Genetic analysis of the DAZ1 transcription factor in *Arabidopsis thaliana*

Thesis submitted for the degree of

Doctor of Philosophy

by
Jaypal Darbar

B.Sc, M.Sc. Botany (INDIA)

M.Sc. Biotechnology (UK)

March 2018

Department of Genetics and Genome Biology

Genetic analysis of the DAZ1 transcription factor in *Arabidopsis thaliana*

by
Jaypalsinhji Darbar

Abstract

The aim of this research was to advance understanding of the molecular and genetic mechanisms by which the transcription factor, DAZ1 (DUO1-ACTIVATED ZINC FINGER 1) supports male germline development in *Arabidopsis thaliana*. The main approach was to analyse the importance of different protein domains in DAZ1 for generative cell division and pollen fertility. The DAZ1 basic region (BR), the DAZ1 conserved region (CR) and the role of individual zinc finger (ZnF) motifs were investigated. *In planta* complementation of mutant *Arabidopsis daz1*$^{-/-}$ *daz2*$^{+/-}$ plants with deletion and substitution variants of DAZ1 revealed important functional roles for the BR, CR and for ZnF3 in male germline development. DAZ1 variants with BR or CR mutations showed a significant but limited ability to rescue male germ cell division and transmission of *daz1 daz2* pollen. A second objective was to study the cause of unexpected seed abortion observed in *daz1*$^{-/-}$ *daz2*$^{-/-}$ mutants that express DAZ1 variants with reduced function. Observations revealed ovule enlargement without development of embryo or endosperm tissues indicating early post-fertilisation defects arising from defective gametes. In a third objective, a genetic screen was conducted in Arabidopsis for suppressors of reduced fertility in chemically mutagenised plants expressing a C-terminally truncated DAZ1. Several suppressor lines were identified which were stably maintained through multiple generations. Overall, these studies uncover the importance of protein domains which support the role of DAZ1 in male germline development. This work also generated putative genetic suppressors which could provide access to novel regulators that orchestrate germ cell division and sperm differentiation in *Arabidopsis*.

Acknowledgements

It is a pleasure to extend my deepest thanks to many people who are part of this journey directly and indirectly, because of them, I achieved my goals in research.

First and foremost, I would like to express sincere gratitude bottom of my heart to my supervisor Professor David Twell for his endless excellent guidance, patience and commitment throughout my PhD. I learned many positive things, which will help in future to build my career from our legendary long discussions, recognised as a small chat in the office. Also, would like to thanks to him to monitoring my experiments and provide solutions, which saved a lot of time in the lab. Specially, I would like to thank him for his support during my financial unrest. I have been involved in great science and I am proud to have been part of it. Special thanks to Prof Pat-Haslop-Harrison and Dr Richard Gornall for feedback about research and for examining me over the years.

I would like to thank Dr Michael Borg (Banya – means brother), Dr Nichols Rutley and Sanjeev Kumar. Specially, I would like to thank Mike to for guidance, for explanations and for providing advice. You were always willing to help and give me suggestions to improve my scientific background. In addition, of course, sharing knowledge about the great Indian curry recipes and Indian places.

Additionally, I would like to acknowledge Twell lab members over the years, not just for the discussion, advice and assistance, but also for creating a supportive and friendly atmosphere in the lab. I would like to special thanks to Dr Nadia Taimur, Dr Nicholas Rutley, Dr Ugur sari, Dr Süleyman Avci, Dr Chetan Patokar, Dr Yanfeng Kuang, Dr Mingmin Zhao, Safia Saada, Charles Solomon and Christina Fischer for showing friendship and invaluable experimental practice. Specially, I would like to thank Sanjeev Kumar and Ramesh Patel for helping in my research by teaching me different methods and for maintaining plants in growth cabinets and the greenhouse. I would like to special thanks to Manish Mansuria who acted as big brother and friend, he always tried to find solutions for my problems. So, really thank you very much Manishji bottom of my heart.

Specially, I would like to thank to Dr Dieter Hackenberg for providing me excellent grounding lab information and for his suggestions, trouble shooting, discussions and ideas during my research. You were always willing to help and give suggestions in the right direction. In addition, our random talks about gold, cars, lottery, especially my Bollywood music that I used to sing loudly, and he mimic in German.

I would like to Amal Alyamani to be my good friend who come forward when I struggled to sort out my financial difficulties. Also, would like to thanks to Amal for giving me new name ‘Joubal’, which quite exciting. Also thanks for making some random jokes and laughing loudly in lab when no one was around. And for sharing a lot information from Saudi Arabian and Indian cultures so, now I learned more about those places, which I will visit in future.

For, Yosra Al-Hakeem I have no word to describe her support throughout my PhD. She made endless efforts inside and outside lab to make this done and stand next to me in most difficult part of my journey. So, I would like to huge thanks to you for being special friend and for supporting me during my financial difficulties because, without you it was not possible to achieve that most important deadline. Thanks for giving me guidance and information for research that help me in my future. Also, I would like to thank to her daughter ‘SADAN’ I called her ‘angel’ and ‘Baby Lioness’ because, she makes me relax from busy work life by explaining about her school life, her friends, piano lessons, belle dance class and much more, whenever she saw me outside lab start playing with me. I really enjoyed her company by playing and laughing. Also, would like to thank Yosra for sharing Iraqi cuisines, cultures and other useful information about Islam that I did not know before, it was great.

Finally, words are not enough to say anything about my parents and my sisters who day and night support me during my journey. Those efforts and advice always encourage me. I would not be where I am today, without their support. I dedicate this thesis to my parents and sisters. Thank you very much mom, for your endless advice about eating healthy, loving, and caring and some hard advice to stay on the track. Thank you very much sisters for your unconditional love, advice and being my guardian during my PhD journey. You both make it possible, we argued a lot over the

phone, but you both stand next to me when all situations against me, without your hard work it was impossible to achieve this goal.

Many thanks to my friends and relatives from India, UK, and Germany who helped me during my PhD study. Specially, when I am struggling for finance, at that time you helped me without saying anything, and resolved problem. Therefore, I would like to say thank you to my brother-in-law Nilesh Patel (UK), sister Nikita Patel (UK), Nileshbhai Khatri (India), Rasminbhai Patel (India) and Mr. Ian Kiru Karan (Germany).

Dedication

‘This thesis is dedicated to my hardworking, supportive and most respectable parents Mrs. Veenaba N Darbar, Mr. Natavarsinh J Darbar, my sisters who inspired me throughout my study Ms. Asha N Darbar and Ms. Gopi N Darbar for their unconditional love, care, endless support and encouragement.’

Abbreviations

At	– *Arabidopsis thaliana*
AP3	– APETALA3
AG	– AGAMOUS
Br	– *Brassica rapa*
Bd	– *Brachypodium distachyon*
BC	– Back crosses
BIM1	– BES1 Interaction MYC-like protein 1
bHLH	– Basic-helix-loop-helix
bZIP	– Basic-leucine zipper
BP	– Base pair
BR	– Basic region
BCP	– Bicellular pollen
Col-0–wt	– Columbia-0 wild type
CaMV	– Cauliflower mosaic virus
CR	– Conserved region
Cyt	– Cytoplasmic
CDKA	– Cyclin-dependent kinase
CYCB1,1	– Cyclin B1,1
CRY 1	– Crypto-chrome 1
CRY 2	– Crypto-chrome 2
CRP810	– Cysteine – rich peptide
COP	– Constitutive photomorphogenic
CDS	– Coding sequence
DAPI	– 4’,6’– diamidino – 2 – phenylindole
DUO1	– DUO POLLEN 1
DAZ1	– DUO1 – ACTIVATED ZINC FINGER 1
DAZ2	– DUO1 – ACTIVATED ZINC FINGER 2
DAT	– DUO1 – activated target
DNA	– Deoxyribonucleic acid

