi-cellular pollen in complemented *MtrDUO1* stop T1 pollen grains. Mature pollen was stained with DAPI to determine the percent tri-cellular and bi-cellular pollen. Pollen grains (1000) from 10 wildtype (WT) and 10 *DUO1/duo1* (duo1) plants were counted as controls. Pollen grains (200) from each T1 line (1-14) containing the MtrDUO1-stop construct were from a DUO1/duo1 plant were counted

In addition nine T1 lines contained pollen with an average of 86.8% tri-cellular pollen (Figure 3.4). This unexpected number could be due to multiple independent insertions of the complementation construct, meaning more than 50% of the *duo1* pollen had an insertion and was complemented in these T1 lines. These results show that the mitotic division mutant phenotype at PMII appears to be rescued by *MtrDUO1* stop allowing tri-cellular pollen grains to form.

Pollen from twenty eight *MtrDUO1:GFP* T1 lines was either over 95% tri-cellular (16 lines) or close to 50% tri-cellular (12 lines). The lines with 50% tri-cellular pollen are likely to be *DUO1/duo1* and not have been complemented. As plants had been verified to contain the GFP complementation construct, pollen for all lines was analysed for nuclear localised GFP. However no GFP was observed from any lines suggesting these constructs were not being expressed so could not complement. The reason the addition of GFP impacted upon expression is unknown, but was not pursued as the data with *MtrDUO1* stop clearly showed complementation.

3.2.4 *duo1* differentiation is rescued by *MtrDUO1*

Having established the mitotic division defect is restored in complemented *MtrDUO1* stop T1 lines it is then of importance to determine if the differentiation of the male generative cell into two functional sperm cells has also been restored. This is due to DUO1 not only being required for PMII but also for sperm cell differentiation and the expression of key genes required for fertilization (Brownfield *et al.*, 2009; Borg *et al.*, 2011). Biologically this is significant as without two functional sperm cells double fertilization cannot occur and the male germ-line is unable to contribute to the future generation.

To test sperm cell function, male transmission of the *duo1* allele from plants complemented with MtrDUO1 was analysed. As the *duo1* mutation blocks the ability of pollen to undergo mitosis, it is not transmitted through the male germ-line and can only be maintained in female heterozygote lines. Any new *DUO1*/*duo1* plants in a crossed F1 progeny must inherit the *duo1* mutation from the male germ-line of the complemented T1 line as the female is WT *DUO1/DUO1*, thus showing rescued male

transmission of the complemented *DUO1/duo1* T1 lines due to inheritance of the *duo1* mutation.

Pollen from the four complemented *MtrDUO1* stop T1 (Figure 3.4) was crossed onto emasculated (anthers removed to prevent self-pollination) wildtype flowers. A limited number of seeds was collected from these crosses as only 1 silique forms from each cross. Seeds were sown and F1 plants were allowed to grow, and the pollen phenotype analysed by DAPI staining.

Table 3.2: Transmission of the *duo1* allele from *DUO1/duo1* plants complemented with *MtrDUO1* stop when crossed onto wildtype females. Plants that showed restored mitotic function with the *MtrDUO*1 stop complementation construct were crossed with wildtype female flowers to see if the male transmission of the *duo1* allele was restored. Ten F1 (lines 1-10) were grown and pollen grains stained with DAPI and scored for the frequency of the bi-cellular and tri-cellular phenotype in 200 pollen grains.

Line	Nuclei Counted	Bi-cellular pollen	Bi-cellular (%)	Tri-cellular pollen	Tri-cellular (%)	Deduced Genotype
1	200	42	21	158	79	*duo1 + comp*
2	200	47	23.5	153	76.5	*duo1 + comp*
3	200	55	27.5	145	72.5	*duo1 +comp*
4	200	22	11	178	89	Wt
5	200	9	4.5	191	95.5	Wt
6	200	6	3	194	97	Wt
7	200	14	7	186	93	Wt
8	200	0	0	200	100	Wt
9	200	11	5.5	189	94.5	Wt
10	200	20	10	180	90	Wt

Seven of ten plants had predominantly tri-cellular pollen, so likely received the DUO1 allele from the male germline (Table 3.2). The remaining three plants had approximately 75% tri-cellular pollen. This is consistent with plants having received both the *duo1* allele and the MtrDUO1 stop insertion from the complemented male germline. Statistical analysis shows that the 3:1 tri-cellular / bi-cellular ratio observed in the crossed F1 generation is statistically different to the 100% tri-cellular phenotype of the maternal wild type pollen (χ^2 =350.0, ρ= <0.001). This shows that sperm cell function was restored in *duo1* pollen by *MtrDUO1* stop as the sperm cells could successfully fertilize. However due to low number of seeds generated the efficiency of rescue by *MtrDUO1* could not be determined.

3.3 Conclusion

The results of this section show that the candidate *MtrDUO1* gene identified is a functional orthologue of *AthDUO1*. The ability of MtrDUO1 to complement the cell cycle division mutant phenotype at PMII in *DUO1/duo1 A. thaliana* provided confirmation. Complemented *MtrDUO1* pollen grains also demonstrated that there was restored cell specification occurring in pollen to form two functional sperm cells allowing the complemented *duo1* pollen to have male transmission and pass on the *duo1* phenotype. This shows that MtrDUO1 is targeting the same key downstream gene regulatory networks in complemented *A. thaliana DUO1/duo1* lines as of what AthDUO1 would regulate in wildtype pollen.

Chapter 4.

Medicago truncatula DUO1 promoter analysis

Results of this chapter have contributed to Figure 4 in the following publication

Peters, B., Casey, J., Aidley, J., Zohrab, S., Borg, M., Twell, D. and Brownfield, L. (2017) A Conserved *cis*-Regulatory Module Determines Germline Fate through Activation of the Transcription Factor DUO1 Promoter. *Plant Physiology*. 173. 1 280-293

4.1 Introduction

Once it was established that *MtrDUO1* is functionally conserved in *A. thaliana*, I next asked if regulation is also conserved. Studies have been undertaken on the regulation of *A. thaliana* DUO1 reporter constructs with *promAthDUO1* first visualised in the generative cell soon after the asymmetric division at PMI (Brownfield *et al*., 2009). Furthermore, the small *ROD1* region has been shown to be both necessary and sufficient for the germ-line expression of *DUO1* (Peters *et al.*, 2017). Shown using a combination of a 5′ deletion series and small regions of *promAthDUO1* to drive expression of reporter constructs in *A. thaliana* pollen (Brownfield *et al.*, 2009; Peters *et al*., 2017). *ROD1* is -153: -83 upstream of the ATG and contains multiple CRE motifs that may allow transcriptional regulation (Figure 1.4; Peters *et al*., 2017).

It has also been shown that a sequence upstream of *ROD1* termed *ROD2* increases *DUO1* expression (Peters *et al*., 2017). Alignments of the *DUO1* promoter sequence from various plant species have shown there is high conservation within *ROD1* between evolutionarily distinct plants (Figure 1.4). While these sequences are conserved the regulation of *DUO1* has not been widely studied in other species with no functional testing of *ROD1* being explored from other plants.

This part my study thus aims to build knowledge on the evolutionarily conserved regulation mechanisms of DUO1 in other plant species too *A. thaliana*. The use of *M. truncatula DUO1* promoter (*promMtrDUO1*) reporter expression constructs were used to test if the functional architecture and expression pattern of *promMtrDUO1* is conserved with the known male germ-line specific expression of *promAthDUO1*. A 5′

promoter deletion series of the *M. truncatula* promoter was used to identify the minimal region needed to drive germ-line specific expression. It is also of interest to test if *ROD1* promoter sequences from plants such as *M. truncatula* and *O. sativa* can drive male germ-line specific expression of promoter reporter constructs.

4.2 Results

4.2.1 Generation of promoter reporter lines

To analyse the activity of various promoter sequences, promoter reporter expression vectors were constructed (Figure 4.1). These expression vectors allow promoter sequences to be placed in front of histone *H2B* and *GFP* sequences through site specific recombination of *att* sites (Section 2.2.5).The histone H2B localises GFP to the nucleus enabling the easy distinction of any vegetative and sperm cell expression providing a confined region for the quantification of expression of promoter activity. All constructs were created using Gateway® Cloning (Karimi *et al*., 2002). The high adenine content around the ATG of *MtrDUO1* meant the promoter fragments started

at the -5 bp position relative to the ATG. The longest promoter fragment able to be cloned was 726 bp, due to an incomplete genome sequence on the annotated

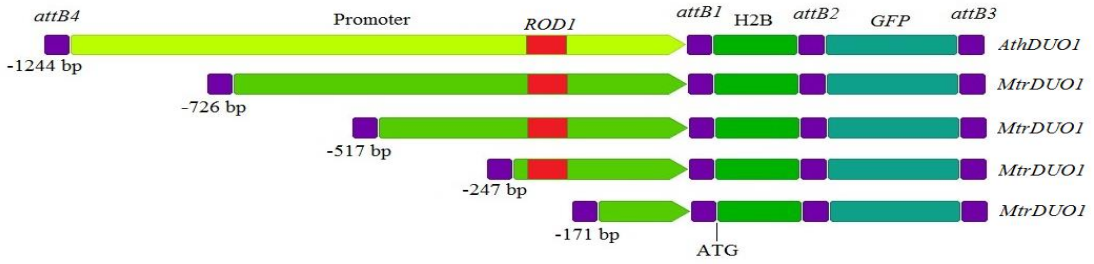


Figure 4.1: Schematic of *MtrDUO1* promoter reporter constructs. Promoter sequences drive expression of histone H2B which provides nuclear localisation and GFP which provides visualisation of expression. The ATG of H2B provides a translation initiation site. *A. thaliana DUO1* promoter used as a positive control is shown in light green (top) and *M. truncatula* promoter fragments are shown below ranging from -726 bp to -171 bp in length from *MtrDUO1*'s ATG. The conserved *ROD1* region is shown in red. Purple linker *att* sequences between the coding regions of H2B and GFP are a result of Gateway® recombination cloning which provides site specific positioning of DNA fragments. (Reporter construct lengths not to scale with each other)

M. truncatula genome available from Phytozome (Goldstein *et al*., 2012). *DUO1/duo1 A. thaliana* plants were transformed with promoter reporter constructs (Figure 4.1; Table 2.10) and T1 plants were selected for herbicide resistance and fluorescence in pollen grains was analysed.

4.2.2 Male germ-line specific expression of *DUO1* is conserved between *A. thaliana* and *M. truncatula*

To determine if the promoters of *AthDUO1* and *MtrDUO1* have the same male germ-line specific expression, the longest length promoter reporters were visualised with epifluorescence microscopy. The *A. thaliana* construct consisted of a 1244 bp fragment of *DUO1* promoter attached to *H2B* and *GFP*; *promAthDUO1*:*H2B-GFP*. The *M. truncatula* construct consisted of a 726 bp fragment of *DUO1* promoter attached to *H2B* and *GFP*; *promMtrDUO1:H2B-GFP* (Figure 4.1). Ten individual T1 lines of both promoter reporters showed fluorescence of GFP in sperm cells of mature pollen with no vegetative cell expression visualised (Figure 4.2). This shows that the *promAthDUO1:H2B-GFP* and *promMtrDUO1:H2B-GFP* both provide male germ-line specific expression in mature pollen in *A. thaliana*.

PromAthDUO1:H2B-GFP and *promMtrDUO1:H2B-GFP* expression patterns were then observed in the developmental stages of pollen (Figure 1.1 shows these stages). Inflorescences containing buds with pollen were dissected out at various stages of development, from microspores through to mature pollen grains (Section 2.5.2) from selected T1 plants. This allowed identification of when *promMtrDUO1* is first expressed in development by analysing when GFP is first visualised. The initial analysis was performed by epifluorescence microscopy (Figure 4.2) and DAPI was used to stain nuclei to determine stage of development and nuclear localisation. DAPI also allows easy identification of developmental stage as the condensed generative cell nuclei stain bright and de-condensed vegetative cell nuclei stain lighter (Figure 4.2). Two plants from both *promAthDUO1:H2B-GFP* and *promMtrDUO1:H2B-GFP* were analysed for GFP expression from microspore through to mature pollen (Figure 4.2). No GFP expression was seen within the microspore (MS; Figure 4.2 A, C). Germ-line specific expression of GFP was then first observed in the early bi-cellular stage from

both reporter constructs (Figure 4.2). Germ-line specific expression of GFP was maintained throughout development through to mature pollen. No vegetative cell GFP expression was seen at any stage of development. Thus *promMtrDUO1* appears to regulate expression similar to *promAthDUO1*.

Epifluorescence microscopy however did not provide the sensitivity required to confidently confirm germ-line specific expression. Confocal laser scanning microscopy (CLSM) with increased sensitivity over epifluorescence was then used to confirm the expression pattern through analysing dissections of buds from two T1 plant lines for each of *promAthDUO1:H2B-GFP* and *promMtrDUO1:H2B-GFP* constructs. Pollen of the various developmental stages was mounted in 0.3 M mannitol which was used as an osmotic protectant to help prevent rupture of pollen grains (Figure 4.3 A-C and 4.4 A-C). DAPI was no longer used as it can enhance auto-fluorescence by causing cell death. The lack of DAPI also eliminates any potential fluorescence bleed-through.

The *promAthDUO1:H2B-GFP* and *promMtrDUO1:H2B-GFP* reporters showed the same developmental expression pattern as each other in developing pollen grains (Figure 4.3 (A) and Figure 4.4 (A)). No GFP is visible in the microspore (MS) nucleus in either constructs (Figure 4.3 and 4.4). Weak GFP is first visualised exclusively in the male generative cell after the asymmetric division (PMI) from both promoter constructs in the early bi-cellular (EBC) phase (Figure 4.3 and 4.4). Stronger GFP expression is then seen in the male generative cell in the late bi-cellular (LBC) phase leading up to PMII (Figure 4.3 and 4.4). GFP is also only seen in the two male germ-line cells in the early tri-cellular (ETC) phase and also in the differentiated sperm cells of mature pollen (MP) grains (Figure 4.3 and 4.4). No vegetative cell GFP expression is seen throughout pollen development from either promoter construct. These similarities in expression of GFP during pollen development show that the *promMtrDUO1* behaves in the same male germ-line specific manner in pollen development as the *promAthDUO1* promoter. The developmental expression of *promAthDUO1:H2B-GFP* and *promMtrDUO1:H2B-GFP* was also similar to *promAthDUO1:DUO1-RFP* previously reported by Brownfield *et al*., 2009.

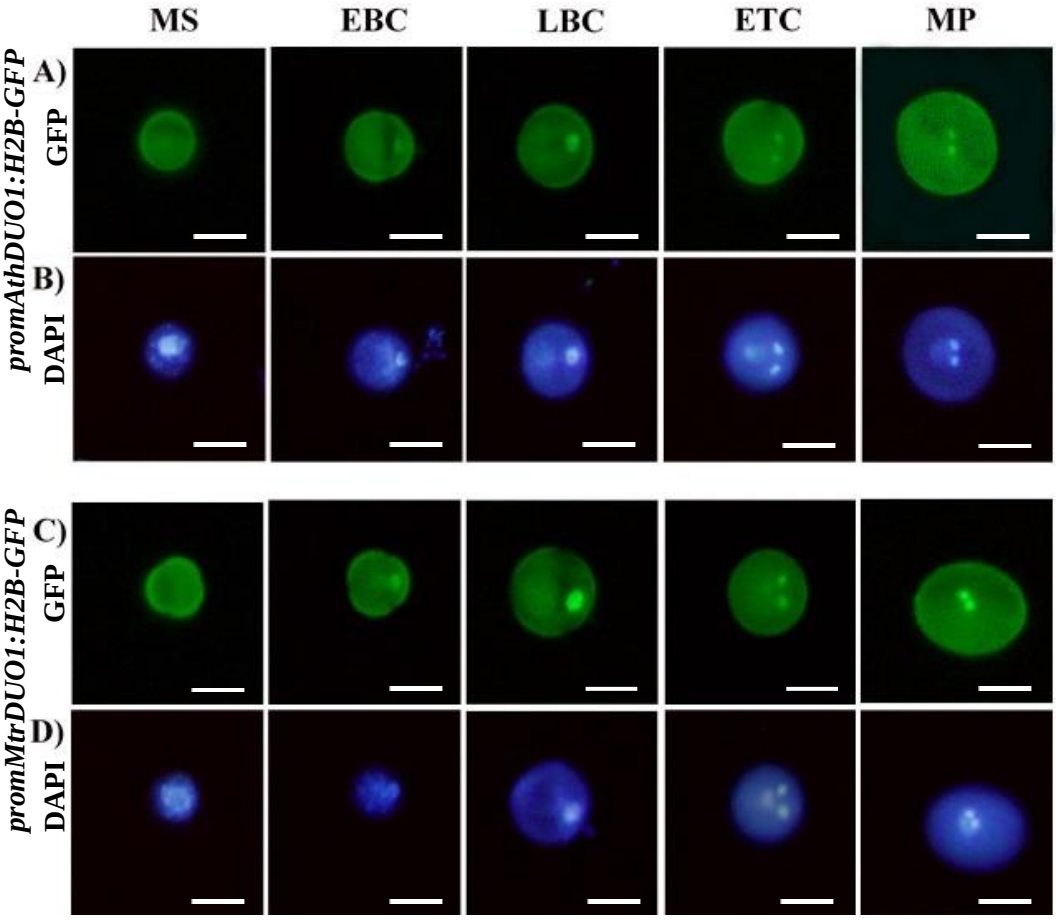


Figure 4.2: *AthDUO1* promoter and *MtrDUO1* promoter reporters during pollen development viewed by epifluorescence microscopy. Expression of *promAthDUO1:H2B-GFP* (A-B) and *promMtrDUO1:H2B-GFP* (C-D) throughout wildtype pollen development. Left to right show different stages of pollen development; Microspores (MS), early bi-cellular pollen (EBC), late bi-cellular pollen (LBC), early tri-cellular pollen (ETC) and mature pollen (MP). Epifluorescence microscopy was used to visualise GFP expression (A and C) and DAPI staining was used to visualise nuclear content to determine developmental stage (B and D). Auto-fluorescence of the pollen cell wall can be seen with green colouration of the whole pollen grain. Scale bar represents 12.5 µm.

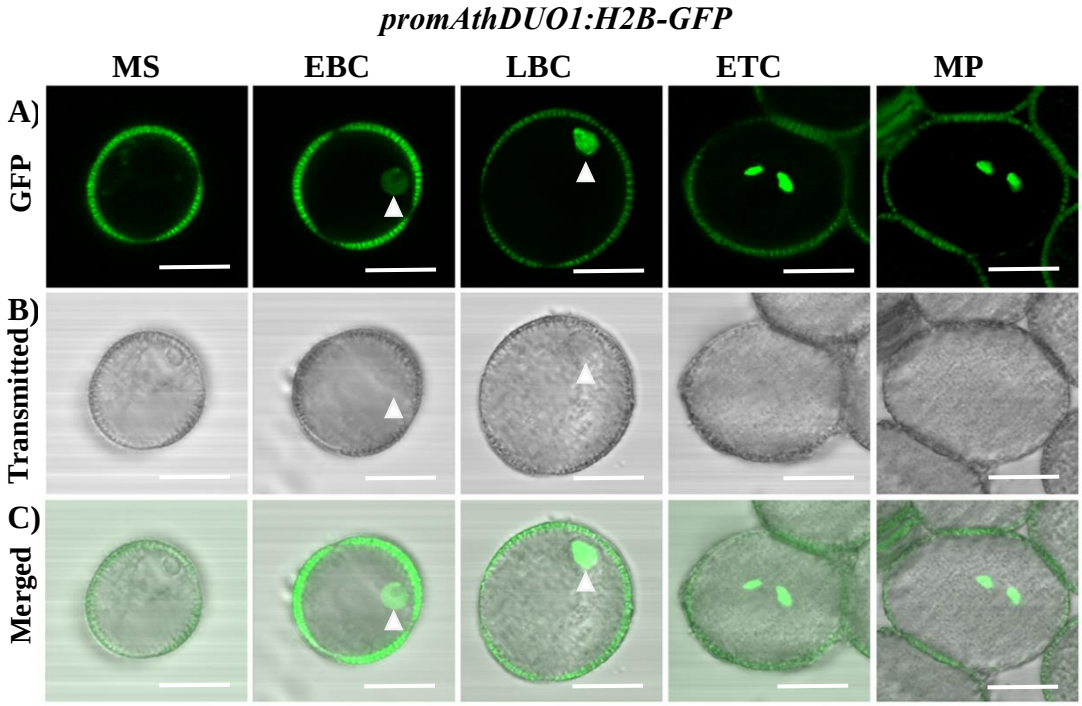


Figure 4.3 Activity of the *AthDUO1* promoter reporter during pollen development viewed by CLSM. Expression of *promAthDUO1:H2B-GFP* throughout wildtype pollen development. Left to right show different stages of pollen development; Microspores (MS), early bi-cellular pollen (EBC), late bi-cellular pollen (LBC), early tri-cellular pollen (ETC) and mature pollen (MP). **A)** GFP is first visualized in the generative cell (White arrowhead) of EBC pollen. GFP expression is maintained in the generative cell in LBC pollen, and also seen in ETC and MP pollen. Pollen cell wall auto-fluorescence is seen as an outer circle throughout development from MS - MP. **B)** Transmitted light images of developmental stages of pollen grains analysed in A. **C)** Merged images of GFP and transmitted light showing localisation of GFP in germ cells in EBC and LBC and in sperm cells in ETC and MP. As a range of signals were shown detector settings were not consistent across all images and a direct comparisons of intensity cannot be made between images. Scale bar represents 12 µm.

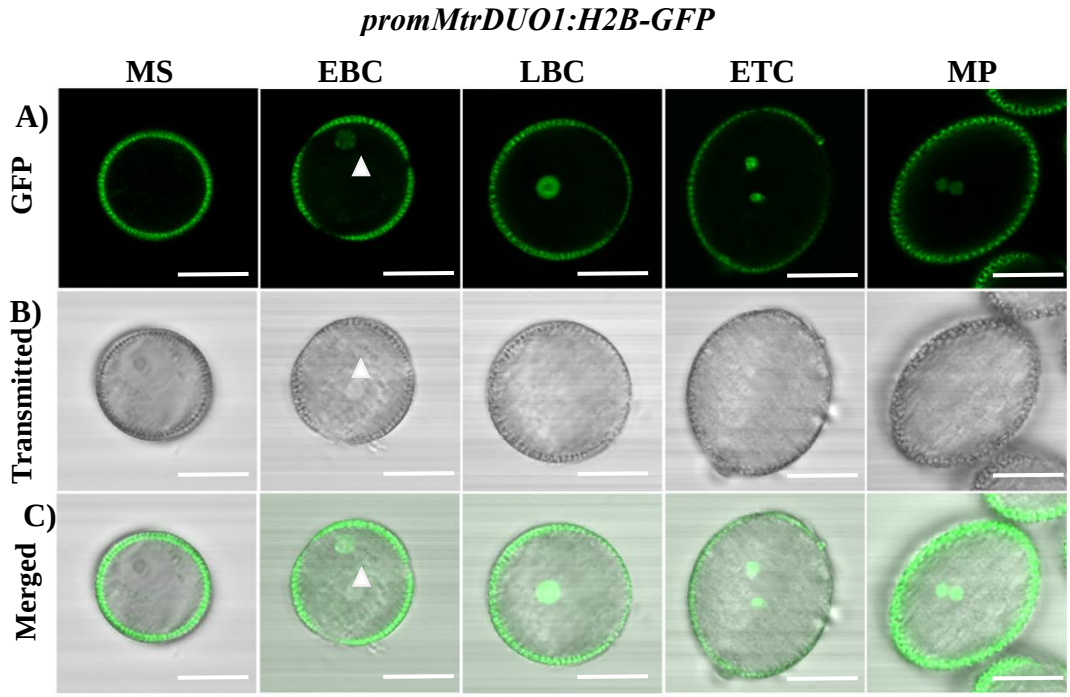


Figure 4.4: Activity of the *MtrDUO1* promoter reporter during pollen development viewed by CLSM. Expression of *promMtrDUO1:H2B-GFP* throughout wildtype pollen development. Left to right show different stages of pollen development; Microspores (MS), early bi-cellular pollen (EBC), late bi-cellular pollen (LBC), early tri-cellular pollen (ETC) and mature pollen (MP). **A)** GFP is first visualized in the generative cell (White arrowhead) of EBC pollen. GFP expression is maintained in the generative cell in LBC pollen, and also seen in ETC and MP pollen. Pollen cell wall auto-fluorescence is seen as an outer circle throughout development from MS - MP. **B)** Transmitted light images of developmental stages of pollen grains analysed in A. **C)** Merged images of GFP and transmitted light showing localisation of GFP in germ cells in EBC and LBC and in sperm cells in ETC and MP. As a range of signals were shown detector settings were not consistent across all images and a direct comparisons of intensity cannot be made between images. Scale bar represents 12 µm.

4.2.3 5′ Promoter deletion series of *M. truncatula* promoter identifies regions important for germ-line expression

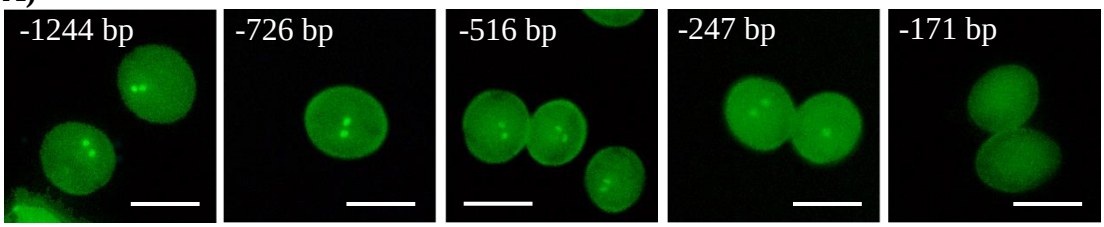
To determine important regions needed for expression of *promMtrDUO1* a 5′ deletion series was made with four varying size promoter fragments (Figure 4.1). One thousand mature pollen grains from 10 T1 *A. thaliana* transgenic plants (100 pollen grains each T1 line) from each of the promoter reporter constructs were used to measure the intensity of fluorescence in mature pollen sperm cells. The amount of GFP protein present is dependent on the transcriptional activity of the promoter region being analysed. This means higher levels of transcriptional activity lead to more GFP mRNA transcripts turning into protein leading to stronger fluorescence activity. Thus, the intensity of GFP provides an insight into the transcriptional activity of the promoter investigated.

In all cases GFP expression was male germ-line specific with fluorescence observed only in sperm cell nuclei and not in vegetative cell nuclei. The relative median fluorescence intensity was used to quantify expression and a two tailed Mann Whitney U test was used to test significance as it was previously shown that the AthDUO1 promotor deletion series had non-normal distributed fluorescence (Aidley, 2012) and the mean provides a poor measure of central tendency where the frequency distribution is skewed (Jones *et al*., 2007). Fluorescence intensity was then normalised to the *promAthDUO1* 1244 bp construct which was used as a positive control.

The longest *promMtrDUO1* 726 bp construct had the strongest relative median fluorescence intensity of the *M. truncatula* constructs (Figure 4.5). A drop in fluorescence intensity was then seen with the *promMtrDUO1* 516 bp and 247 bp constructs both having very similar relative fluorescence intensities and A two tailed Mann Whitney U test showed that the difference in fluorescence between the 516 bp and 247 bp constructs was not statistical significant (U-value 30, p-value 0.14156) but the difference between the 726 bp and 516 bp constructs was significant (U-value 0, p-value 0.00018; Figure 4.5). The *promMtrDUO1* 171 bp reporter construct had no GFP expression in pollen grains and no measurements were taken. Thus it was also not possible to perform a statistical test comparing the expression levels between the 171 bp and 247 bp constructs.

The varying length deletions showed there were two regions of promoter that affected the fluorescence intensity of *promMtrDUO1*. The first of these is the region between -516 and -726 bp which appears to act as an enhancer increasing fluorescence intensity in sperm cells. The second is the minimal promoter length needed to drive germ-line specific expression of -247 bp. The -247 bp fragment contains the conserved *ROD1* whereas the 171 bp fragment did not, which shows there is a likely conserved germ-line specific transcriptional function for *ROD1*.