dNTPs – Deoxyribonucleotide triphosphate

DIC – Differential interference contrast

DAP – Day after pollination

DB – Destruction box

DEFL – Defendin – like

DET – De-Etiolated

EAR – Associated amphiphilic repression

ECM – Extra cellular matrix

EDTA – Ethylenediaminetetraacetic acid

E. Coli – *Escherichia coli*

EMS – Ethyl methanesulfonate

ERF – Ethylene response element-binding factor

EtBr – Ethidium bromide

FLuc – Firefly luciferase

fmol – Femtomole

FAR1 – Far-red impaired response 1

FRI – FRIGIDA

FAY3 – Far-red elongated hypocotyl 3

FUS – Fusca

g – Gravity

gai – gibberellin-insensitive

GBF1 – G-box binding factor 1

GCS1 – GENERATIVE CELL SPECIFIC 1

GEX – GAMETE EXPRESSED PROTEIN

gem1 – Gemini pollen 1

gem2 – Gemini pollen 2

Gen – Gentamicin

GFP – Green fluorescence protein

GUS – Beta glucuronidase

GV3101 – *Agrobacterium tumefacians*

HAP – Hours after pollination

HAP2 – HAPLESS 2

HDACs – Histone deacetylases

HFR1 – Long hypocotyl in FAR-RED 1

HOX – Homeobox

HY5 – Elongated hypocotyl 5

Hyg – Hygromycin

K – Lysine

Kan – Kanamycin

KPL – Kokopelli

L – Leucine

LRE – Light responsive element

LAF1 – Long after FAR-RED light 1

LB media – Luria-bertani media

LUG – LEUING

mm – Millimeter

MSA1 – More sulphur accumulation 1

MET 1 – methyl transferase 1

MS – Murashige and Skoog

MYBs – MYB binding sites

MYO – Million years ago

N – Asparagine

ng – Nanogram

Nt – *Nicotiana tabacum*

NLS – Nuclear localisation signal

NPCs – Nuclear pore complexes

Nuc – Nuclear

NUPs – Nucleoporins

Os – *Oryza sativa*

PAR1 – Phytochrome rapidly regulated

PCR – Polymerase chain reaction

pH – Potential of Hydrogen

PHYA – Phytochrome A

PI – PISTILATA

PMI – Pollen mitosis I

PMII – Pollen mitosis II

POEM – Pollen tube-dependent ovule enlargement morphology

PPT – Phosphenothricin

PTC – Pollen tube contents

R – Arginine

RanGAP – Ran-GAPase-activating protein

RF – Rescued fertility

RFP – Red fluorescence protein

Rif – Rifampicin

RLuc – Renilla luciferase

ROI – Region of interest

RPM – Round per minutes

S – Serine

SAGE – Serial analysis of gene expression

SAM – S-adenosyl methionine

Sb – *Sorghum bicolor*

Sl – *Solanum lycopersicum*

St – *Solanum tuberosum*

Spec – Spectinomycin

SRF – Serum response factor

SUP – SUPERMAN

SV40 – Simian virus 40

TAIR – The Arabidopsis Information Resource

TCP – Tricellular pollen

TE – Tris acetate EDTA

TGA-1A – TGAGC-sequence specific binding proteins 1A

TGA-1B – TGAGC-sequence specific binding proteins 1B

TILLING – Targeting induced local in genomes

TPI – Total pixel intensity

TPL – TOPLESS

Tm – Melting temperature

UM – uninucleate microspore

UNM – Uninucleate microspore

UV – Ultra violet

WT – Wild type

WUS – WUSCHEL

ZFP – Zinc finger protein

Zm – *Zea mays*

Table of contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The plant life cycle and alternation of generations

Plants undergo a unique and complex life cycle that cycle alternates between a haploid gametophyte and a diploid sporophyte generation that differ functionally and morphologically (Maheswari 1950). The dominance of each generation varies across the plant kingdom. For example, in bryophytes (mosses, liverworts, and hornworts) the principal generation is the gametophyte which produces the gametes while in vascular plants the diploid generation is dominant. The immature sporophyte in angiosperms develops to produce the main plant structures, the roots, shoots, and flowers. Meiotic division leads to the production of haploid microspores and megaspores, both of which characterise the beginning of the haploid gametophyte phase, eventually leading to the production of differentiated male and female gametes. During land plant evolution the haploid male gametophyte generation that develops from single haploid spores has progressively reduced in complexity, shifting from multicellular free-living plants in mosses and ferns to endosporic development in aquatic ferns such as *Marsilea*, in which micro gametophytes develop entirely within the microspore wall (Berger and Twell 2011). In angiosperms the mature male gametophyte is the most-highly reduced and three-celled, while the female gametophyte is typically seven-celled. Following double fertilisation, involving fusion of one sperm with the egg and a second sperm diploid central cell nucleus, embryo and endosperm structures develop (Figure 1.1). Maternal tissues also contribute to the developing seed which upon germination and growth of a new sporophyte generation completes the plant life cycle.

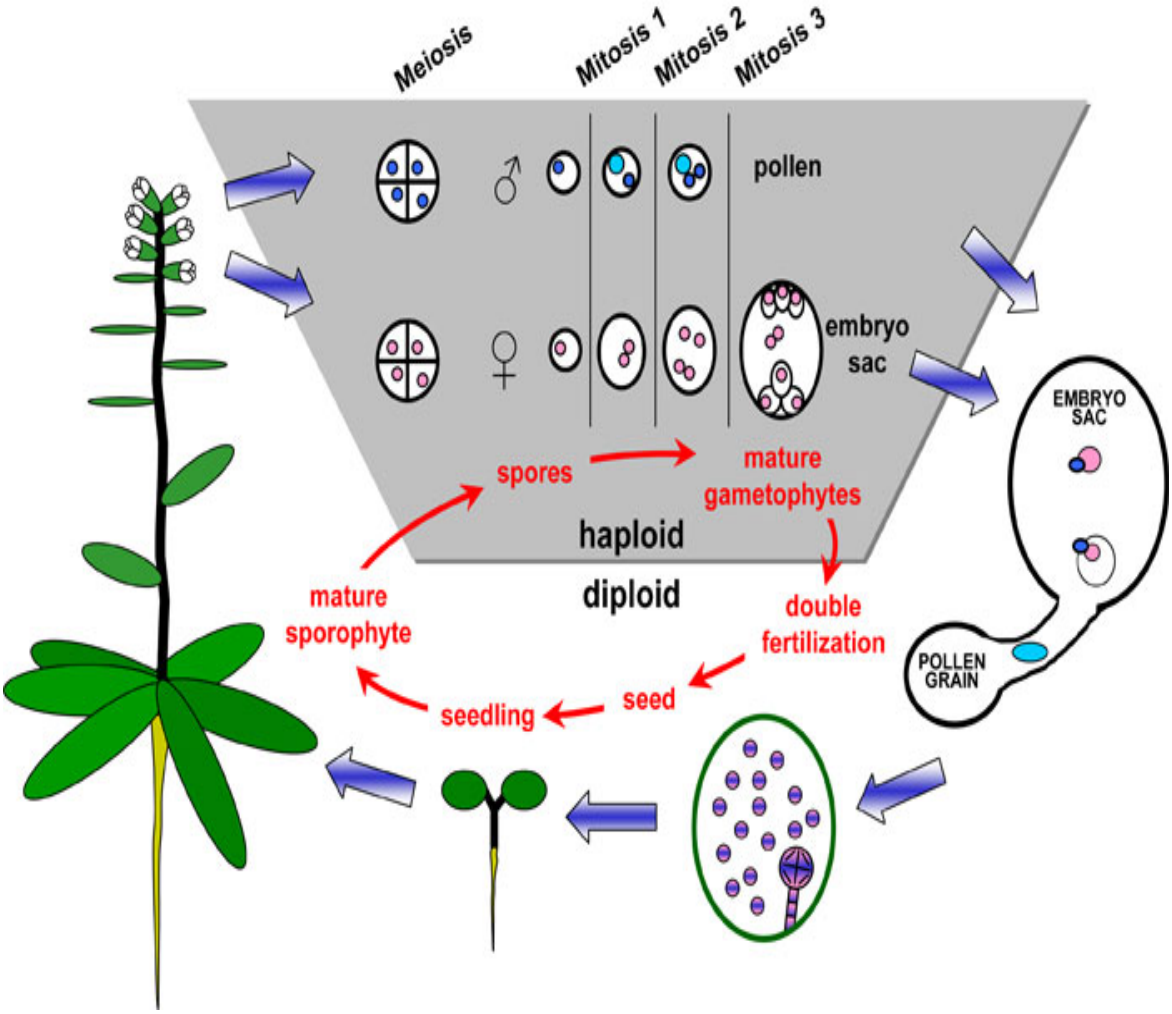


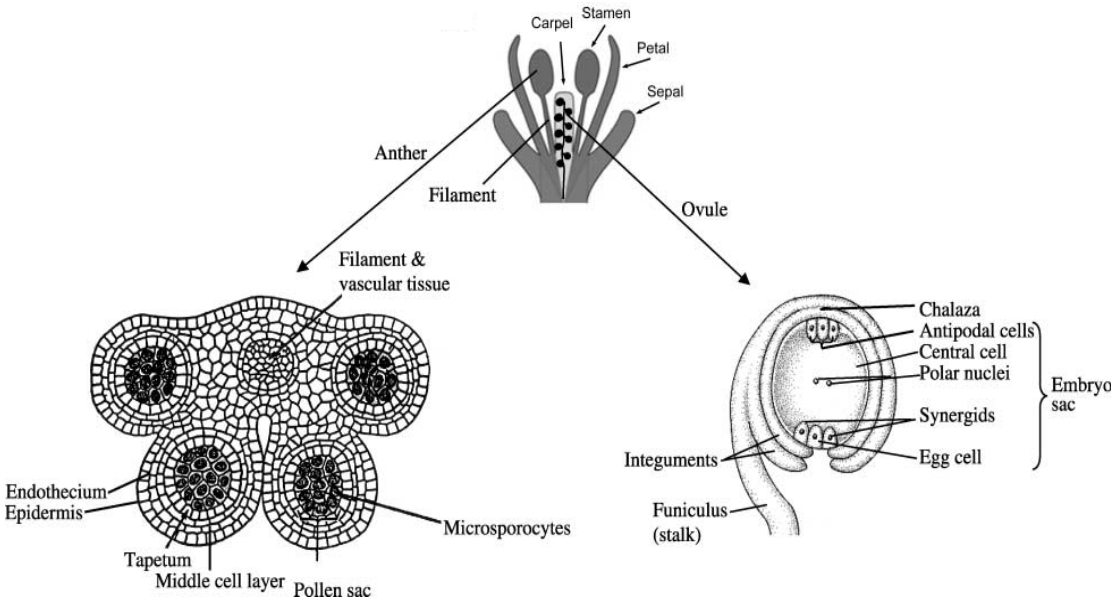
Figure 1.1. The life cycle of *Arabidopsis thaliana*. Diagram showing diploid sporophytic and haploid gametophytic phases (Berger *et al.* 2006).