A)



B)

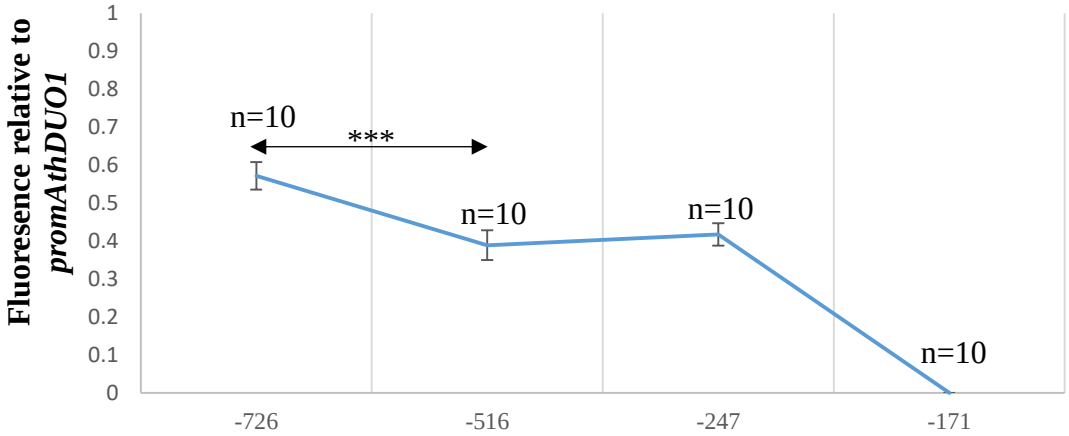


Figure 4.5: Fluorescence analysis of *A. thaliana* pollen sperm cell nuclei transformed with *MtrDUO1* 5´ promoter deletion reporter constructs. A) Images containing pollen grains expressing GFP from the promoter in sperm cell nuclei. Left to right; *promAthDUO1* -1244 bp, *promMtrDUO1 -726 bp*, *promMtrDUO1* -516 bp, *promMtrDUO1* -247 bp, *promMtrDUO1* -171 bp. Images were captured with an exposure of 142.8 milliseconds. Scale bar represents 25 µm. **B)** The relative median fluorescence of the *promMtrDUO1* -726 bp, *promMtrDUO1* -516 bp, *promMtrDUO1* -247 bp, *promMtrDUO1* -171 bp constructs are normalised to the relative median fluorescence of the *promAthDUO1* -1244 bp construct. A thousand pollen sperm cell nuclei fluorescence was measured from 10 individual lines (n=10) for each promoter reporter construct from Figure 4.1. Error bars show 95% confidence intervals for each construct. Deletions with significant differences to the one before are indicated by *** (indicates p-value <0.001, calculated by Mann-Whitney U test).

4.2.4 Construction of *ROD1* promoter reporter constructs

The *MtrDUO1* promotor deletion series showed that the minimal region needed to drive germ-line specific expression contained the evolutionarily conserved *ROD1*. *ROD1* sequences from the evolutionary diverse species *M. truncatula* (Eudicot) and *O. sativa* (Monocot) (Figure 3.1) were cloned to determine if the *ROD1* region from MtrDUO1 (*MtrROD1*, -247 to -150) and the promoter of OsaDUO1 (*OsaROD1*) was sufficient for germ-line specific expression, similar to that of *AthROD1* (-153 to -61). A nucleotide alignment of the *ROD1* sequences (Figure 4.6) cloned shows the conservation of the sequences and conserved CRE motifs maintained across these three evolutionary diverse plant species.

```
AthROD1 .. -AG----------AGAA-ACTTGGGGAGTGGG---GTACTGAAAGTGGAGGGTAATGGTAATGTGGGTGA
MtrROD1 .. -TG----------ACAA-ACT-----AGTGGA---GGAC-----GT----GGGCATGTTTAGTGGGGTGA
OsaROD1 .. GAGCGTACGCGCTAGAATTCT--CG--GTGGAACCGGACCCGGCGT--TCGTGGGGGCTGGTGGTGGCGA

AthROD1 .. -----TAACCGCATCCAACCGCCA----CTACAAAACCCAAAGAA-A------------
MtrROD1 .. GGCAGTAACTGC-TACAACCGCAAAAGTCTCCAAAA--CATGGAATAGTGTATGTGTGC
OsaROD1 .. GGCGGTTAGTCG-TGCCGCCGCGA-A---T-CAAACC-C---GAC-CGT---------C
```

Figure 4.6: Nucleotide alignment of cloned *ROD1* sequences from *A. thaliana*, *M. truncatula* and *O. sativa*. *ROD1* sequences cloned from promotors of DUO1 orthologues from *A. thaliana*, *M. truncatula* and *O. sativa* were aligned using MaliN multiple nucleotide alignment software. *Cis*-regulatory element motifs conserved are underlined as follows; AGAA motif is shown in red. GTGG motif is shown in yellow. AACCYG (Reverse complement TTGGC in *O. sativa*) motif is shown in green.

To test *ROD1* function promoter reporter constructs were made that contained four copies of the *ROD1* sequences linked together via restriction sites (Section 2.2.3) giving *MtrROD1x4* and *OsaROD1x4* (Figure 4.7). As a positive control *AthROD1x4* was also cloned. *ROD1x4* entry clones were then linked to *Min35S:H2B* and *GFP* through Gateway® cloning to create expression constructs (Figure 4.7). *Min35S* is a 46 bp sequence from the *CamV35S* promoter that acts as a minimal promoter region needed for transcriptional initiation complexes to bind allowing transcription. H2B then also allows nuclear localisation with the histone *H2B* sequence. The full length CamV35S and *Min35S* sequences have been shown not to have expression in pollen (Wilkinson *et al*., 1997; Zhang *et al*., 2016)

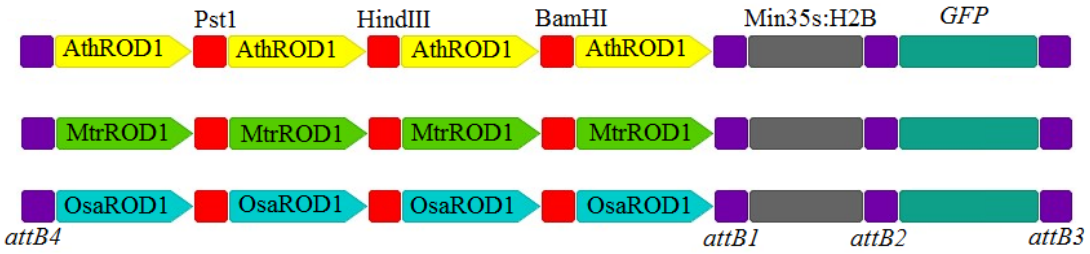


Figure 4.7: Schematic of *ROD1* tetramer promoter reporter constructs. The schematic shows the structure of three promoter reporter expression constructs. *ROD1* promoter sequences from *AthDUO1*, *MtrDUO1* and *OsaDUO1* have been linked together through restriction enzyme sites (Red) to form a *ROD1* tetramer (*ROD1x4* sequence). These *ROD1* tetramer promoter sequences drive expression of *Min35S:H2B* which provides a transcription start site along with nuclear localisation and GFP which provides visualisation of expression. Purple linker *att* sequences between the coding regions of *H2B* and *GFP* are a result of Gateway® recombination cloning which provides site specific positioning of DNA fragments

4.2.5 *MtrROD1* produces germ-line expression in *A.thaliana*

Analysis of the *AthROD1x4* and *MtrROD1x4* reporter constructs through epifluorescence (Figure 4.8) and CLSM (Figure 4.9, 4.10) microscopy showed that *AthROD1* and *MtrROD1* provide similar germ-line expression patterns. Epifluorescence expression analysis during inflorescence development for *AthRODx4* and *MtrROD1x4* showed that no GFP expression is seen in the microspore stage (Figure 4.8 A,C; MS) and that GFP is first visualised in the early bi-cellular stage (Figure 4.8 A, C; EBC) and is male germ-line specific from both *ROD1* reporters. Expression continues specifically in the male germ-line throughout development to mature pollen grains reflecting expression patterns of the full length promoters of *AthDUO1* and *MtrDUO1* (Figure 4.2). CLSM imaging of the *AthROD1x4* (Figure 4.9) and *MtrROD1* (Figure 4.10) constructs through pollen development showed the same expression patterns with increased sensitivity. This shows that the male germ-line specific expression of *ROD1* is conserved between *MtrROD1* and *AthROD1*.

ROD1x4 reporter constructs were also quatified to determine the difference in expression of GFP levels in mature pollen from *AthROD1x4*, *MtrROD1x4* and *OsaROD1x4* transformed plants (Figure 4.11). The median fluorescence units from ten

individual T1 lines harbouring the *AthROD1x4, MtrROD1x4* constructs had the GFP intensity measured in 100 sperm cell nuclei for each line (Figure 4.11 B). This showed that the median fluorescence intensity of the GFP between the two dicotyledonous species *AthROD1x4* and *MtrROD1x4* constructs was very similar and appeared to have maintained the same mechanism of regulation. (Figure 4.11 B). Over all this analysis of *MtrROD1* verse *AthROD1* shows that both sequences are necessary and sufficient for male germ-line specific regulation and likely share an evolutionary conserved mechanism.

4.2.6 *ROD1* in *O. sativa*

OsaROD1x4 showed no visible GFP expression in mature pollen of seven individual T1 lines. This could be due to a potential evolutionary difference in regulation of DUO1 between dicotyledonous (*A. thaliana* and *M. truncatula*) and monocotyledonous (*O. sativa*) plants. The *OsaROD1* sequence is less conserved than *MtrDUO1* to *AthDUO1* and lacks the two AACCG CRE motifs that potentially play a role in auto-activation.

To confirm the presence of the *OsaRODx4* promoter reporter construct in the analysed plants DNA was extracted from leave tissue. A PCR screen was then used to check the presence of the promoter reporter construct. DNA sequencing of the amplicon was used to confirm the sequence of the transformed reporter construct showing the correct reporter construct was present in the analysed plants.

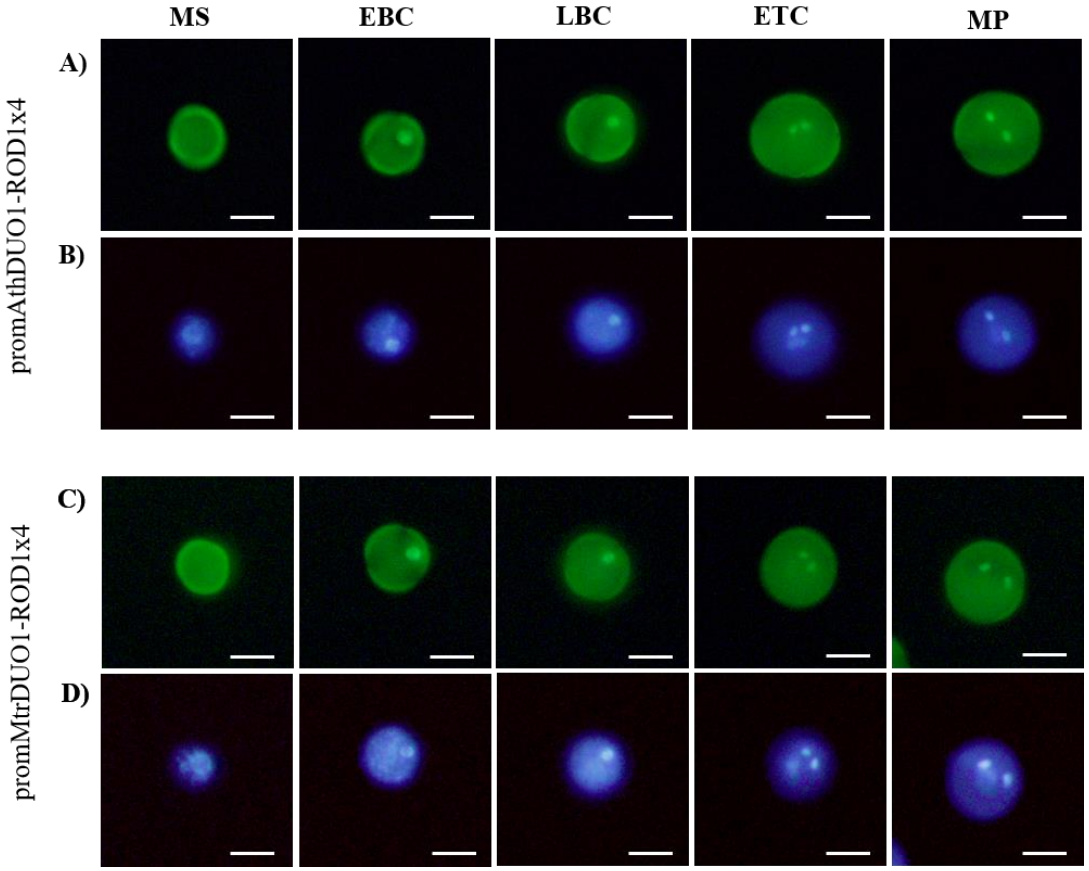


Figure 4.8: Activity of the *AthROD1x4* and *MtrROD1x4* promoter reporters during pollen development viewed by epifluorescence. Expression of *promAthROD1x4:H2B-GFP* and *promMtrROD1:H2B-GFP* throughout wildtype pollen development. Left to right show different stages of pollen development; MS, EBC, LBC, ETC and MP*. Epifluorescence microscopy was used to visualise GFP expression (A and C) and DAPI staining was used to visualise nuclear content (B and D). A-B) Isolated pollen grains from *promAthROD1:H2B-GFP* transgenic plants shows GFP expression in A and DAPI images in B. C-D) Isolated pollen grains from promAthROD1:H2B-GFP transgenic plants shows GFP expression in C and DAPI images in D. Both promoter reporter constructs (A and C) show similar expression patterns when visualised. No GFP expression is seen in the nucleus of the MS, GFP is first visualised in EBC after the asymmetric division and is maintained throughout LBC-MP. Scale bar represents 12.5 µm. * MS; Microspore, EBC; Early bi-cellular, LBC; Late bi-cellular, ETC; Early tri-cellular, MP; Mature pollen.

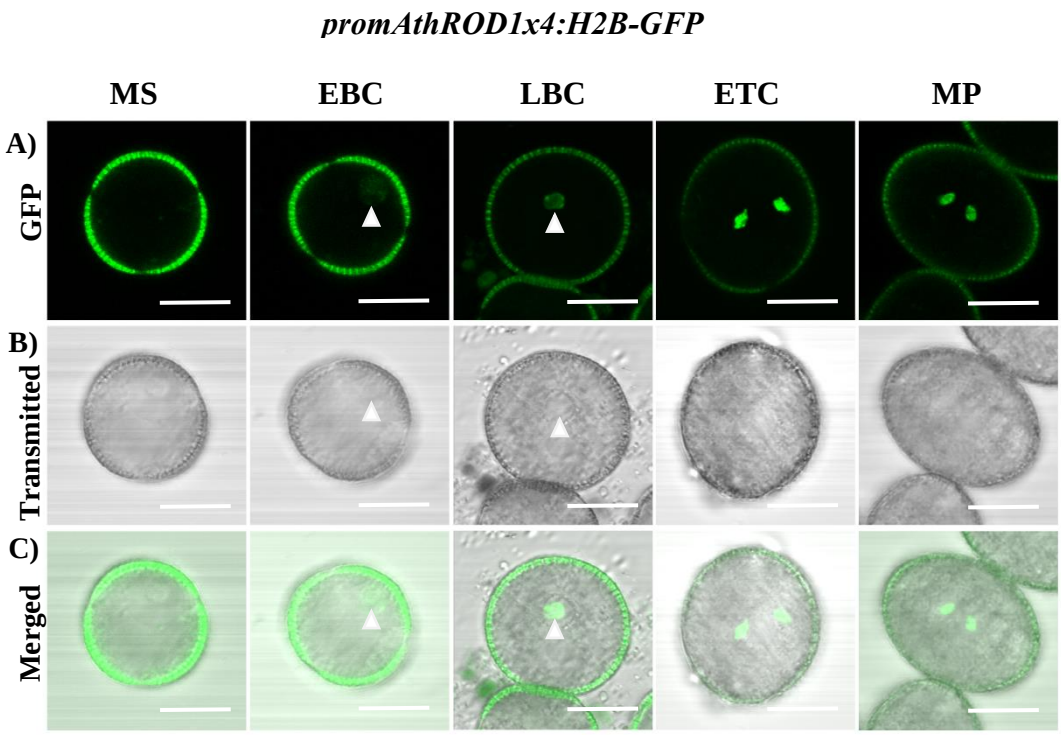


Figure 4.9: Activity of the *AthROD1x4* Promoter during pollen development viewed by CLSM. Expression of *promAthROD1x4:H2B-GFP* throughout wildtype pollen development. Left to right show different stages of pollen development; MS, EBC, LBC, ETC and MP*. **A)** Pollen cell wall auto-fluorescence is seen throughout development MS-MP. GFP is first visualized in the generative cell (White triangle) of EBC pollen soon after the asymmetric division of PMI. GFP expression is maintained in the generative cell in LBC pollen, and also seen in ETC and MP pollen in developing and mature sperm cells after PMII. **B)** Bright field images of developmental stages of pollen grains analysed in A. **C)** Merged images of GFP and bright-field showing localisation of GFP in germ cells in EBC and LBC and in sperm cells in ETC and MP. * MS; Microspore, EBC; Early bi-cellular, LBC; Late bi-cellular, ETC; Early tri-cellular, MP; Mature pollen. Scale bar represents 12.5 µm

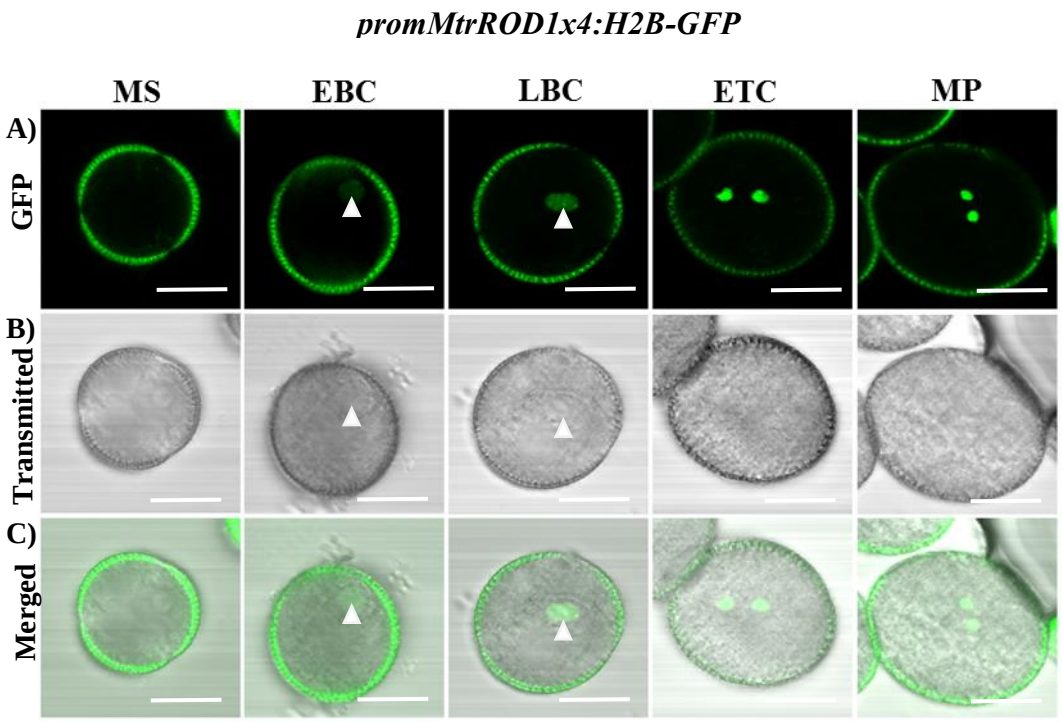


Figure 4.10: Activity of the *MtrROD1x4* promoter during pollen development viewed by CLSM. Expression of *promMtrROD1x4:H2B-GFP* throughout wildtype pollen development. Left to right show different stages of pollen development; MS, EBC, LBC, ETC and MP*. **A)** Pollen cell wall auto-fluorescence is seen throughout development MS-MP. GFP is first visualized in the generative cell (white triangle) of EBC pollen soon after the asymmetric division of PMI. GFP expression is maintained in the generative cell in LBC pollen, and also seen in ETC and MP pollen in developing and mature sperm cells after PMII. **B)** Bright field images of developmental stages of pollen grains analysed in A. **C)** Merged images of GFP and bright-field showing localisation of GFP in germ cells in EBC and LBC and in sperm cells in ETC and MP. * MS; Microspore, EBC; Early bi-cellular, LBC; Late bi-cellular, ETC; Early tri-cellular, MP; Mature pollen. Scale bar represents 12.5 µm

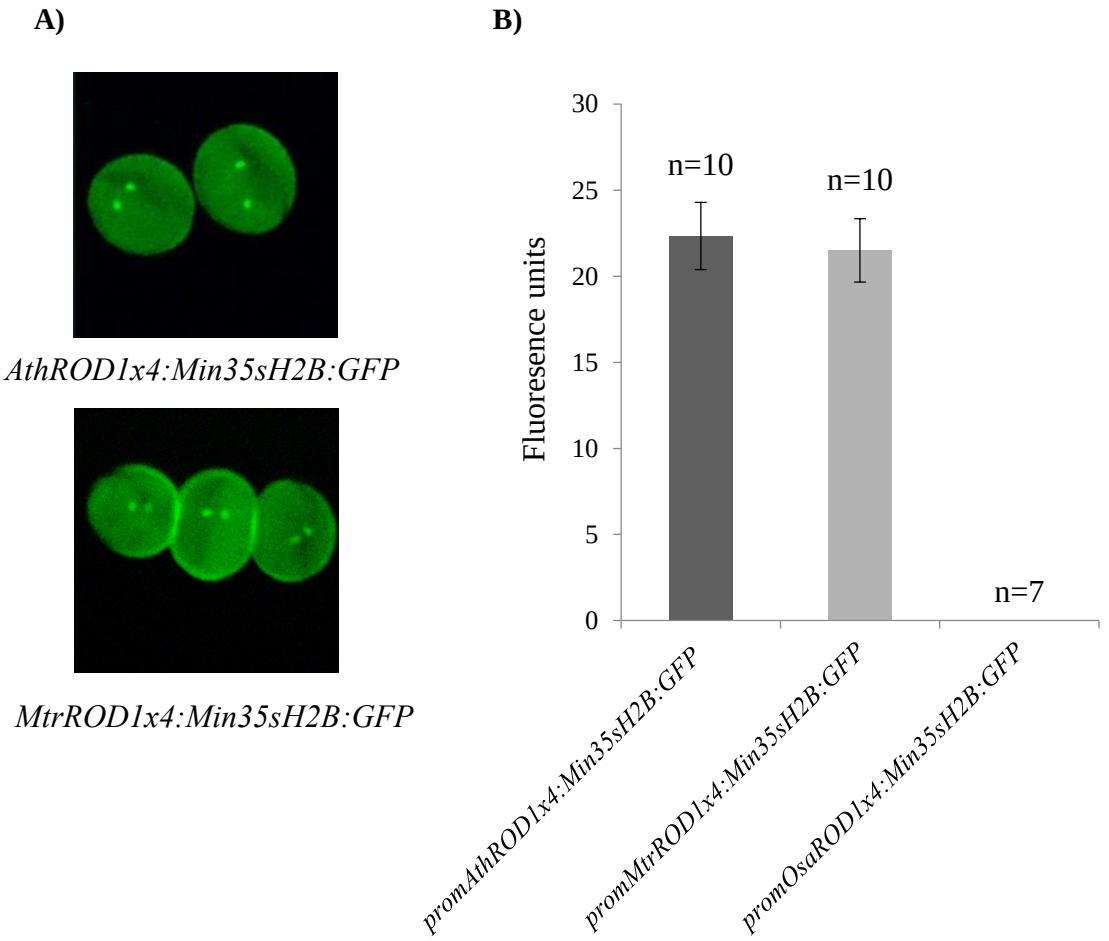


Figure 4.11: Fluorescence analysis of *ROD1x4* promoter reporter constructs. A) Images of mature pollen from *AthROD1x4* (Top) and *MtrROD1x4* (Bottom) expressing GFP specifically in the male sperm cells. Images were captured with an exposure of 142.8 m/sec **B)** Fluorescence quantification of *AthROD1x4*, *MtrROD1x4* and *OsaRODx4*. 1000 sperm cell nuclei was measured from 7 to 10 T1 lines of each promoter reporter. *AthROD1x4* and *MtrROD1x4* show similar levels of GFP expression and whereas no expression was seen from *OsaROD1x4*. Error bars indicate the standard deviation between the numbers of lines shown.

4.3 Conclusion

The results of this section show that the regulation of *promMtrDUO1* has been conserved with that of *promAthDUO1*. First *promMtrDUO1* was shown to have GFP expressed specifically in the sperm cells of pollen. Developing pollen grains were then analysed at different stages to see similar germ-line specific expression of GFP between the *promAthDUO1* and *promMtrDUO1* reporter constructs. A 5′ deletion series of *promMtrDUO1* showed there is a minimal region needed for sperm cell specific expression. A region of the promoter between -726 bp and -516 bp was also shown to enhance germ-line specific GFP expression similar to the *ROD2* region of the *A. thaliana* promoter. The *ROD1* sequence of *promMtrDUO1* was shown to be sufficient for germ-line specific GFP expression in pollen. Collectively this shows there is likely an evolutionarily conserved mechanism of regulation within *ROD1* of *A. thaliana* and *M. truncatula*.

Chapter 5.

Auto-activation of DUO1

5.1 Introduction

Auto-activation is the process in which a protein product can activate transcription of its encoding gene. Transcriptional auto-activation occurs when a transcription factor activates the expression of the gene encoding it by binding to the upstream promoter region. DUO1 is a R2R3 MYB transcription factor and has been characterised to upregulate target genes through binding to the CRE AACCG (Borg *et al.*, 2011). The same CRE motif is present twice within *ROD1* of the *DUO1* promoter from *A. thaliana* (Peters *et al*., 2017) and shows high sequence conservation across other diverse plant species (Figure 1.4), thus auto-regulation of DUO1 is possible. Consistent with this a yeast-one-hybrid analysis has shown that *A. thaliana* DUO1 can bind to its own promoter (Peters & Brownfield unpublished).

Another way to test an interaction between a transcription factor and a specific DNA sequence is a transient dual luciferase assay in plant leaf tissue. Luciferase based assays have been shown to provide a simple, quick and sensitive method for the study of promoter activation (Koncz *et al*., 1990) and been demonstrated to be effective in the analysis of pollen gene promoters (Bate and Twell, 1998).

Dual luciferase assays rely on infiltration of plant leaf tissue with *Agrobacterium* (*A. tumerfaciens*) carrying two suitable expression constructs. One construct contains an effector such as a transcription factor under control of a constitutive promoter. This expression construct is transiently expressed following infiltration in the plant leaf tissue. The other construct contains a reporter in which a promoter sequence is used to drive expression of a Firefly luciferase gene. Firefly luciferase is an enzyme which is capable of catalysing adenosine triphosphate (ATP) dependent oxidation of its substrate luciferin to produce light (Deluca and McElroy, 1978). If the effector binds to the promoter sequence and results in transcription there will be increased expression of Firefly luciferase with the amount quantified by measuring the emitted light. The reporter construct also contains a luciferase called Renilla which catalyses a luminescent reaction by utilising O_2 and coelenterate luciferin (coelenterazine)

(Mathews *et al*., 1977). Renilla is under the control of the Cauliflower mosaic virus (*CaMV-35S*) constitutive promoter and used as an internal control by normalising the activity of Firefly luciferase to that of Renilla (Sherf *et al*., 1996). This normalisation accounts for variation between transformation efficiency of different infiltrations. The amount of Firefly luciferase (FLuc) and Renilla (Ren) present in samples is quantified by analysing the amount of luminescence given off at the optimal wave length for each luciferase after the addition of their substrates. The amount of luminescence corresponds to the degree of interaction between the transcription factor (effector) and promoter (reporter) tested; High luminescence readings are a result of more firefly luciferase produced due to a strong interaction between transcription factor and promoter sequence tested. This in theory provides a quick way to study genes and genetic networks in plants.

Such an approach has been used once before to test the interaction of *A. thaliana* DUO1 and a 91 bp region of *promAthDUO1* (Aidley, 2011). This tested luciferase activities of promoter reporter constructs with a single AACCG MYB binding site, or with a mutated MYB site with a sequence of CATGA, which DUO1 is known not to bind (Borg *et al.*, 2011) were analysed in the presence of DUO1. Results of this showed that the construct with the native AACCG MYB site had a higher level of expression than the construct with the mutated MYB site. These assays however only tested the effect of one MYB binding site on *DUO1* auto-activation although there are two AACCG motifs in *ROD1* (Aidley, 2011) and were only conducted once.

This chapter thus aims to analyse the effect that both MYB CRE's have on *DUO1* auto-activation and test the evolutionary conservation of the potential auto-activation role of DUO1 in *A. thaliana, M. truncatula* and *O. sativa*. The *A. thaliana DUO1* promoter contains two AACCG sites where as *M. truncatula* contains one AACCG MYB binding site and one AACTG site and *O. sativa* has two sites with a reverse complement of the MYB binding site of CGGTT within *promOsaDUO1* (Figure 4.6). A dual luciferase approach similar to what has been used before was used to test promoter constructs from *A. thaliana, M. truncatula* and *O. sativa* where different combinations of the presence of the MYB CRE's were tested.