1.2 Anther and pollen structure and function of *A. thaliana*

Male gametophyte development in higher plants is a complex process that requires the coordination of various cell and tissue types and their specifically associated gene expression patterns (McCormick 2004, Twell 2011). This process takes place inside the anther which is part of the stamen of the flower. The development of the stamen involves production of the anther which contain the microsporangia supported by the filament (Figure 1.2A). In *Arabidopsis thaliana* at floral stage 13, filaments elongate rapidly and anthers dehisce to release viable pollen onto the

stigma for pollination and subsequent fertilisation. Disruptions in stamen development frequently lead to male sterility (Sanders *et al.* 1999); (Ge *et al.* 2010).

(A)



(B)

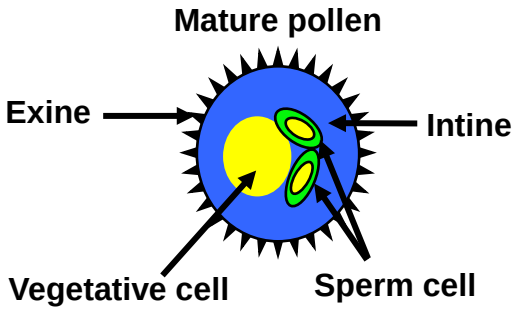


Figure – 1.2. Male and female reproductive organs in *A. thaliana*. (A) Arabidopsis floral, anther, and ovule structure (Wilson and Yang 2004). (B) Arabidopsis pollen structure.

Pollen grains are the micro-gametophytes of seed plants that produce the male gametes required for sexual reproduction. Angiosperm pollen contains either two (bicellular) or three (tricellular) cells when shed from the anther and in the progamic phase of tricellular pollen species such as *Arabidopsis thaliana* (Figure 1.2B), the germ cell produces two sperm cells during pollen tube growth (Durbarry *et al.* 2005). Pollen also protects the germline from environmental injury, promotes gamete

dispersal and ensures delivery of a pair of sperm cells to the female gametes via the pollen tube. Pollen is found with different sizes, shapes and surface characters in diameter with round, elliptical and multifaceted pollen type (Borg *et al.* 2009). The exine protects the reproductive cells from the impacts and environmental injury and plays an important role in attachment to insect pollinators and adhesion to stigmatic surface. The exine is composed of sporopollenin, whereas the intine is largely composed of pectin and cellulose. Exine synthesis and patterning are largely under control of the sporophytic or diploid cells of the anther. The exine is developed through the contribution of the inner anther wall cell layer, the tapetum, which has a dominant role in the synthesis and deposition of sporopollenin on the pollen surface and the early microspore cytoplasm (Figure 1.2B) (Blackmore *et al.* 2007).

1.3 Pollen development and gametogenesis

In flowering plants male gametogenesis is restricted to a simple cell lineage of two cell divisions following meiosis that results in the production of two non-motile sperm cells (Durbarry *et al.* 2005). Whereas animal cells destined to become germline cells are determined early in embryogenesis, in angiosperms germ cells are only developed late, after the transition to flowering in the gametophyte. A single mitotic division of the generative cell and cell specification produce functional gametes capable of double fertilisation (Berger and Twell 2011).

The simple cell lineage and highly orchestrated development of the male gametophyte characterises is crucial for sexual reproduction and represents microcosm of cellular development (Borg *et al.* 2009). The first division of the microspore at pollen mitosis I (PMI) is asymmetric and gives rise to a large transcriptionally active vegetative cell and a smaller generative cell with condensed chromatin and fewer organelles. Moreover, formation of the generative cell marks the establishment of a discrete male germ cell lineage (Figure 1.3) (Tanaka 1988, Twell *et al.* 1998). After, PMI, the two daughter cells follow different developmental pathways that are characterized by the differential control of the cell cycle and gene

expression. Whereas, the vegetative cell exits the cell cycle in G1 phase and will produce the pollen tube, the generative cell completes a further cell cycle to form the two sperm cells (Brownfield and Twell 2009).

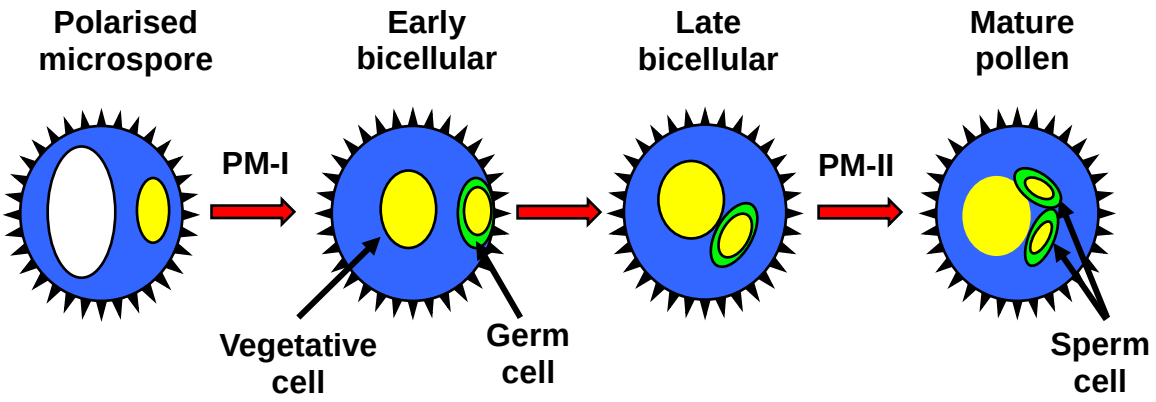


Figure 1.3. Stages of male gametophyte development in *A. thaliana*. Development is illustrated starting with haploid microspores, the production of vegetative and generative cells after asymmetric division and division at pollen mitosis II to produce twin sperm cells (Borg *et al.* 2009).

During microgametogenesis, the released microspores enlarge, and a single large vacuole is produced (Owen and Makaroff 1995, Yamamoto *et al.* 2003). These microspores then undergo an asymmetric cell division known as pollen mitosis I (PMI) (Figure 1.3). This asymmetric division is essential for the correct cellular patterning of the male gametophyte, since the two daughter cells each harbour a distinct cytoplasm and possess unique gene expression profiles that confer their distinct structures and cell fates (Twell *et al.* 1998). Finally, during pollen maturation a physical association between the sperm cells and the vegetative nucleus is established that is referred to as the male germ unit (MGU) (Lalanne and Twell 2002). This MGU is common in species which shed either bicellular or tricellular pollen and is important for the coordinated delivery and fusion of the sperm cells with the female gametes in the embryo sac (Dumas *et al.* 1998).

1.4 *Arabidopsis thaliana* as a genetic model

Arabidopsis thaliana (thale cress, mouse-ear cress or *Arabidopsis*) is small flowering plant, native to Eurasia and plant belongs to the Brassicaceae family. This family contains over 330 genera and approximately 3,700 species. Brassicaceae contain important crop plant species such as *Brassica napus* (rapeseed). *Arabidopsis thaliana* is popular model plant organism in plant biology, molecular biology and in genetics due to its small size, rapid life cycle and comparatively small genome of approximately 125 Mb. Arabidopsis an important tool for understanding molecular and genetic aspects of other flowering plants including crop species (Ref. www.arabidopsis.org). The Arabidopsis Information Resource (TAIR) maintains Arabidopsis genome data. Plant transformation techniques are facile and well developed in Arabidopsis and involve using *Agrobacterium tumefaciens*-mediated transfer of DNA via binary T-DNA vectors. This allows transgenic approaches to be used to investigate understand gene expression and functions in all aspects of development (Ref. www.arabidopsis.org).

1.5 Conservation and functional role of genes in plants

Plant nuclear genomes are extremely flexible like chromosome numbers, the degree of gene clustering and chromosome size can all differ by as much as an order of magnitude, even between closely related species (Kellogg and Bennetzen 2004). Some variation is generated so rapidly that two different allelic versions of a chromosomal segment can be dissimilar in gene content and arrangement even within a single plant species like maize (Fu and Dooner 2002, Kellogg and Bennetzen 2004).

Recent results suggest that evolution frequently occurs on contemporary timescales, repeatedly within observable time periods (Stockwell *et al.* 2003). Thus, contemporary evolution is associated with the same aspects that are driving the present extinction crisis, habitat loss and degradation, overharvesting and exotic

species loss (Stockwell *et al.* 2003). Stuart et al (2003) demonstrated and identified pairs of genes that are co-expressed over 3182 DNA microarrays from humans, flies, worms and yeast and found that 22,163 such co-expression relationships, each of which has been conserved across evolution. Therefore, this conservation implies that co-expression of these gene pairs confers a selective advantage and therefore, these genes are functionally related (Stuart *et al.* 2003). The current accessibilities of large DNA microarray datasets for humans, files, worms, and yeast makes it possible to measure evolutionary conserved co-expression on a genome wide-scale (Teichmann and Babu 2002, Alter *et al.* 2003, van Noort *et al.* 2003).

Evolutionary developmental genetics is a novel scientific concept, which assumes that changes in developmental control genes are a foremost characteristic of evolutionary changes in morphology, which provide understanding the phylogeny of developmental regulator genes may thus help to understand evolution of plants and mammals (Theissen *et al.* 2000). Flowers and phylogenetic ancestors are an ideal model system to understand the linkage between development, genes and evolution (Doyle 1994). Most animals are probably relatively closely related members of a monophyletic group because their diverse body shapes evolved in a relatively short period of time about 540 million years ago (MYA) that process has been termed the 'Cambrian explosion' (Philippe *et al.* 1994). For example, mammals specify their body plans in a similar way, involving homeobox genes (HOX genes) which are organised in clusters (Slack *et al.* 1993, Rounsley *et al.* 1995, Raff 2012). However, the absence of the HOX clusters in plants explains that the existence of these genes is not an absolute necessity for the evolution of composite multicellular body plans, for example, the oldest acknowledged seed plant (*Elkinsia*) has been conserved in the fossil record of the time break (late Devonian, about 365 MYA), and different transitional stage in the evolution of the ovule have been found in the fossil record of the lower carboniferous, about 350 MYA (Theissen *et al.* 2000).

Homeotic transformation is observed in floral patterning mutants which arise from 8mutations in meristem and organ identity genes. The phenotypes of three classes

of mutants were used to construct the well-established ABC model of floral patterning (Bowman *et al.* 1991). In Arabidopsis three homeotic functions A, B, and C were proposed. A function contributed by both APETALA1 (AP1), and APETALA2 (AP2), B function by APETALA3 (AP3) and PISTILATA (PI), and C function by AGAMOUS (AG) (Theissen *et al.* 2000). Except AP2 the ABC genes all share a highly conserved MADS-box, which encodes the DNA-binding domain of these MADS-domain transcription factors (Schwarz-Sommer *et al.* 1990, Sommer *et al.* 1990, Coen and Meyerowitz 1991, Weigel and Meyerowitz 1994, Theißen *et al.* 1996), , (Yanofsky *et al.* 1990). MADS domain is an acronym for the four founder proteins MCM1 (from yeast), AGAMOUS (from *Arabidopsis*), DEFICIENS (from *Antirrhinum*), and SRF (from human). These MADS-domain transcription factors control a gene network contributing to flower development and MADS-box genes are not only dominant among the organ identity genes, but also well represented at other levels, i.e. among meristem identity genes, intermediate genes, cadastral genes, and possibly even downstream genes (Theissen *et al.* 2000). In contrast to the HOX genes of mammals, which are organised in genomic clusters, the MADS box genes of plants are scattered throughout the genome (Fischer *et al.* 1995), (Liljegren *et al.* 1998).

1.6 DUO POLLEN 1 - a male germline-specific transcription factor

Male germline development in flowering plants is initiated with the development of the generative cell that later produces two twin sperm cells which are required for double fertilisation. DUO1 is the first male germline-specific transcription factor to be identified in plants and has an essential role in cell division and cell specification (Durbarry *et al.* 2005), (Rotman *et al.* 2005), (Brownfield *et al.* 2009). It belongs to R2R3-MYB gene family which is one of the largest transcription factor families in Arabidopsis. DUO1 protein accumulates in generative and sperm cell nuclei (Rotman *et al.* 2005), (Brownfield *et al.* 2009).

The mutant *duo1-1* was identified by ethyl methane sulfonate (EMS) mutational analysis in *A. thaliana* and results showed male germ cells unable to divide instead

of dividing to form two sperm cells (Durbarry *et al.* 2005, Rotman *et al.* 2005). Further, studies showed that *duo1* alleles block CYCB1;1 expression during G2-phase in the generative cell and prevented cell differentiation and the expression germ-cell-specific marker genes (Brownfield *et al.* 2009). DUO1 was shown to regulate HTR10, a male germline-specific histone H3.3 (HTR10), GEX2, a membrane-associated protein and GCS1/HAP2, an ancestral membrane-associated protein required for the gamete fusion (Brownfield *et al.* 2009). The DUO1 MYB domain was shown to bind to conserved sequences in the *HTR10* promoter and in other candidate target genes such as GCS1/HAP2 and GEX2, showing that these represent direct DUO1 targets (Borg et al, 2011). DUO1 also regulates the expression of other putative transcription factors, including DAZ1, DAZ2 (DUO1-ACTIVATED ZINC FINGER 1 / 2) proteins which belong to the C_2H_2-type zinc finger family of transcription factors (Englbrecht *et al.* 2004).

1.7 Overview of conservation and expansion of C_2H_2 zinc finger proteins

The *DAZ1* and *DAZ2* genes are thought to be regulated by DUO1 binding to their promoter regions (Borg *et al.* 2014). C_2H_2 zinc finger proteins constitute an abundant family of DNA binding proteins in the genomes of higher and lower eukaryotes. Investigations showed that 176 zinc finger proteins are present in *Arabidopsis thaliana*, hence this family constitutes the most abundant family of putative transcriptional regulators. Only 33 *A. thaliana* zinc finger proteins are conserved in other eukaryotes and most C_2H_2 zinc finger proteins in Arabidopsis are plant-specific (Figure 1.4). The classical C_2H_2 zinc finger domain participates in a wide range of functions and can bind to DNA, RNA and proteins (Englbrecht *et al.* 2004).

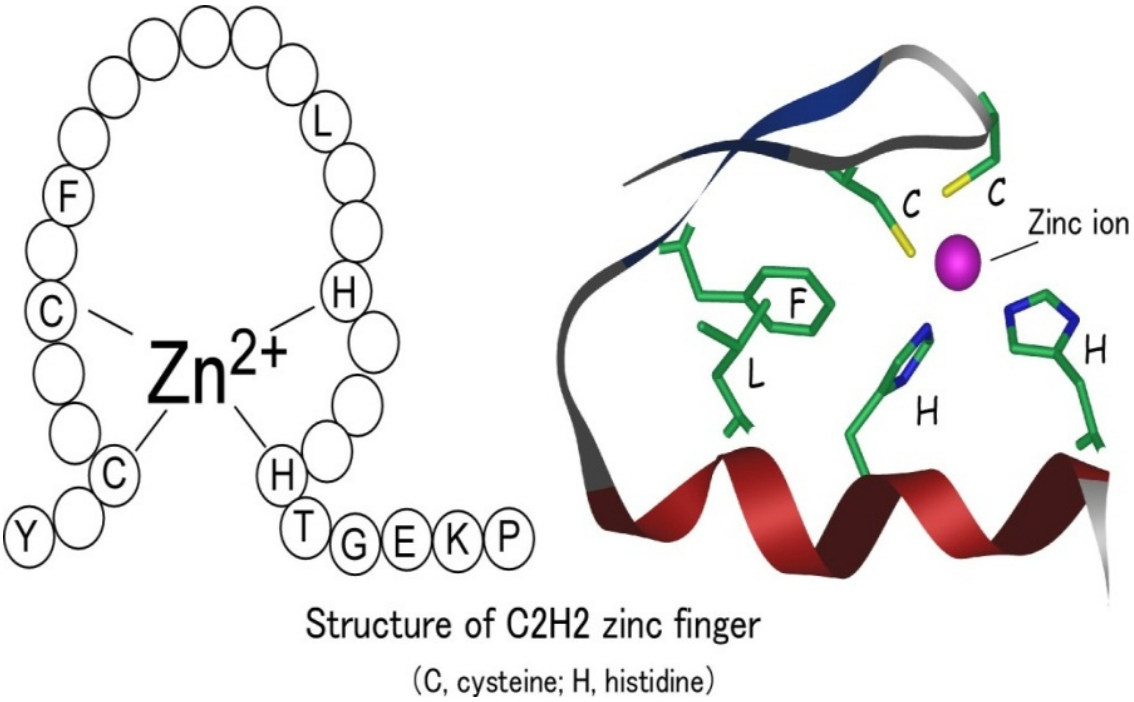


Figure 1.4. C_2H_2 Zinc finger protein structure. Structure of C_2H_2-type zinc finger, two cysteine molecules and two histidine molecules coordinate the zinc ion (Purugganan 1997).