5.2 Results

5.2.1 Generation of expression constructs used in dual luciferase assays

To analysis the role of auto-activation of DUO1 effector and reporter constructs were made. Effector constructs contained either *CaMV-35S* or *promUB14* which are constitutive promoters driving a cDNA copy of *DUO1* (Table 2.10). As DUO1 contains a functional recognition site for the microRNA miR159 (Palatnik *et al*., 2007) a miR159 resistant version of *DUO1 cDNA* that had the miR159 recognition site mutated was cloned and termed *mDUO1* (Borg, 2011)(Table 2.10). *CamV-35S*:mDUO1 effector expression construct is shown in Figure 5.1, the *promUB14:mDUO1* vector was cloned after initial assays to try to optimise further assays (Discussed section 5.2.3).

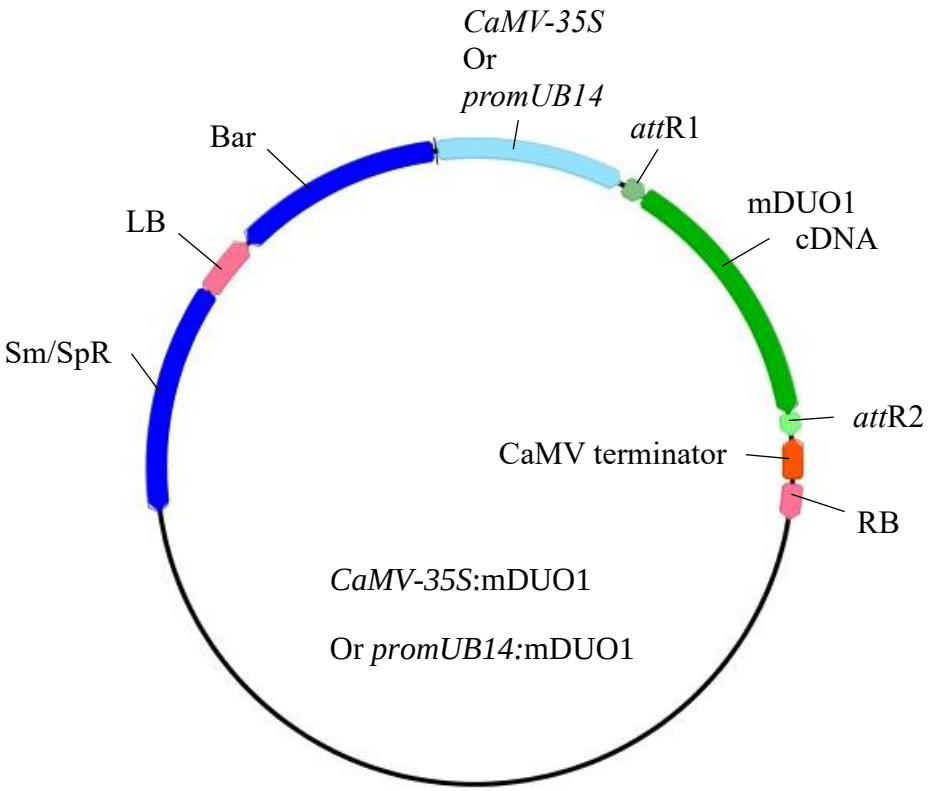


Figure 5. 1. Schematic of the vector design of effector constructs. The cDNA of *DUO1* or a mutated cDNA version (*mDUO1*) (shown) resistant to miRNA mDUO1 was inserted into the destination vector pB2GW7 through Gateway® Cloning. The *CaMV-35S* or *promUB14* promoter drives constitutive expression of *mDUO1* cDNA.

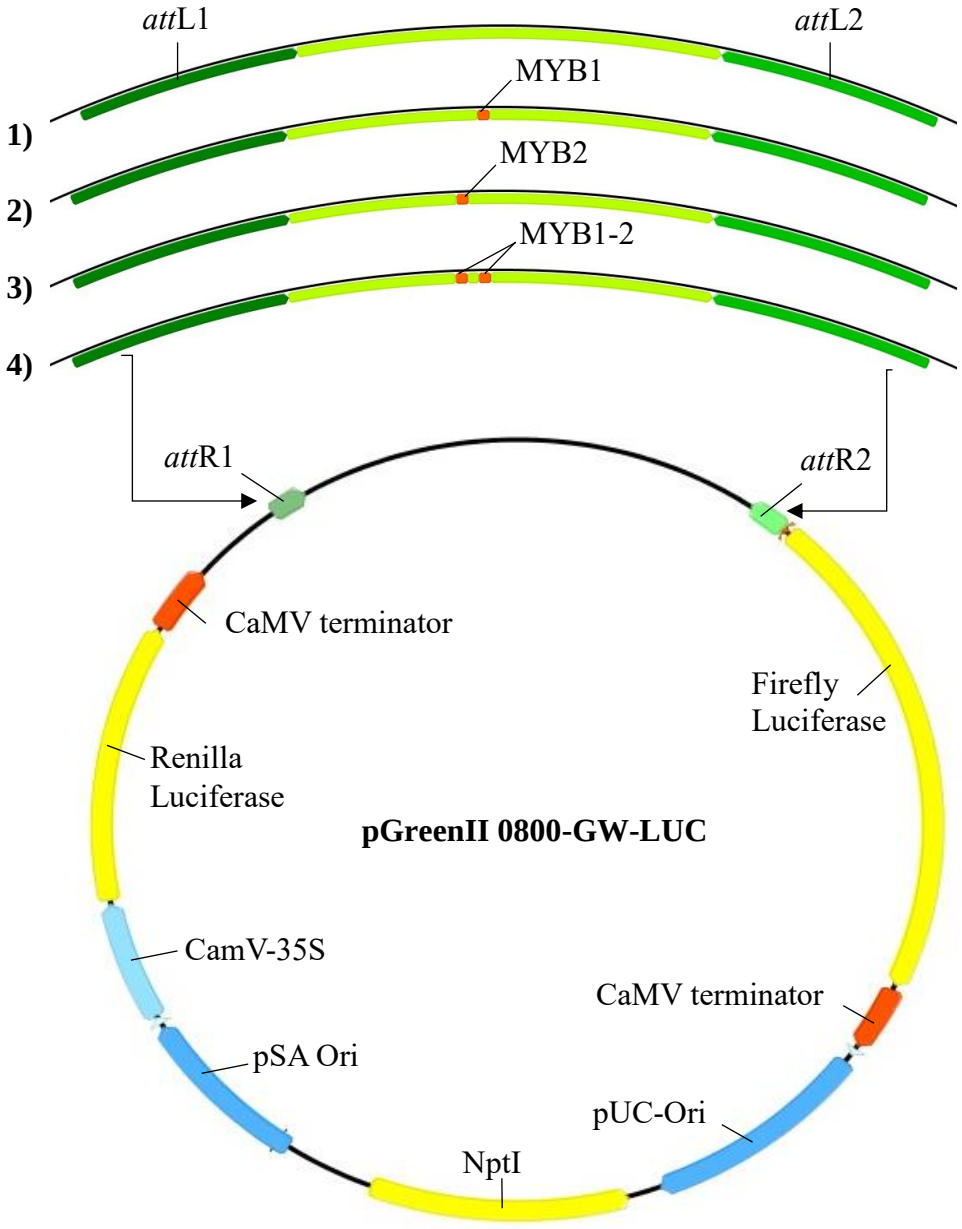


Figure 5.2. Schematic of the vector design of luciferase promoter reporter constructs. *DUO1* promoter sequences from *A. thaliana*, *M. truncatula* or *O. sativa* with *att*L sites either side are recombined via LR reaction into the destination vector pGreenII 0800-GW-LUC. 1) Unmodified promoter fragment of either *AthDUO1* (198 bp), *MtrDUO1* (243 bp) or *OsaDUO1* (308 bp). 2-4) *AthDUO1*, *MtrDUO1* or *OsaDUO1* with the MYB binding sites mutated to CATGA or TCATG for *O. sativa* are shown in orange. 2) MYB binding site closest to ATG termed MYB1. 3) MYB binding site distal from ATG termed MYB2. 4) Both MYB binding sites knocked out termed MYB1-2. pGreenII 0800-GW-LUC contains *att*R sites for recombination which are in-frame of Firefly luciferase with a CamV terminator. Renilla luciferase is constitutively driven by the *CamV35S* promoter.

Effector constructs contained promoter sequences from *A. thaliana*, *M. truncatula* and *O. sativa* with no mutations to the MYB binding sites (*AthDUO1* (198 bp), *MtrDUO1* (243 bp) and *OsaDUO1* (309 bp)) or with mutations. When the 5´ MYB binding site was mutated it was termed MYB1. When the 3´ MYB binding site was mutated it was termed MYB2, and mutation of both MYB binding sites was termed MYB1-2 (Figure 5.2). All fragments were cloned upstream of Firefly luciferase (Figure 5.2). MYB binding sites were mutated to the CATGA sequence previously described (Borg *et al.*, 2011) or the reverse complement TCATG for *O. sativa*.

Cloning of the fragments occurred two ways; Gateway® cloning and Restriction enzyme cloning (See Figure 5.8 for schematic diagram of the restriction cloning vectors). Gateway® cloning was used initially as the destination vector (pGreenII 0800-GW-LUC) backbone was ready available in the lab and compatible with PDONR vectors to easily form reporter constructs through LR reactions. Restriction enzyme cloning into the pGREEN 0800-5-LUC destination vector to create reporter constructs was then used to try optimise expression seen from reporters (Discussed in 5.2.4)

5.2.2 Dual luciferase assays using Gateway® effector constructs

The initial dual luciferase assays to test DUO1 auto activation used Gateway® cloned effectors and reporters. Infiltrated *N. benthamiana* leaf samples (Figure 5.3) were analysed for FLuc and Ren activity 3 days post infiltration with low luminescence levels detected from reporters in the presence of effectors e.g. *mDUO1*:*AthROD1*. Negative controls used were reporter constructs co-infiltrated with another vector constitutively expressing a protein (*35S*-GUS). The GUS protein is expected not to bind the *DUO1* reporter resulting in low (background) levels of luciferase activity. This negative control shows basal levels of activity of the reporter constructs as luciferase detected from negative controls may be due to native transcription factors present in *N. benthamiana* leaves binding to the *DUO1* promoter.

Luciferase measurements were analysed through Fluc/Ren ratios to account for transformation efficiency between different leaves by normalising to the internal *35S-Renilla* control. An increase in FLuc/Ren ratio can be seen when *AthDUO1* (198) was co-infiltrated with mDUO1 compared to being co-infiltrated with the negative control

35S-GUS (Figure 5.3). This shows the presence of mDUO1 results in activation of the *AthDUO1* (198) reporter. When looking at the fold enrichment change of FLuc/Ren of *AthDUO1* (198) in the presence of *mDUO1* verse *35S-GUS*, *mDUO1* resulted in a 3.7 fold increase (0.411/0.11). However, this is less than half of the fold change reported previously for the 91 bp *AthDUO1* promoter fragment tested in Aidley, 2011. It was observed that when *mDUO1*:*AthMYB1* (198) and *mDUO1*:*AthMYB1-2* (198) were co-infiltrated together these assays showed reduced levels of luciferase activity compared to *mDUO1*:*AthDUO1* (198) (Figure 5.3). Indicating the mutant MYB sites may influence DUO1 binding and auto-activation by reducing the luciferase expressed from this reporter.

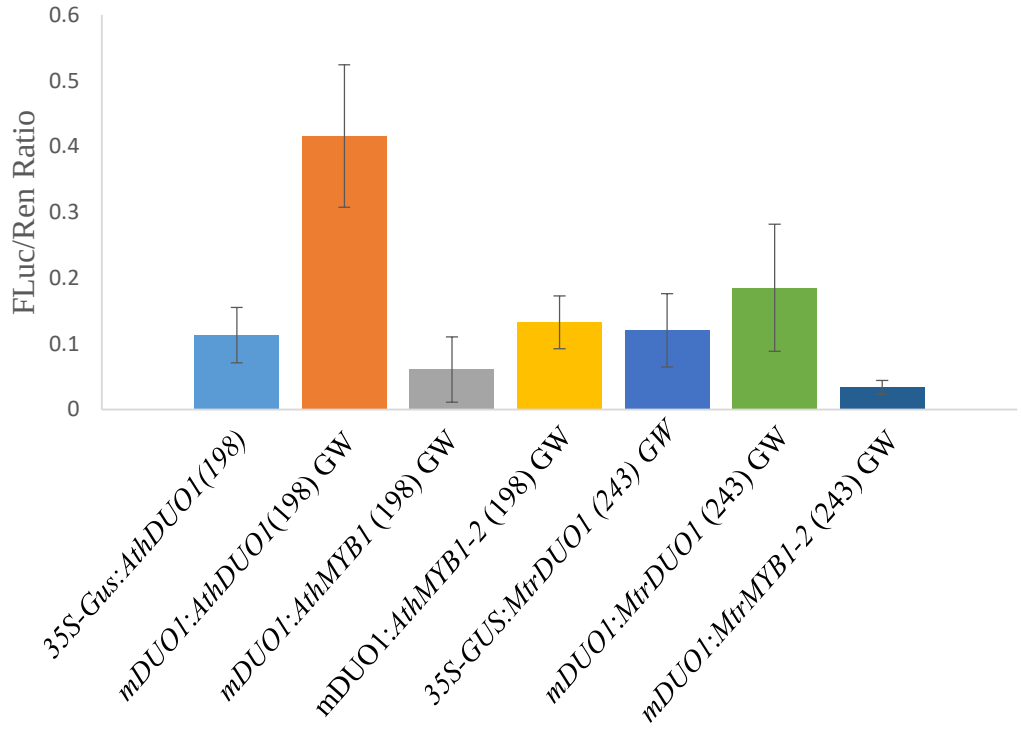


Figure 5.3 FLuc/Ren Ratio of Gateway® cloned reporter constructs. *N. benthamiana* leaf tissue infiltrated for 3 days with various combinations of expression constructs was measured for Firefly luciferase activity and Renilla. Resulting FLuc:Ren ratios are averages of four measurements taken from two individually infiltrated leaves (two disks from each leaf); (N=4). Error bars = S.E of the four samples.

The *MtrDUO1* (243) reporter also had a small level of activation in the presence of mDUO1 with a slight increase in FLuc/Ren activity over *35S-GUS:MtrDUO1 (243)* (Figure 5.3). *MtrMYB1-2* infiltrated with mDUO1 showed reduced luciferase expression compared to *35S-GUS:MtrDUO1 and* mDUO1:*MtrDUO1* showing that the mutant MYB sites reduced expression of luciferase.