A. thaliana C_2H_2-type zinc fingers are classified based on the arrangement and spacing of zinc fingers and the numbers of amino acids coordinating histidine residues. In DAZ1 and DAZ2 zinc fingers are dispersed and classified as class C, but are further divided into subclasses C1, C2, and C3 (Englbrecht *et al.* 2004). These subclasses are characterized by ZnF types that differ in spacing between the two-invariant zinc coordination histidine residues by three C1, four C2 and five C3 amino acid residues. Subset C1 comprises 77 ZFPs containing ZFs with HX3H or HX3C spacing in subset C2. Subset C3 contains 16 out of 22 ZFPs which are conserved in eukaryotes other than plants. Subclass C1-1i has 33 members found specifically in plants (Englbrecht *et al.* 2004). The 6-amino acid sequence of the α helixes of the zinc fingers proximal to the pair of histidines place DAZ1 and DAZ2 in subgroup C1 – 3iC, having the sequence KALGGH, for the first ZnF QALGGH for the second and third ZnFs respectively (Rutley, 2014), (Englbrecht *et al.* 2004).

Furthermore, around 55 % of C_2H_2-type ZnF protein genes were shown to be expressed during male gametophyte development (Honys and Twell 2004).

1.7.1 Overview of the Basic (BR) and Conserved (CR) region domains in DAZ1

DAZ1 and DAZ2 both possess a short region of basic amino acids (PKKRKV) basic region (BR) in the N-terminal region of each protein. The BR resembles a monopartite nuclear localisation signals (NLS) that are known to mediate the transport of proteins into the nucleus and first discovered in SV40 large T-antigen (Dingwall and Laskey 1991). The monopartite NLS-mediated nuclear import pathway exists in yeast and other eukaryotes and is referred to as a 'classical' NLS pathway (Ba *et al.* 2009). These classical NLSs show characteristic patterns of basic residues loosely matching two consensus sequences, K(K/R) X (K/R) and $KRK_{10\text{-}12}$ KRXX, termed 'monopartite' and 'bipartite' classical NLS (Poon and Jans 2005, Lange *et al.* 2007). Recognition of these classical NLSs on nuclear proteins is through the importin α subunit, which is recognised by the importin β subunit. Some NLS families are independent of importin α and may bind directly to one of the members of the importin β superfamily (Lange *et al.* 2007).

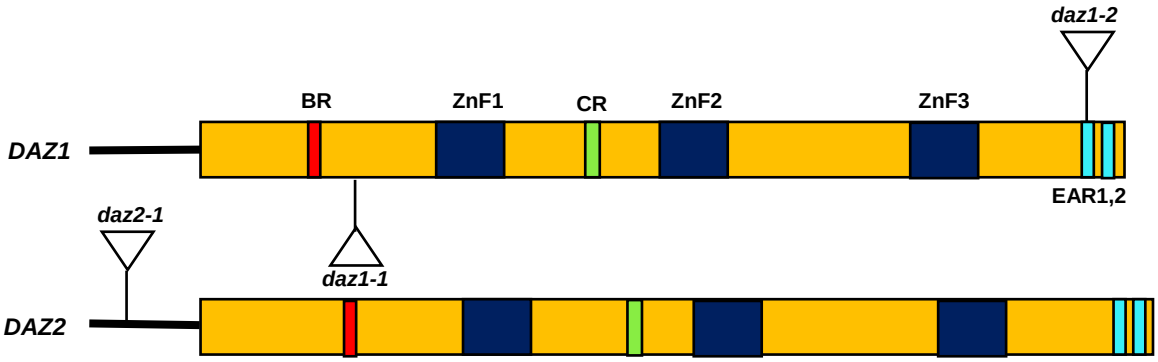


Figure 1.5. DAZ1 and DAZ1 protein domain organisation. DAZ1 and DAZ2 are members of subgroup C1-3iC within the C_2H_2-type family that are characterised by three dispersed zinc finger protein domains and a conserved region containing a CLLM motif between the first and second zinc fingers (Borg *et al.* 2014).

In an initial phylogenetic study of DAZ1, an N-terminal basic region (BR) was found to be present in *DAZ1* orthologues from eudicots (Figure 1.5), but absent from *DAZ1* orthologues from several monocots (Borg *et al.* 2014). Alignment of DAZ1 sequences from further families of eudicots and monocots confirmed that the BR is conserved in eudicot species (eg. *Solanum lycopersicum*), but is absent from monocot species (eg. *Sorghum bicolor, Oryza sativa, Zea mays)* (Rutley, 2015). Interestingly, the BR was found to be absent from a DAZ1 orthologue from *Amborella trichopoda,* the earliest divergent extant Angiosperm(Albert *et al.* 2013), indicating that this region evolved after the divergence of the monocot clade. The importance and role of the basic region in dicot DAZ1 proteins remained unknown and this is addressed in Chapter 3.

The conserved region CLLM is located between ZnF1 and ZnF2 domains in both DAZ1 and DAZ2. CLLM is conserved in DAZ1 homologues throughout the angiosperms in both monocot and eudicot species (Figure 1.5). However, *A. trichopoda* DAZ1 shows sequence differences in the conserved region compared with homologues from monocot and eudicot species. The importance of the conserved region in Arabidopsis DAZ1 proteins is addressed in Chapter 3.

1.7.2 The functional role of DAZ1/DAZ2 in the male germline

As described earlier DAZ1 and DAZ2 form a distinct subgroup of C_2H_2-type Zinc finger proteins that are categorised by three zinc finger domains (Englbrecht *et al.* 2004, Borg *et al.* 2014). Recent data shows that DUO1 is able to activate the transcription of its direct target genes by binding to specific sequences in their promoter regions termed MYB binding sites (MBSs). Mutagenesis of these MBSs in revealed their importance for *trans*-activation of the DAZ1 and DAZ2 promoters by DUO1 (Borg *et al.* 2014). These data provide strong support for a direct role for DUO1 in DAZ1 and DAZ2 transcription. Furthermore, in planta analysis of protein accumulation revealed the phased expression of DAZ1 and DAZ2, which peak

following that of DUO1. Therefore, DUO1 directly determines the male germline-specific accumulation of DAZ1 and DAZ2.

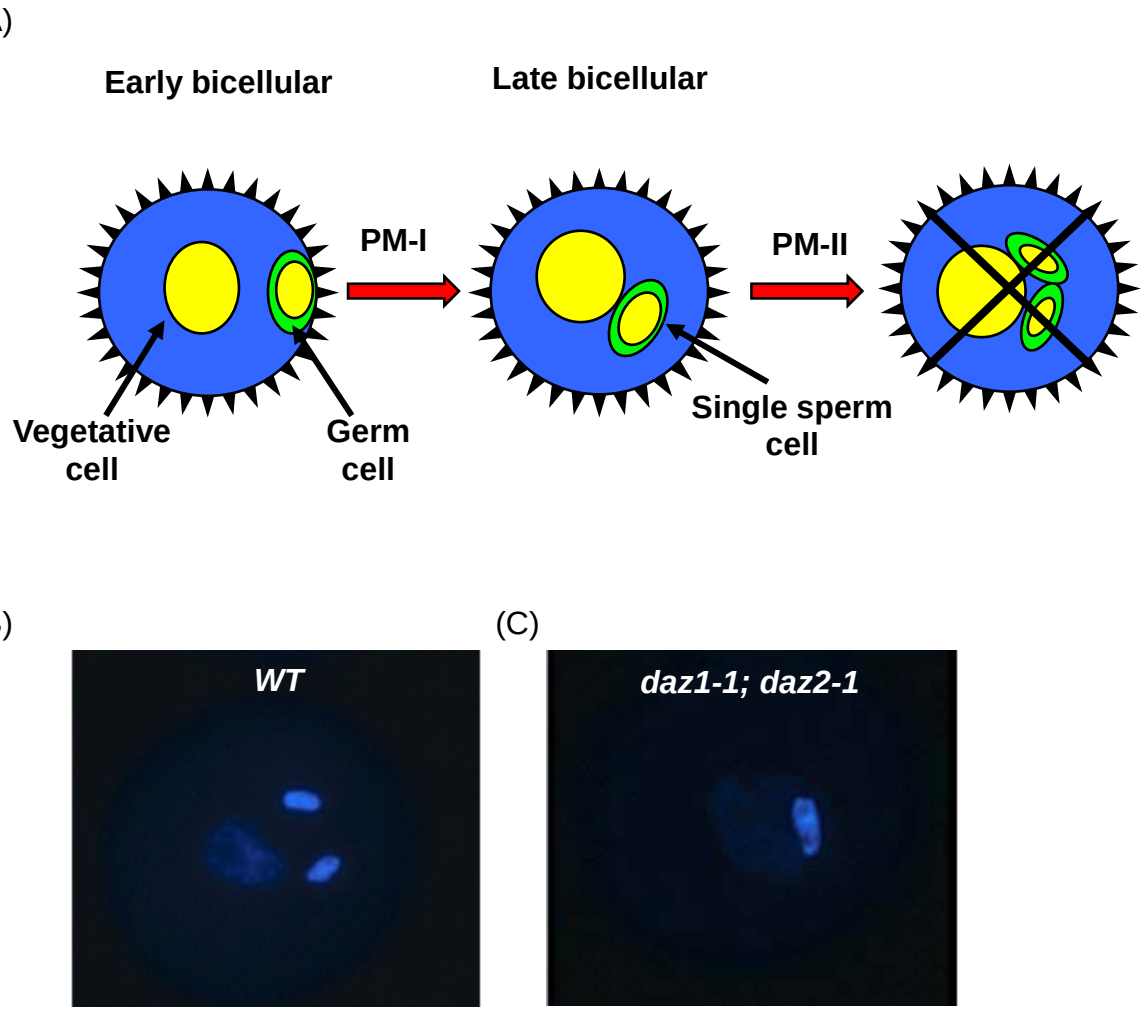


Figure 1.6. Phenotype of wild type and *daz1-1 daz2* pollen. (A) Diagram illustrating that mutant *daz1-1 daz2-1* pollen is bicellular as the germ cell fails to enter pollen mitosis II. Images shows two different pollen phenotypes after DAPI-staining. (B) Wild type tricellular pollen, (C) *daz1-1 daz2-1* bicellular pollen.

The transition of cells from G2 to mitotic phase is recognised by the accumulation of a specific class of mitotic or B-type cyclins (Borg *et al.* 2014). It was previously shown that a CYCB1:1:DB-GUS reporter in which the CYCB1:1 promoter and N-terminal

mitotic destruction box was fused with the *GUS* gene, does not accumulate in *duo1-1* germ cells (Brownfield *et al.* 2009). A similar CYCB1:1:DB-GFP fusion was monitored in *daz1-1 daz2-1* germ cells and was shown to be reduced but not absent in *daz1-1 daz2-1* germ cells (Borg *et al.* 2014). Thus, *daz1-1 daz2-1* germ cells may fail to divide (Figure 1.6, A) because of other promoting and inhibitory factors including CYCB1;2 in addition to reduced CYCB1;1 (Borg *et al.* 2014).

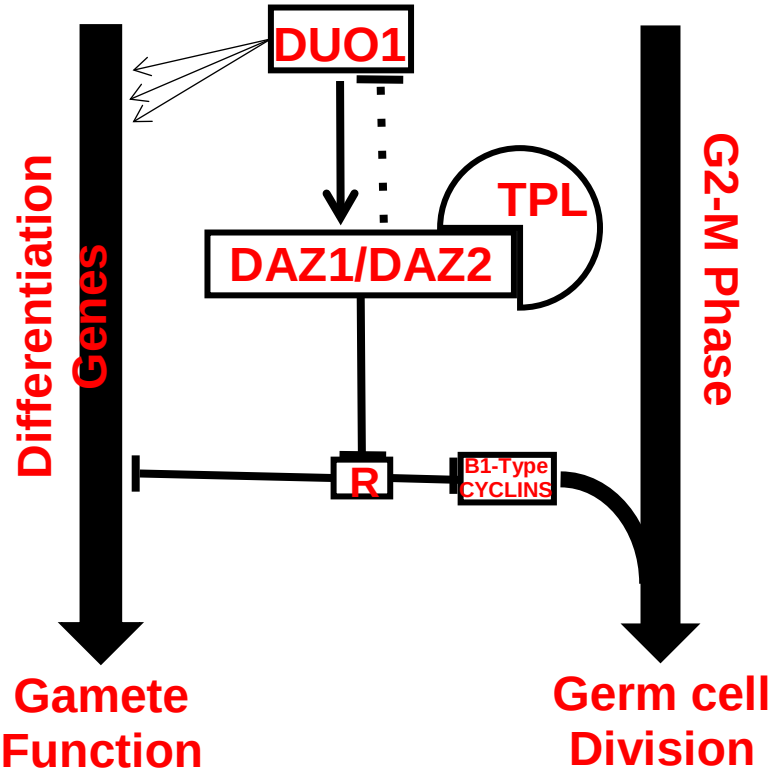


Figure 1.7. Model of DUO1-DAZ1/DAZ2 role in male gametogenesis. DAZ1/DAZ2 form a regulatory node down steam of DUO1 which is required for the germ cell division and gamete function. DAZ1/DAZ2 interact with the corepressor TPL (TOPLESS) which is proposed to supress an unknown negative regulator (R) to control gene expression. DAZ1 and DAZ2 have been shown to physically interact with the co-repressor TPL in yeast 2-hybrid assays and this interaction is dependent on the EAR motifs (Borg *et al.* 2014).

The presence of DAZ1 and DAZ2 are required for the mitotic division of the generative cell (Figure 1.6, B). Several lines of evidence support this hypothesis, first, the failure of *daz1-1; daz2-1* germ cells to enter mitotic phase does not result from in complete replication, since mutant cells seem to re-enter S-phase. Further, *daz1-1; daz2-1* shows reduced mitotic B1 cyclin expression and endo reduplication is known to relate to reduced G2/M – cyclin – dependent kinase activity (Figure 1.7) (Cebolla *et al.* 1999, Kiang *et al.* 2009, Borg *et al.* 2014). Second, DAZ1/DAZ2 germ cells fail to enter mitosis even though DUO1 is still expressed and importantly the expression of DAZ1 in *duo1* mutant germ cells can restore mitotic division independently of gamete differentiation (Figure 1.7) (Borg *et al.* 2014).

1.7.3 EAR motif function and conservation

Ethylene-responsive element binding-factor-associated amphiphilic repression (EAR) motif-mediated transcriptional repression is emerging as one of the principal mechanisms of plant gene regulation (Kagale and Rozwadowski 2011). The EAR motif was the first active repression motif reported in plants. It was originally identified almost a decade ago in a subset of class II ERFs and TFIIIA- type zinc finger proteins as a small motif with a consensus of L/FDLNL/F(x) P (Ohta *et al.* 2001). The EAR motif, defined by the consensus of either LxLxL or DLNxxP, is the predominant transcriptional repression motif so far identified in plants. Additionally, this motif is highly conserved in transcriptional regulators known to function as negative regulators in a broad range of developmental and physiological processes across evolutionarily diverse plant species. C2H2-type zinc finger proteins that contain the ERF-associated amphiphilic repression (EAR) domain are thought to play an important role in regulating the defence response of Arabidopsis to abiotic stress (Ohta *et al.* 2001, Kazan 2006). Recent discoveries of co-repressors interacting with EAR motifs, such as TOPLESS (TPL) and AtSAP18, have begun to unravel the mechanisms of EAR motif-mediated repression (Kagale and Rozwadowski 2011). Approximately 72 % of these proteins contain an LxLxL type of EAR motif, 22% contain a DLNxxP type of EAR motif, and the remaining 6% have a motif where LxLxL and DLNxxP are overlapping. *In vitro* and *in planta*

investigations indicate that approximately 40 % of these proteins may function as negative regulators of gene expression (Kagale and Rozwadowski 2011).

1.7.4 TOPLESS gene repression in Arabidopsis

Gene expression is elementary to an organism's capability to react to its environment and this underpins the processes of development and differentiation (Causier *et al.* 2012). The Groucho (Gro)/Tup1 family, first identified in *Drosophila* and *Saccharomyces*, represents an archetypal class of co-repressors that are recruited by a range of DNA-binding transcription factors to elicit repression of transcription (Liu and Karmarkar 2008). The first Gro/Tup1 family member to be isolated in plants was *LEUNIG* (LUG) (Conner and Liu 2000). LUG was identified due to its repressive activity, restricting the expression domain of the floral homeotic gene AGAMOUS (Conner and Liu 2000). Another, Gro/Tup1 corepressor family, TOPLESS (TPL/TPR), has been identified in plants, which interacts directly or indirectly with transcription factors (Kieffer *et al.* 2006, Long *et al.* 2006).