5.2.3 Testing and optimising of controls for the dual luciferase assay

The first assays initially performed however had no known external positive control to confirm the dual luciferase assay was working. To identify that dual luciferase assays were functioning correctly a known effector and reporter construct pair with a described interaction identified by Helens *et al*., 2005 was used as a positive control for further assays (Figure 5.4). The effector construct contained the MYB75 transcription factor (PAP) and the reporter construct contained the promoter sequence from the gene At5g42800 (*DFR*) driving Firefly luciferase. Negative controls were also included in the assay where *A.tumefaciens* cells (GV3101) that either contain no effector or a 35S- GUS effector with no known interaction with *DFR* or

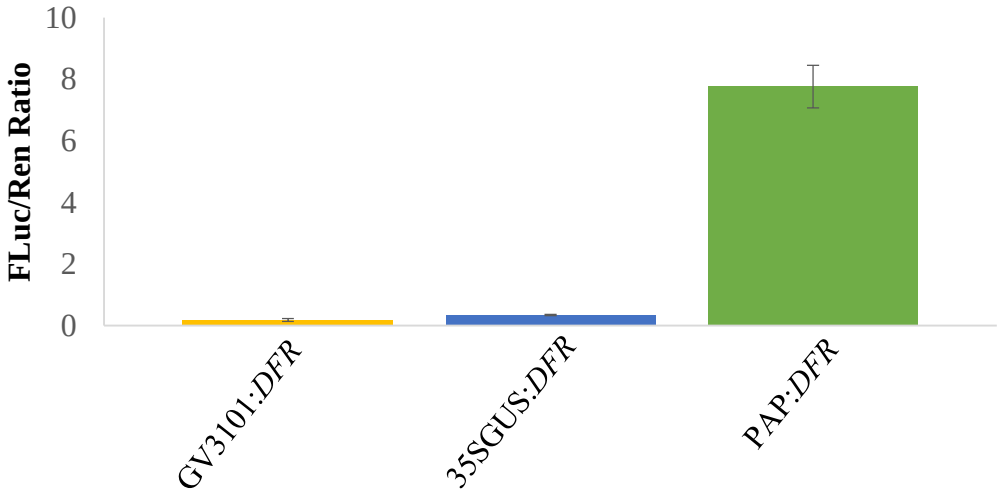


Figure 5.4. FLuc/Ren ratio from transiently expressed positive control PAP:*DFR* infiltrations. Dual luciferase controls are tested for FLUC/Ren ratio 3 days post infiltration. Negative control infiltrations contained GV3101 with no effector or 35SGUS was used as a non-binding effector to test interactions. Positive control tested for the system was PAP:*DFR*. Ratios are averages of four measurements from two individually infiltrated leaves (two disks from each leave); (n=4). Error bars = S.E of the four samples.

promDUO1(198) are infiltrated with these respective reporter constructs (Figure 5.4). The negative control infiltrations (Figure 5.4; GV3101:*DFR* and GUS:*DFR)* show the basal level of luciferase activity from the *DFR* reporter construct due to native transcription factors present in the leaf tissue that could activate the *DFR* reporter construct. PAP:*DFR* infiltrated plants had an increase in the FLuc/Ren ratio (Y axis) over plants infiltrated with the negative controls GV3101:*DFR* and 35S-GUS:DFR (Figure 5.4). This shows trans-activation of *DFR* occurring in the presence of PAP and confirmed that the dual luciferase system was working to test interactions between transcription factors and promoter sequences.

5.2.4 Redesign of DUO1 promoter reporter and effector vectors

While some activation was seen in initial assays reported, the level of activation was much lower than previously reported by Aidley, 2011 and was not able to be repeated in subsequent assays (Data not shown). This lead to a redesign of the effectors and the positive control reporters used in an attempt to improve the assays. The mDUO1 effector vector was redesigned with another strong constitutive promoter *promUB14* used to drive expression of mDUO1 replacing the *CamV35S* promoter (Figure 5.1). This was done to minimise any potential detrimental effect multiple *CamV35S* promoters (mDUO1 and Renilla both under control of *CamV35S*) may have within infiltrated leaves e.g. post transcription gene silencing (Vaucheret *et al*., 2001).

The reporter constructs were also redesigned and cloned into a new destination vector backbone (pGREEN 0800-5-LUC) using restriction enzyme (RE) cloning to allow the promoter of interest to be inserted upstream of Firefly luciferase creating *AthDUO1* (198) RE, *MtrDUO1* (243) RE, *OsaDUO1* (308) RE reporters. This was done as the Gateway® destination vector (pGreenII 0800-GW-LUC) used initially had been modified in-house (Lee, Unpublished) and not previously been used in luciferase assays whereas the pGreenII 0800-5-LUC has been used. It was noted with pGreenII 0800-GW-LUC that the distance between the terminator of the Renilla gene was particularly close to the promoters being tested; RE cloning allowed a larger separation.

Also cloned into the pGREEN 0800-5-LUC vector were new potential positive controls; the full length *promAthDUO1*(1244), *promAthROD1x4* and *promMGH3* (downstream target of DUO1; Brownfield *et al*., 2009). *promAthDUO1* (1244) sequence was used to determine if the length of the *DUO1* promoter influenced auto-activation (Shown to provide luciferase expression in Aidley, 2011). *promAthROD1x4* was used determine if multiple MYB binding sites (8 in total within the 4 joined *ROD1* sequences) would enhance luciferase activity through multiple DUO1 binding events. The *MGH3* promoter has been used in dual luciferase assays previously (Borg *et al*., 2011) and was shown to have a 350 fold increase in luciferase activity when in the

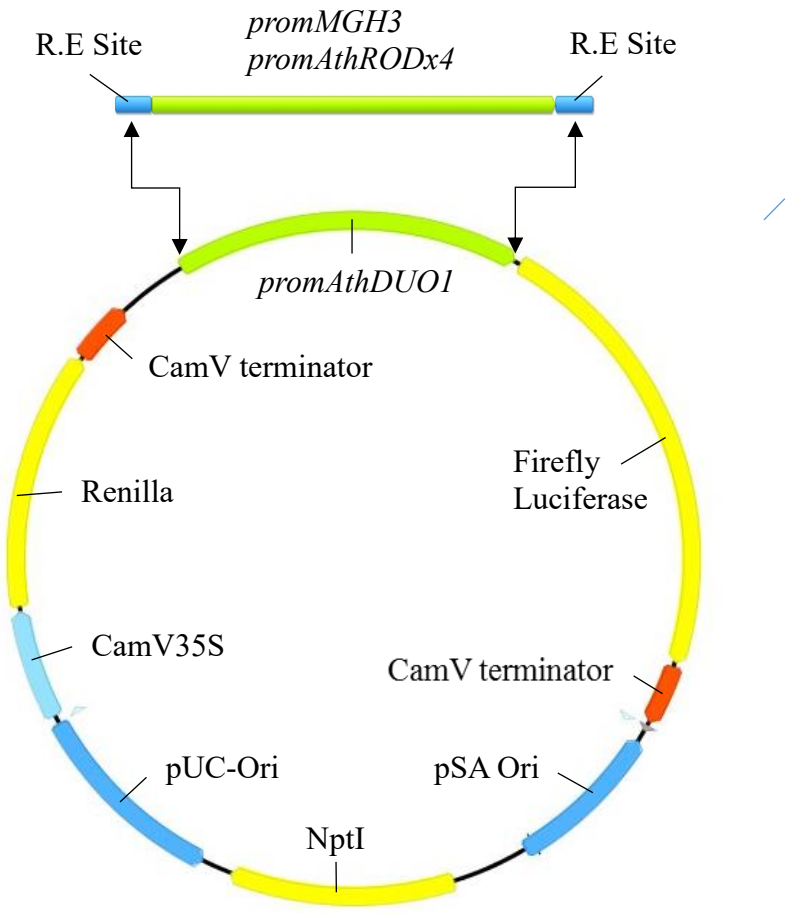


Figure 5.5: Schematic of the re-designed positive luciferase reporter controls. Three new positive reporters were made; *promAthDUO1* (1244) full length promoter, *AthROD1x4* (four *ROD1* promoter regions joined together) and *MGH3* full length promoter sequence (known downstream target of DUO1). All promoter sequences were restriction digest cloned into the pGREEN 0800-5-LUC destination vector backbone.

presence of mDUO1 showing a strong interaction. The *promMGH3* was thus included in further assays to confirm if mDUO1 was functional, as a strong interaction should occur between the mDUO1 effector and *MGH3* reporter.

The results of assays using the RE cloned reporter constructs showed the external system controls *PAP:DFR* had an interaction with an increase in Fluc/Ren ratio over GV3101:*DFR*. This positive interaction shows the dual luciferase assay system was functioning as previously described in other assays but at a reduced level (Fluc/Ren ratio of 7.7 from Figure 5.4 vs Fluc/Ren ratio of 1.6 from Figure 5.6). However when

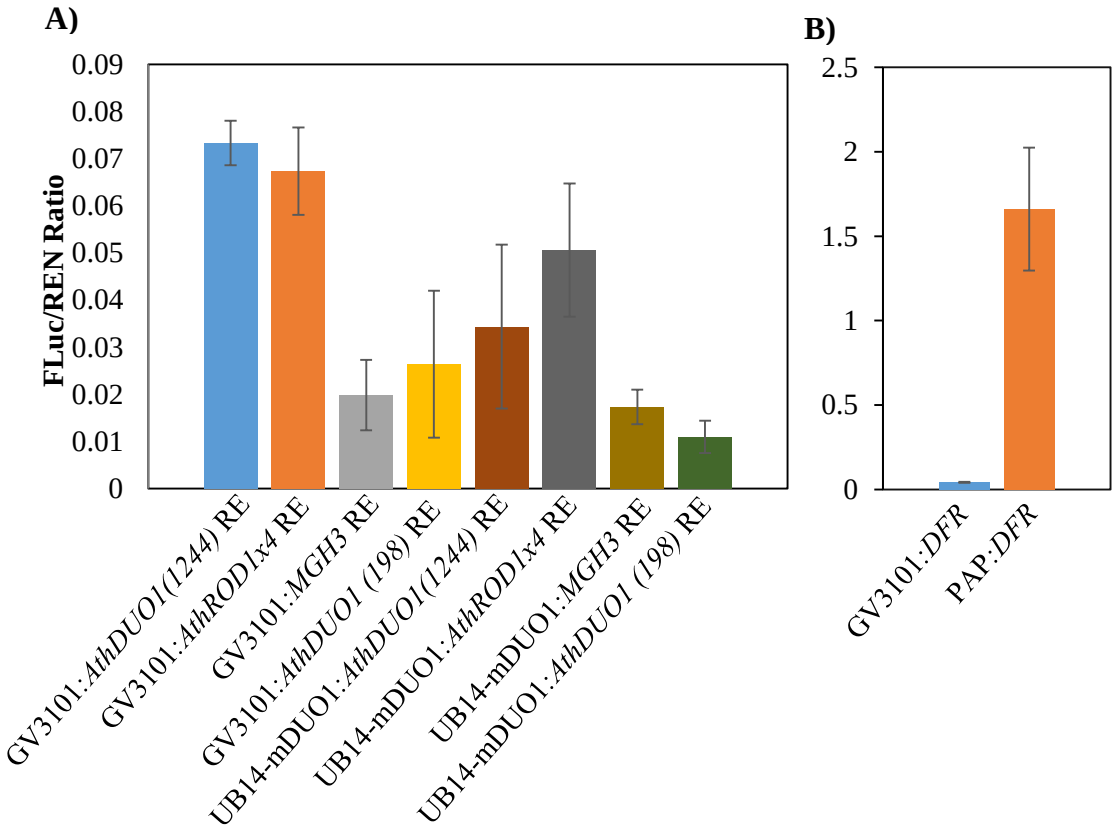


Figure 5.6: FLuc/Ren Ratio of restriction enzyme cloned reporter constructs.
A) Restriction enzyme cloned promoter reporters are tested for FLUC/Ren ratios. Negative control infiltrations contained reporter constructs co-infiltrated with GV3101 and no effector. **B)** Positive control tested for the system was PAP:*DFR*. All ratios are averages of four measurements from two individually infiltrated leaves (two disks from each leaf); (n=4). Error bars = S.E of the four samples.

analysing reporter constructs testing DUO1 binding the expected negative controls (GV3101:*AthDUO1* (1244), GV3101:*AthROD1x4*, GV3101:*MGH3*, GV3101:*AthDUO1* (198) had a higher FLuc/Ren ratio than the corresponding ratios from the reporter constructs infiltrated with the UB14:mDUO1 effector construct (Figure 5.6A). This shows the luciferase of these effector constructs was not being expressed by the binding of mDUO1. This assay was repeated once more with similar results seen (Data not shown).

5.2.5 Future ways to improve dual luciferase assays for DUO1 autoregulation

The results achieved were inconclusive and further time for modifcations and testing of the dual luciferase system are needed to confirm the role of autoregulation of DUO1. If more experimental time was available the following approaches may be used to eliminate any issues effecting the dual luciferase approach;

One issue effecting the efficiency of the transient expression of the constructs transformed via *Agrobacterium* is post transcriptional gene silencing (PTGS) which is a mechanism similar to RNAi that plants use as an antiviral defense (reviewed in Vance and Vaucheret, 2001). Silencing proteins such as the virus encoded P19 silcening protein can limit the amount of transiently expressed mRNA degraded by PTGS through P19 binding to siRNA's limiting degradation (Lakatos *et al*., 2004). The *Agrobacterium* strain in which all the reporter constructs were transformed into was thought to contain a plasmid which encoded the P19 silencing protein. However, when the *Agrobacterium* strain was PCR screened for the P19 plasmid by another lab member the plasmid was not present. Co-transforming all reporter constructs with the P19 plasmid into the suitable *Agrobacterium* strain (GV3101) and re-doing the assays may reduce any PTGS occuring from the transient epression of the effector and reporter constructs used in the dual luciferase assays.

Another experiment to check transient expression would be to perform a time course RT-PCR series with primers designed to detect DUO1 transcripts from cDNA isolated at different time points of infiltrated leaf tissue. This would allow identification of whether or not DUO1 is being transiently expressed and the peak time of transient expression from the effector constructs (*CamV35S*:*mDUO1* and *UB14*:*mDUO1*).

Results from such an experiment will show if the effector construct is functional and when peak expression of mDUO1 is occuring and time the dual luciferase measurements accordingly from these results.