1.7.5 Nuclear Localisation Signals (NLS)

A nuclear localisation signal (NLS) is a short stretch of amino acids that mediates the transport of nuclear proteins by binding to their receptors, known as importins or karyopherins (Görlich and Kutay 1999). Positively charged residues are common in NLSs and some of these bind to importins, which recognised by the nuclear pore complex (Richardson *et al.* 1988, Garcia-Bustos *et al.* 1991, Conti *et al.* 1998). Changing the positive charge in this region is often the simplest way to disrupt nuclear import; however, there are glycine-rich motifs with few positive charges (Bonifaci *et al.* 1997). Two main approaches have been used to test the role of putative NLS motifs, (1) deletion of the NLS or substitution of basic amino acid residues, (2) fusion of a putative NLS to a normally non-nuclear protein (Tinland *et al.* 1992, Moede *et al.* 1999). For instance, spatially regulated nuclear localization of the Drosophila Dorsal protein dictates the development of dorsal-ventral pattern (Roth *et al.* 1989, Rushlow *et al.* 1989, Steward 1989).

In plants, nuclear localisation of some proteins involved in mediating photo-morphogenesis is light-regulated (Harter *et al.* 1994, von Arnim and Deng 1994). In 1998, Chattopadhyay lab studied that in *Arabidopsis thaliana HY5* gene has been defined genetically as a positive regulator of photo morphogenesis and shown to encode a basic leucine zipper type of transcription factor and reported that *HY5* gene is constitutively nuclear localised and is also involved in light regulation of transcriptional activity of the promoter containing the G-box, a well-characterised light responsive element (LRE) (Chattopadhyay *et al.* 1998). Additionally, *HY5* is a nuclear protein and its nuclear localisation pattern is independent of light stimuli and cell types, therefore, light modulation of *HY5* activity most likely occurs within the nucleus rather than by influencing its nucleo cytoplasmic distribution, which is distinct from that reporter for other G-box binding factors, such as, GBF2 (Harter *et al.* 1994, Terzaghi *et al.* 1997). Furthermore, the FAR1 (FAR-RED IMPAIRED RESPONSE 1) sequence also contains a basic regions, which could act as a nuclear localisation signal (NLS). The FAR1 amino acid sequence contains the monopartite NLS-RKRK, which together with other basic residue nearby, could contribute to nuclear localisation (Hudson *et al.* 1999).

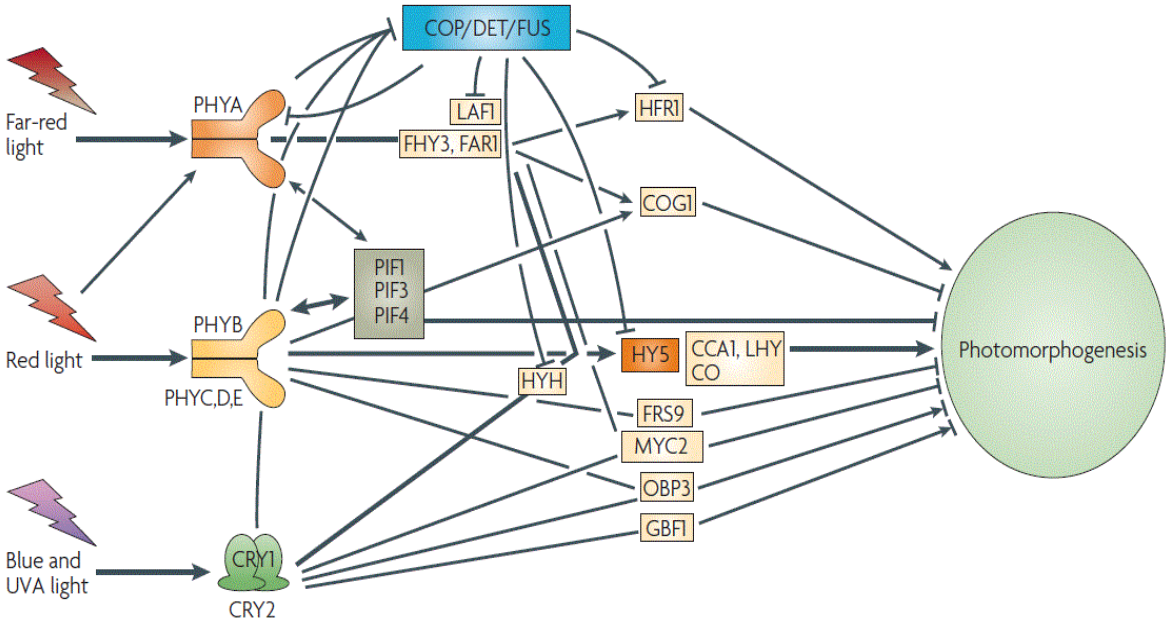


Figure 1.8. Transcriptional networks for seedling photo-morphogenesis. A simplified overview of the network involved in this process in shown. Importantly key regulators of this-regulated transcriptional network have been identified in *Arabidopsis thaliana* and suggest that existence of separate intermediate network that are dedicated to each photoreceptor group. A group of PIF transcription factors interact directly with phytochromes and function mainly as repressors of photo-morphogenesis. Key transcription factors, such as HY5, serve as single integration points of major branches downstream of all photoreceptor. The COP/DET/FUS class of factors act as light-inactivatible repressors of photo-morphogenesis. In figure bold lines indicates the convergence pathway (Jiao *et al.* 2007).

Genetic and genomic analyses suggest the existence of several signalling pathways downstream of PHYA in photo-morphogenesis (Figure 1.8) (McCormac and Terry 2002, Wang *et al.* 2002). FAR-RED IMPAIRED RESPONSE 1 (FAR1) and FAR-RED ELONGATED HYPOCOTYL 3 (FAY3) are both novel transposon derived putative transcription factors, which interact with each other and are specific to far-red light (Hudson *et al.* 1999, Hudson *et al.* 2003), whereas LAF1 is a transcription factor that is homologous to the R2R3-MYB family of DNA binding

proteins(Ballesteros *et al.* 2001). Loss of function mutants of *far1, fhy3 or laf1* show developmental defects in PHYA mediated seedling photo-morphogenesis in response to far-red light, whereas they show no obvious phenotype under other light qualities (Hudson *et al.* 1999, Wang and Deng 2002, Hudson *et al.* 2003). Additionally, similar light-hyposensitive phenotype was initiate, under both far-red and blue light conditions (Fairchild *et al.* 2000, Duek and Fankhauser 2003) in mutants of the *LONG HYPOCOTYL IN FAR-RED* (*HFR1*) gene, implying its role in both PHYA and crypto chrome signalling (Jiao *et al.* 2007). FHY3 and FAR1 most likely act more upstream in the network, close to PHYA (Wang *et al.* 2002), whereas, HFR1 probably acts further downstream as microarray studies found that HFR1 controls the expression of a smaller subset of genes (Kim *et al.* 2002, Wang *et al.* 2002, Yang *et al.* 2003). HY5, a bZIP transcription factor, might represent another branch under FHY3/FAR1, although its function is not limited to far-red light (Figure 1.8) (Jiao *et al.* 2007).

1.8 *Arabidopsis* gene expression during male germline development

In recent years, highly advanced technologies have uncovered male gametophyte gene expression on large scale. Some early studies exploited serial analysis of gene expression (SAGE) technology (Lee and Lee 2003) and 8K Affymetrix AG microarray (Becker *et al.* 2003); (Honys and Twell 2003) to analyse the pollen transcriptome. The Affymetrix 23K *Arabidopsis* ATH1 array technique covers roughly 80% of *Arabidopsis* genes. Publicly available data provide more insight into male gametophyte gene expression which span four stages of male gametophyte development (1) uninucleate microspores, (2) bicellular pollen, (3) tricellular pollen, and (4) mature pollen from ecotype Landsberg *erecta* (Honys and Twell 2004). Further two datasets are derived from the Columbia ecotype (Zimmermann *et al.* 2004); (Pina *et al.* 2005).

Taken together these datasets indicate that the number of genes expressed in pollen is 5000 to 7000 however, further analysis involving pollen development, extends the total number to approximately 14,000 genes (Honys and Twell 2004, Twell *et al.*

2006). The developmental transcriptome showed a decrease in transcriptome size from 11,565 genes in uninucleate microspore (UNM) to 7235 genes in mature pollen. On the other hand, the percentage of the pollen-specific genes increased from 6.9 % at UNM to 8.6 % in mature pollen reflecting differentiation and functional specialisation of mature pollen (Rutley and Twell 2015). Interestingly, 612 out of 1350 predicted transcription factors present on the *Arabidopsis* ATH1 genome array are expressed in developing male gametophytes (Twell *et al.* 2006). Combined transcriptomic and mutational analyses further indicate that groups of such transcription factors initiate a late pollen regulatory network that is likely to play important roles in pollen development (Borg *et al.* 2009).