Varying the amount of the *Agrobacterium* carrying the mDUO1 effector used when infiltrating leaf tissue may also influence *N. benthamiana* leaves wellbeing as ectopic expression of mDUO1 has been seen to result in cell death and growth defects (Brownfield *et al*., 2009; Palatnik *et al*., 2007). This is due to mDUO1 lacking the mi159 binding site making mDUO1 resistant to miRNA regulation of gene expression. As DUO1 is a transcription factor accumulation of the mDUO1 protein may then influence gene expression resulting in senesence of the tissue infiltrated. Another R2R3 MYB has also been seen to be involved in plant leaf senesences resulting in cell death when infiltrating plant leaves (Gao *et al.*, 2017). Thus, finding an *Agrobacterium* bait/prey ratio in which the mDUO1 effector doesn't result in cell death but can regulate the expression of reporter constructs could also improve results.

5. 3 Conclusions

The varying measurements taken through the dual luciferase approach to study the potential auto-regulation role of DUO1 have been inconclusive. The approach adopted replicated what has been shown previous (Aidley, 2011) and building on the information already known. A 3.6 fold enrichment of Fluc/activity was achieved with AthDUO1(198) in the presence of mDUO1 (Figure 5.3) and fold enrichment of FLuc/Ren was reduced in the mutated MYB reporter AthMyb1 and AthMyb1-2 infiltrations. However sample size was small and these results were unable to be replicated. Assays conducted in this study could not confidently replicate what was been previously shown (Aidley, 2011). Numerous factors could be effecting the results obtained and a process of elimination of issues was used to try improve assay measurements. However further time and experiments outside of those described in this thesis are needed to confidently confirm the role of autoregulation of DUO1 via dual luciferase assays.

Chapter 6.

Discussion.

The DUO1 R2R3 MYB transcription factor is a key regulator involved in switching on the male germ-line developmental program in pollen. DUO1 expression after the asymmetric PMI division results in correct division at PMII and differentiation of the resulting daughter cells to form mature sperm cells for double fertilisation (Durbarry *et al*., 2005; Rotman *et al*., 2005). Studies thus far have focused on DUO1 function and regulation in the model organism *A. thaliana* (Durbarry *et al*., 2005; Rotman *et al*., 2005; Brownfield *et al*., 2009; Borg *et al*., 2011). In this thesis, analysis of DUO1 function and regulation extended the understanding of DUO1 into the model legume *M. truncatula*. This showed DUO1 has evolutionarily conserved its male germ-line regulatory function and maintained similar mechanisms of regulation across evolutionary diverse plant species.

6.1 Evolutionary conservation of the male germ-line regulator DUO1

DUO1 was first identified to be an R2R3 MYB transcription factor by Rotman *et al*., (2005). R2R3 MYB transcription factors are specific to plants and one of the most abundant type of transcription factor in plants, with 126 R2R3-MYBs identified in *A. thaliana* genome (Yanhui *et al.,* 2006). It was initially noted by Rotman *et al*., (2005) that DUO1 contained a supernumerary lysine at position 66 (K66) within the R2R3 MYB DNA binding domain that other known plant R2R3 MYB transcription factors lacked (Martin and Paz-Ares. 1997; Stracke *et al*., 2001).

From identification of this supernumerary lysine Rotman *et al*., (2005) also then described putative DUO1 orthologues containing K66 in tobacco, *Zea mays* and *Oryza sativa,* initially suggesting K66 as a signature for the DUO1 family. Borg (2010), then went onto describe additional putative DUO1 orthologues containing K66 across 19 divergent plant species. This included higher vascular plant species through to lower plant species such as non-vascular mosses. The higher plant species (dicots and monocots) appeared to contain the identity and spacing around K66 more so than the low plant species which contained additional amino acids around the location of K66.

This spacing differences suggests through the course of evolution these additional amino acids have been lost whilst K66 has been retained since orthologues from all other species have conserved this signature feature. The steric effect K66 contributes to the confirmation and DNA binding capacity of the DUO1 MYB DNA binding domain may thus be tailored for conserved transcriptional regulation of DUO1 target genes in the male germ-line of diverse flowering plant species (Borg, 2010).

While previously published studies focussed on sequence similarity of DUO1 orthologues containing the supernumerary lysine there has been little focus on testing functionality of this conservation of DUO1 across various flowering plant species. However DUO1 orthologues from rice (OsaDUO1) and tomato (SlyDUO1) have been shown to be functionally equivalent to that of AthDUO1 in complementation studies conducted in *DUO1/duo1 A. thaliana* lines (Twell, unpublished).

Results described in Chapter 3 of this thesis built on the understanding of the functional conservation of DUO1 from the legume *M. truncatula* which is evolutionarily distinct from *A. thaliana* (Figure 3.1). *M. truncatula* contains a single DUO1 orthologue that was identified to contain the conserved K66 lysine (MtrDUO1) (Section 3.2). Complementation testing showed that MtrDUO1 reduced the frequency of *duo1* pollen grains with the mutant mitotic division phenotype at PMII and could restore male transmission of the *duo1* allele. Confirming that the non-native MtrDUO1 orthologue, could activate the same downstream cell cycle pathways and genes necessary for specification of the germ-line cell into sperm cells as AthDUO1 in *A. thaliana*. Collectively, this shows a conserved regulatory function of the male germ-line regulator DUO1 across two evolutionary divergent eudicots and suggests K66 in the R2R3 MYB DNA binding domain is important for function. Understanding the potential pathways these DUO1 orthologues regulate may help further elucidate the underlying male germ-line regulatory mechanisms these plants may share.

This current study has tested function of MtrDUO1 in *A. thaliana* confirming a conserved role. However to further investigate the function of DUO1 from *M. truncatula* a number of experiments using forward/reverse genetics using *M. truncatula* plants could build on the information established from this thesis. When the complementation experiments were designed and carried out a *duo1* knockout

M. truncatula line was not available. Now with the increasing number of *M. truncatula* Tnt1 retrotransposon lines readily available (Cheng *et al*., 2011), a *duo1* knockout would be useful to conduct addition functional tests. First identification of a mutant phenotype at PMII and testing male transmission of the *Mtrduo1* allele would confirm if MtrDUO1 is likely to regulate the male germ-line in *M. truncatula*. Testing the ability of AthDUO1 (or other plant DUO1 orthologues) to then complement a *DUO1/duo1 M. truncatula* line would further build on the understanding of the conserved pathways of the male germ-line regulation in plants.

Since it is now established DUO1 proteins with the K66 supernumerary lysine are functionally conserved (AthDUO1 and MtrDUO1), future research could address which amino acids or conserved domains are required for DUO1 to function as a male germ-line regulator. An early target would be the K66 supernumerary lysine and removing or modifying K66 via site directed mutagenesis. This would identify the importance of K66 in overall DUO1 protein structure and function by determining if K66 is necessary for regulating DAT genes in the male germ-line across various plant species. Similar work has been conducted by Borg (2011) who transiently expressed AthDUO1 truncations in *DUO1/duo1 A. thaliana* lines identifying the acidic carboxyl-terminus as the transactivation domain of the *A. thaliana* DUO1 protein. Testing other domains or truncations of the DUO1 protein would also be of interest to improve overall understanding of the evolution of DUO1 protein as a transcriptional regulator for the male germ-line.

Additional experiments such as transcriptome analysis on a *M. truncatula DUO1/duo1* line vs WT could help identify conserved pathways and DAT genes shared with *A. thaliana* male germ-line regulation along with any potential differences. These experiments would further improve understanding of the genetic networks necessary for male germ-line development in flowering plants and help determine if MtrDUO1 is the master regulatory involved in male germ-line formation of *M. truncatula* as is the case for *A. thaliana*.

6.2 Functional conservation of *ROD1* determines germ-line expression of DUO1

While it is known that expression of AthDUO1 is restricted to the male germ-line during pollen development and the regulation of AthDUO1 has been shown to be transcriptional (Brownfield *et al*., 2009), the regulation of DUO1 from other species is also largely unknown. The work described in chapter 4 shows there is a conserved mechanism that regulates germ-line specific expression between evolutionarily diverse eudicot species with *promMtrDUO1* and *promAthDUO1* resulting in identical germ-line specific promoter reporter expression patterns throughout pollen development (Figure 4.3 & 4.4). This shows that like *AthDUO1*, *MtrDUO1* is also likely to be transcriptionally regulated as the promoter sequence alone can drive male germ-line specific expression. MtrDUO1 like AthDUO1 is also thus likely to be regulated by an unknown asymmetrically inherited transcription factor at PMII that controls DUO1 expression.

The 5´ deletion series of the *promMtrDUO1* and isolation of *MtrROD1* in promoter report constructs identified that *MtrROD1* sequence acts as a *cis*-regulatory module necessary for germ-line specific expression in pollen like that from *AthROD1*. The high sequence similarity shown between that of *AthROD1* and *MtrROD1* (Figure 4.6) (Peters *et al*., 2017) suggests there is selective pressure to retain this *cis*-regulatory module for DUO1 germ-line expression. This is likely due to the tight location and timing of DUO1 expression in the male germ-line being critical for reproductive success of flowering plants maintaining a selective pressure against mutations occurring to preserve the *ROD1* cis-regulatory module in flowering plants.

The *ROD1* module contains a number of putative CRE's which have varying levels of conservation between *AthROD1* and *MtrROD1* plants (Figure 4.6) and other dicotyledonous plants (Figure 1.4). The distal GAGA motif identified in *Arabidopsis* accessions for DUO1 is not as well conserved between *A. thaliana* and *M. truncatula* with a motif of GACA present in *MtrROD1* (Figure 4.5). The *AthROD1* GAGA motif has been shown not to be necessary for germ-line specific expression via site directed mutagenesis of this motif in reporter constructs with germ-line expression still observed (Zohrab, 2015). Thus, the expression seen from the mutated GAGA

AthROD1 reporter along with relatively poor conservation is consistent with this CRE unlikely to be critical for male germ-line specific expression of DUO1.

Three repeats of a GTGG motif present in *AthROD1* have however been shown to be vital for germ-line specific expression through site directed mutagenesis of *AthROD1* reporter constructs (Peters *et al*., 2017). The mutagenesis of these sites resulted in no reporter expression (germ-line or vegetative cell specific in pollen) being observed. The GTGG motifs are also likely to be crucial for *MtrROD1* expression as these sites are highly conserved between *AthROD1*, *MtrROD1* and various other dicotyledonous plant species (Figure 1.4; Figure 4.5) (Peters *et al*., 2017). Additional experiments e.g. using site directed mutagenesis of the repeated GTGG sites in an *MtrROD1* reporter construct would confirm if these sites are crucial for *MtrROD1* germ-line specific expression.

There are also two conserved AACYGY motifs present in ROD1 likey to be involved in auto-activation of DUO1 (See section 6.3 below for further discussion).

Collectively the conserved ROD1 CRE's show there is high selective pressure to maintain certain motifs (3 GTGG and 2 AACYGY repeats) within the ROD1 *cis*-regulatory module critical for correct transcriptional control of DUO1 in pollen development.

6.3 Auto-activation of DUO1

The two proximal AACYGY motifs of *AthROD1* and *MtrROD1* are present in all dicotyledonous plants with at least one site containing the AACCG DUO1 binding site (Figure 1.4) (Peters *et al*., 2017). These AACYGY motifs give the potential for auto-activation to occur with evidence of DUO1 binding the AACCG motif (and or AACYGY) to form a positive feedback loop to enhance expression (Peters et al., 2017).

Zohrab (2015) showed that site directed mutations of the AACYGY motifs in fluorescent reporter constructs reduced levels of reporter expression. This suggests that male germ-line specific expression of DUO1 isn't dependent on the AACYGY motifs but instead the motifs likely enhance expression of DUO1. The AACYGY repeats

have potentially evolved due to redundancy and ensure the important male germ-line regulator DUO1 is correctly expressed with sufficient levels in pollen developmental processes after PMI. The high conservation of the AACYGY motif across dicots suggests the auto-activational role is maintained across these diverse species.

This thesis thus intended to determine if the auto-regulatory role of these AACYGY motifs is conserved across evolutionarily diverse species. The aim was to test the role of each motif site individually and in combination from *AthROD1*, *MtrROD1* and *OsaROD1* using a dual luciferase system to describe the specific role each site played in DUO1 expression. *MtrROD1* was thus of interest as it contained one AACCG DUO1 binding site and one AACTG site and a dual luciferase approach was thought to be sensitive enough to identify any differences in expression from the AACYGY motif; E.g. does DUO1 binding at a particular AACYGY site effect expression more so than the other site.

However the dual luciferase approach to test auto-activation was troublesome with inconsistent results obtained and possible improvements to the dual luciferase assays were discussed in Section 5.2.5. An alternative approach of creating stable transgenic *A. thaliana* lines with *MtrROD1* fluorescent promotor reporter constructs with various mutations of the above motifs could be used to confirm the transcriptional function of the CRE's within *MtrROD1*. This approach however wasn't chosen as creating stable *A. thaliana* mutant lines is time consuming and is the same technique as used in Chapter 3 and 4. The dual luciferase assays used different laboratory techniques and in theory could provide quicker results for interactions between a protein and DNA target sequence.

6.4 Conclusions

The DUO1 R2R3 MYB transcription factor is a conserved regulator that is a master switch for male germ-line differentiation and specification in flowering plants. The high conservation of the DUO1 protein across dicotyledonous plants and the ability of MtrDUO1 to complement the *A. thaliana duo1* mutant shows MtrDUO1 is a functional orthologue that has been evolutionarily maintained across diverse plant species.

The transcriptional regulation of DUO1 has also been functionally maintained across these diverse plant species with the highly conserved *cis*-regulatory module *ROD1* present in dicotyledonous plants. Here *MtrROD1* has been shown to be functionally equivalent to that of *AthROD1* and provides identical male germ-line expression in pollen development and mature sperm cells. Three different CRE's are highly conserved within ROD1; GAGA, GTGG, AACYGY. The current model proposed suggests GTGG motifs are vital for germ-line specific expression although what binds this motif to regulate DUO1 is still unknown. The AACYGY tandem repeats are then likely responsible for auto-regulation of DUO1, however further work needs to clarify the individual role of these sites.

The increased understanding of DUO1 function and regulation in key crop plants such as legumes is of importance as this provides approaches to manipulate plant fertility through the male germ-line. Providing options to either try increase crop yields through higher efficiency fertilization and increased seed production or on the other hand prevent spread of unwanted species through creation of infertile pollen or aborted seed production.

References

Aidley, J. (2012). Hunting for cis-regulatory elements in the *Arabidopsis thaliana* DUO1 promoter. University of Leicester Department of Genetics. MSc Thesis.

Arnosti, N. and Kulkarni, M., (2005). Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *Journal of Cellular Biochemistry*, 94(5), 890-898.

Barker, D., Bianchi, S., Blondon, F., Datte, Y,. Duc, G., Essad, S., Flament, P., Gallusci, P., Genier G., Guy P., Muel, X,. Tourneur, J., Denarid, J., and Huguet, T. (1990) *Medicago truncatula*, a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. *Plant Molecular Biology Report*, 8, 40-49.

Bate, N. and Twell, D. (1998). Functional architecture of a late pollen promoter: Pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. *Plant Molecular Biology*, 37, 859-869.

Berger, F., and Twell, D. (2011) Germline specification and function in plants. *Annual Review Plant Biology* 62, 461-84

Berger, S. (2007). The complex language of chromatin regulation during transcription. *Nature*, 447(7143), 407-412.

Bhatt, A., Canales, C., and Dickinson, H. (2001) Plant meiosis: The means to 1N. *Trends Plant Science. 6*, 114-121.

Borg, M. (2011) Regulatory mechanisms controlling male germline development in *Arabidopsis thaliana*. Doctor of Philosophy thesis, University of Leicester.

Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M., and Twell, D. (2011) The R2R3 MYB transcription factor DUO1 activates a male germline- specific regulon essential for sperm cell differentiation in Arabidopsis. *Plant Cell. 23*, 534-549.

Borg, M., Brownfield, L., and Twell, D. (2009) Male gametophyte development: A molecular perspective. *Journal Experimental Botany. 60*, 1465-1478.

Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijó, J., and Becker, J. (2008) Comparative transcriptomics of Arabidopsis sperm cells. *Plant Physiology. 148*, 1168-1181.

Brownfield, L., Hafidh, S., Borg, M., Sidorova, A., Mori, T., and Twell, D. (2009) A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genetics. 5*, e1000430.

Martin, C. and Paz-Ares, J. (1997) MYB transcription factors in plants. *Trends in Genetics*, 13(2), 67-73

Cheng, X., Wen, J., Tadege, M., Ratet, P., Mysore K.. (2011) Reverse genetics in *Medicago truncatula* using Tnt1 insertion mutants. Methods Molecular Biology. 678:179-90.

Cheo D., Titus S., Byrd D., Hartley J., Temple G., Brasch M. (2004) Concerted assembly and cloning of multiple DNA segments using in vitro site-specific recombination: functional analysis of multi-segment expression clones. *Genome Research* 14: 2111-2120. 10.1101/gr.2512204

Cowell, I. (1994). Repression versus activation in the control of gene transcription. *Trends in Biochemical Sciences*, 19(1), 38-42.

Cook, D. (1999) *Medicago truncatula* — A model in the making! *Current Opinion in Plant Biology*, 2(4), 301-304

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010) MYB transcription factors in Arabidopsis. *Trends Plant Science 15*, 573-581.

Deluca, M. and McElroy, W. (1978). Purification and properties of firefly luciferase. *Methods in Enzymology*, 57, 3-15.

Durbarry, A. (2005) Male germ line development in *Arabidopsis*. *duo* pollen mutants reveal gametophytic regulators of generative cell cycle progression. *Plant Physiology.* 137, 297-307.

Eady, C., Lindsey, K., and Twell, D. (1995) The significance of microspore division and division symmetry for vegetative cell-specific transcription and generative cell differentiation. *Plant Cell*. 7, 65-74.

Engel, M., Chaboud, A., Dumas, C., and Mccormick, S. (2003) Sperm cells of Zea mays have a complex complement of mRNAs. *Plant Journal.* 34, 697-707.

Engel M., Holmes-Davis R., McCormick S. (2005) Green sperm. Identification of male gamete promoters in Arabidopsis. *Plant Physiology*. 138, 2124-33.

Guo, J., Ling, H., Ma, J., Chen, Y., Su, Y., Lin, Q., Gao, S., Wang, H., Que, Y., Xu, L. (2017) A sugarcane R2R3-MYB transcription factor gene is alternatively spliced during drought stress. *Nature Scientific Reports* **7**, Article number: 41922

Goodstein D, Shu S, Howson R, Neupane R, Hayes R, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al (2012) Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research* 40: D1178–D1186

Haerizadeh, F., Singh, M., and Bhalla, P. (2006) Transcriptional repression distinguishes somatic from germ cell lineages in a plant. *Science.* 313, 496-499.

Hartley, J., Temple, G., and Brasch, M. (2000) DNA cloning using in vitro site-specific recombination. *Genome Research.* 10: 1788-1795.

Hellens, R., Allan, A., Friel, E., Bolitho, K., Grafton, K., Templeton M., Karunairetnam, S., Gleave, A., Laing, W. (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods*. Dec 18;1:13.

Hohn, T., Vazquez, F. (2011) RNA silencing pathways of plants: Silencing and its suppression by plant DNA viruses. *Gene Regulatory mechanisms.* Dec. 588-600.

Horvitz, H., and Herskowitz, I. (1992) Mechanisms of asymmetric cell division: Two bs or not two bs, that is the question. *Cell.* 68, 237-255.

Iwakawa, H., Shinmyo, A., and Sekine, M. (2006) Arabidopsis CDKA:1 , a cdc2 homologue, controls proliferation of generative cells in male gametogenesis. *Plant Journal.* 45, 819-831.

Jia, L., Clegg, M. T., and Jiang, T. (2004) Evolutionary dynamics of the DNA- binding domains in putative R2R3- MYB genes identified from rice subspecies indica and japonica genomes. *Plant Physiology.* 134, 575.

Jiang, C., Gu, J., Gu, X., and Peterson, T. (2004) Ordered origin of the typical two- and three-repeat Myb genes. *Gene* 4; 326: 13-22.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28* (12), 1647-1649.

Kim, H. J., Oh, S. A., Brownfield, L., Hong, S. H., Ryu, H., Hwang', I., Twell, D., and Nam, H. G. (2008) Control of plant germline proliferation by SCF.sup.FBL17 degradation of cell cycle inhibitors. *Nature.* 455, 1134.

Koncz, C., Langridge, W., Olsson, O. et al. (1990). Bacterial and firefly luciferase genes in transgenic plants: Advantages and disadvantages of a reporter gene. *Developmental Genetics*, 11(3), 224–232.

Landy, A., (1989) Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annual Revue Biochemistry*. 58: 913-949.

Lakatos, L., Szittya, G., Silhavy, D. and Burgyán, J. (2004) Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombus viruses. *The EMBO Journal*. 23(4): 876–884.

Ma, H., De Phamphilis, C. (2000) The ABCs of Floral Evolution. *Cell* Vol.101 (1) 5-8

MacAlister C., Ohashi-Ito K., Bergmann D., (2007) Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature.* 445(7127):537-40.

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

Martinez-Trujillo, M., Limones-Briones, V., Cabrera-Ponce, J., Herrera-Estrella, L. (2004) Improving transformation efficiency of *Arabidopsis thaliana* by modifying the floral dip method. *Plant Molecular Biology* Report 22: 63-70

Matthews, J., Hori, K. and Cormier, M., (1977). Purification and properties of Renilla reniformis luciferase. *Biochemistry*, 16(1), 85-91.

Mori T., Kuroiwa H., Higashiyama T., Kuroiwa T., (2006) Generative cell specific 1 is essential for angiosperm fertilization. *Nature Cell Biology* 8, 64-71

Mori T., Igawa T., Tamiya G., Miyagishima SY., Berger F. (2014) Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. *Current Biology.* 20;24(2):170-5.

Ogata K., Hojo H., Aimoto S., Nakai T., Nakamura H., Sarai A., Ishii S., Nishimura Y. (1992) Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proceedings of the National Academy of Science.* 89 pp. 6428-6432

Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994) Solution structure of a specific DNA complex of

the myb DNA- binding domain with cooperative recognition helices. *Cell.* 79, 639-648.

Okada T., Endo M., Singh M., Bhalla P., (2005) Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. *The Plant Journal* 44, 557-568.

Okada, T., Bhalla, P. L., and Singh, M. B. (2006) Expressed sequence tag analysis of *Lilium longiflorum* generative cells. *Plant and Cell Physiology.* 47, 698-705.

Owen H., Makaroff C., (1995) Ultrastructure of microsporogenesis and microgametogenesis in *Arabidopsis thaliana* (L.) Heynh. Ecotype Wassilewskija (Brassicaceae). *Protoplasma* 185, 7-21.

Pabo, C. and Sauer, R. (1992) Transcription Factors: Structural Families and Principles of DNA Recognition. *Annual Review of Biochemistry*. 61, 1053-1095.

Pacini E. (1996) Types and meaning of pollen carbohydrate reserves. *Sexual Plant Reproduction.* 9, 362-366.

Palatnik, J. F., Wollmann, H., Schommer, C. et al. (2007). Sequence and expression differences underlie functional specialization of Arabidopsis microRNAs miR159 and miR319. *Developmental Cell*, 13(1), 115-125.

Park S., Howden R. and David T. (1998) The *Arabidopsis thaliana* gametophytic mutation *gemini pollen1* disrupts microspore polarity, division asymmetry and pollen cell fate. *Development*. 125, 3789-99.

Peters, B., Casey, J., Aidley, J., Zohrab, S., Borg, M., Twell, D. and Brownfield, L. (2017) A Conserved cis-Regulatory Module Determines Germline Fate through Activation of the Transcription Factor DUO1 Promoter. *Plant Physiology.* 173. 1 280-293

Petricka, J., Van Norman, J., and Benfey, N. (2009) Symmetry breaking in plants: Molecular mechanisms regulating asymmetric cell divisions in Arabidopsis. *Cold Spring Harbor perspectives in biology.* 1, a000497.

Ratcliff, F., MacFarlane, S. and Baulcombe, D.C. (1999) Gene silencing without DNA: RNA-mediated cross protection between viruses. *Plant Cell*. 11, 1207–1215.

Rotman, N., Durbarry, A., Wardle, A., Yang, W. C., Chaboud, A., Faure, J., Berger, F., and Twell, D. (2005) A novel class of MYB factors controls sperm- cell formation in plants. *Current Biology*. 15, 244-248.

Rutley, N., Twell, D. (2015) A decade of pollen transcriptomics. *Plant Reproduction*. 28, 73-89.

Russell, S. (1992) Double fertilization. *Int.l Rev. Cyto.* 140, 357-388.

Russell S., Gou X., Wong C., Wang X., Yuan T., Wei X., Bhalla P., Singh M. (2012). Genomic profiling of rice sperm cell transcripts reveals conserved and distinct elements in the flowering plant male germ lineage. New Phytol; 195(3):560-73.

Sargant, E. (1900). Recent work on the results of fertilization in angiosperms, *Annals of Botany* (14): 689-712.

Scott, R. J. (2004) Stamen structure and function. *The Plant Cell Online. 16,* S46-S60.

Saikumar, P., Murali, R. and Reddy, E. P. (1990). Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. *Proceedings of the National Academy of Sciences*, 87 (21), 8452-8456.

Sherf, B., Navarro, S., Hannah, R. (1996). Dual-luciferase reporter assay: an advanced co-reporter technology integrating firefly and Renilla luciferase assays. *Promega Notes*, 57(2).

Singh, M., Bhalla, P. L., Xu, H., and Singh, M. B. (2003) Isolation and characterization of a flowering plant male gametic cell- specific promoter 1 1 GenBank accession number: AY207012. *FEBS Lett.* 542, 47-52.

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biology*. 4, 447-456.

Tanaka, I. (1997) Differentiation of generative and vegetative cells in angiosperm pollen. *Sexual Plant Reproduction*. 10, 1-7.

Tanikawa, J., Yasukawa, T., Enari, M., Ogata, K., Nishimura, Y., Ishii, S., and Sarai, A. (1993). Recognition of specific DNA sequences by the c-myb proto-oncogene product: role of three repeat units in the DNA-binding domain. *Proc. Natl. Acad. Sci*. 90, 9320-9324.

Twell, D. (2011) Male gametogenesis and germline specification in flowering plants. *Sex Plant Reproduction*. 24, 149-160.

Twell, D., Park, S. K., and Lalanne, E. (1998) Asymmetric division and cell-fate determination in developing pollen. *Trends in Plant Science.* 3, 305-310.

Ueda, K., and Tanaka, I. (1995) The appearance of male gamete-specific histones gH2B and gH3 during pollen development in *Lilium longiflorum*. *Developmental Biology.*169, 210-217.

Vance, V., Vaucheret, H. (2001) RNA silencing in plants—defence and counter defence. *Science* 292: 2277-2280

Vaucheret, H., Béclin, C., Fagard, M. (2001) Post-transcriptional gene silencing in plants. *Journal of Cell Sci.* 114, 3083-3091

Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM. (2009) Expansion and diversification of the *Populus* R2R3–MYB family of transcription factors. *Plant Physiology.* 149, 981-993.

Wilkinson J, Twell D and Lindsey K. (1997). Activities of CaMV35S and nos promoters in pollen: Implications for field release of transgenic plants. *Journal of Experimental Botany*. 48, 265-275

Xu, H., Swoboda, I., Bhalla, P. L., and Singh, M. B. (1999) Male gametic cell- specific gene expression in flowering plants. *Proc.Natl.Acad.Sci.* 96, 2554-2558.

Yamamoto, Y., Nishimura, M., Hara-Nishimura, I., and Noguchi, T. (2003) Behavior of vacuoles during microspore and pollen development in *Arabidopsis thaliana*. *Plant and Cell Physiology.* 44, 1192-1201.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G., and Li-Jia, Q. (2006) The MYB transcription factor superfamily of Arabidopsis: Expression analysis and phylogenetic comparison with the rice MYB family. *Plant.Mol.Biol.* 60, 107-12.

Žárský, V., Garrido D., Říhová L., Tupý J., Vicente O., Heberle-Bors E. (1992). Derepression of the cell cycle by starvation is involved in the induction of tobacco pollen embryogenesis. *Sexual Plant Reproduction.* 5, 189-194.

Zohrab, S.W.S (2015) Functional analysis of candidate *cis*-regulatory elements in the *DUO1* promoter of *Arabidopsis thaliana.* Honours thesis, Plant Biotechnology, University of Otago.

Zhang J, Yuan T, Duan X, Wei X, Shi T, Li J, Russell SD and Gou X (2016). *Cis*-regulatory elements determine germ-line specificity and expression level of an isopentenyl transferase gene in sperm cells of Arabidopsis. Plant Physiology 170, 1524-1534.

Zheng B., He H., Zheng Y., Wu W., and McCormick S. (2014) An ARID Domain-Containing Protein within Nuclear Bodies Is Required for Sperm Cell Formation in *Arabidopsis thaliana*. *PLoS Genetics*. 10(7): e1004421.