In flowering plants, the male germline development creates unique contribution to the pollen transcriptome (Engel *et al.* 2003). Recent research shows the sperm cell transcriptome of 5829 genes was smaller than that of mature pollen (7177 genes) with only 65.4% of genes in common. However, comparing the seedling and pollen transcriptomes approximately 2400 transcripts were enriched in sperm with 642 detected exclusively in sperm (Borges *et al.* 2008). For example, the DNA methyltransferase (MET1) is enriched in sperm, consistent with the active role of MET1 in the epigenetics inheritance of CG-context methylation (Saze *et al.* 2003, Saze 2008, Calarco *et al.* 2012, Rutley and Twell 2015).

1.9 Seed development and abortion in flowering plants

Seed growth and development during reproduction signify a major investment of resources through the maternal plant. Each seed must have enough reserves to go through the initial stages of germination and seedling formation. However, successful establishment will depend, at least in part, on the overall resources of each seed (Mena-AlÍ and Rocha 2005). Previous lab member Dr. Rutley found that DAZ1 expressing mutants ΔEAR1,2 lines (D2.2) shows aborted ovules which was not predicted furthermore, mutant DAZ1 variants (smEAR1,2 and muZnF2) tested and shows shame seed abortion phenotype because, the numbers of seeds within a fruit is the final result of the sequential process that starts with an initial number of

ovules in the ovary and ends with the successful development of fertilised ovules. Therefore, in mutant DAZ1 variants ovule abortion is not usual and predicted thus, need to understand and explore the further insight in seed development stages in flowering plants. The upper limit for seed number per fruit is set by the number of ovules and the seed to ovule ratio is mainly determined by fertilisation success and seed abortion (Calviño 2014). However, even after successful fertilisation, abortion of flower and fruits occurs at different developmental stages. Abortion of developing embryos can substantially reduce the final number of seeds and large-scale abortion of flower and fruits is a common phenomenon in nature (Bawa and Webb 1984, Calviño 2014).

Most plants producing healthy seed set in fruits depend on the fitness of the plants or on the progeny and plants commonly produce large number of aborted seeds with reduced proportions of inbred offspring (Latta 1995). Several studies document that late stages of seed development are less affected by environmental factors, but developing ovules and young embryo are often affected by stress, resulting in aborted ovules and embryos (Sun *et al.* 2004). Failure of seed abortion is also subject to genetic load caused by deleterious recessive mutations, indicating that removal of inbred offspring may prevent the expression of mutations as homozygotes (Latta 1995).

Genetic and molecular factors play a vital role in seed development and abortion in plants. Although extensively studied, current knowledge does not provide complete information about the transcription factors (TFs) that directly or indirectly control seed development. For example, Le et al. (2010) identified 289 seed-specific genes, including 48 that encode transcription factors. Their analysis showed that seven of these seed-specific TF genes encoded known regulators of seed development in *A. thaliana* such as LEAFY COTYLEDON (LEC), LEC1, LEC1-LIKE, LEC2 and FUS3 (Brandon H Le *et al.* 2010). Further, wide knowledge of seed-specific TFs provides valuable information about the extensive gene networks, which coordinate seed development to make viable offspring (Brandon H. Le *et al.* 2010).

In *Arabidopsis thaliana* it has been unclear how pollen tube growth, fertilisation and ovule growth may be linked. In recent research by Kasahara et al (2016), this issue addressed in the discovery of pollen tube-dependent enlargement of ovule morphology (POEM). Their analysis revealed expression of the embryo and endosperm markers, WOX9 and AGL62 soon after fertilisation at 12 hours after pollination (HAP) with wild type pollen, but not after pollination with *gcs1* pollen indicating no fertilisation event. In *A. thaliana gcs1* mutants sperm fail to fuse which leads to pollen sterility showing that that GCS1 (GENERATIVE-CELL-SPECIFIC1) is a critical fertilisation factors in angiosperms (Mori *et al.* 2005). Further, transcriptome analysis of ovules at 12 HAP after crossing with wild type or *gcs1* pollen revealed 24 common up-regulated genes in WT and *gcs1* crosses compared to ovules without pollination (Kasahara *et al.* 2016).

1.10 Mutagenesis in flowering plants

Mutagenesis of flowering plants and breeding have rapidly advanced with improvements in high-resolution molecular mapping and biochemical analytical techniques. Unique mutagenised populations combined with novel screening methods for traits that are difficult to generate and identify by conventional breeding are now being developed and characterised at the molecular level (Sikora *et al.* 2011). Treatment with mutagens alters genes or breaks chromosomes leading to mutations while some gene mutations occur naturally as errors in DNA replication. Although some of these mistakes can be repaired others may be transmitted and result in phenotypic variation in offspring(Novak and Brunner 1992).

Genetic mutations can result in heritable changes in DNA and occur naturally due to DNA damage caused by environmental factors such as UV, radiation and due to errors in DNA replication. Mutations contribute to evolution as alterations that are passed on to offspring can lead to phenotypic variation and novel individual genes. The artificial induction of mutations in wild type drosophila, maize and barley was reported in the 1920s. Since that time, many individual researchers used this technique. Mutations are introduced in plants by exposure of their propagules, like

seeds and meristematic cells, tissue and other organs. The aim of these approaches is to develop more genetic variation with which to allow the directed combinations of new traits or alleles (Novak and Brunner 1992).

There are three ways by which to induce mutations by either using (1) physical agents such as fast neutron, UV and X-ray radiation, and (2) chemical agents such as N-methyl-N-nitrosourea (MNU), 1, 2: 3, 4-diepoxybutane (DEB) or ethyl methane sulfonate (EMS) and (3) biological agents such as transposons and T-DNA (Serrat *et al.* 2014). EMS has become one of the most effective, reliable and frequently used chemical mutagens in plant research (Brockman *et al.* 1984). EMS mostly induces C to T substitutions resulting in C/G to T/A transitions (Krieg 1963, Kim *et al.* 2006) and at a low frequency, G/C to C/G or G/C to T/A transversion through 7-ethylguanine hydrolysis or A/T to G/C transitions through 3-ethyladenine paring errors (Krieg 1963, Greene *et al.* 2003, Serrat *et al.* 2014). Other mutagens such as sodium azide (AZ) and methylnitrosourea (MNU) are also used and often combined into an AZ-MNU solution, which causes GC to AT shifts, or AT to GC shifts (Sikora *et al.* 2011).

The chemical mutagen, ethyl methanesulfonate characteristically introduces dozens of mutations in each plant, and it is conceivable to find a mutation in any specified gene by screening fewer than 5000 plants from the mutagenized M1 generation (Feldmann *et al.* 1994, Greene *et al.* 2003). Isolation of dominant mutations can be found in the M1 generation, but these are mostly rare phenomena and frequently, recessive mutations are recovered (McConnell *et al.* 2001). Typically M1 seeds are planted and grown for maturity and resulting M2 seeds are screened for phenotypes (Østergaard and Yanofsky 2004).

EMS induced mutants have been induced in many plant species including crops species such as rice, maize, barley, and sorghum (Caldwell *et al.* 2004, Till *et al.* 2004, Gilchrist *et al.* 2006, Till *et al.* 2007, Xin *et al.* 2008). However, in comparison with insertional mutagenesis, chemically induced mutations generate a greater

diversity of alleles and are high productive because, each distinct line can bear single point missense and nonsense alternates in hundreds of genes (Homikoff, Till, and Comai, 2004). Therefore, an allelic series of induced mutations with dissimilar effects on gene function can be easily inaccessible by screening a comparatively small population of mutated plants (Østergaard and Yanofsky 2004).

An alternative to performing mutagenesis screen is to exploit natural variation that is recognised to occur among the hundreds of divergent ecotypes of Arabidopsis that have been isolated from around the world. One such example is the allelic variation among Arabidopsis ecotypes at the FRIGIDA (FRI) locus that was shown to be an important determinant of flowering time (Johanson *et al.* 2000, Østergaard and Yanofsky 2004).

1.11 Aims and objectives

The abundance of genetic and genomic resources available for *Arabidopsis thaliana* and other plant species provide increasing opportunities to understand male gametophyte development at the molecular level. The combination of transcriptome studies and genetic analysis led to the identification of DAZ1 and DAZ2, which act redundantly to regulate generative cell division and differentiation in *Arabidopsis thaliana* (Borg *et al.* 2011, Borg *et al.* 2014).

The overall aim of this research was to advance understanding of the molecular mechanisms by which DAZ1/DAZ2 proteins support male germline development in *Arabidopsis thaliana*. In DAZ1 and DAZ2 a basic region (BR) shared by these proteins near the N-terminus and a second conserved region (CR) is present between ZnF1 and ZnF2, but the functional roles of these regions are unknown. Previous work has shown that truncation of DAZ1 (DAZ1ΔEAR1,2) leads to reduced fertility and has a severe effect on male transmission (Borg *et al.* 2014). Moreover, the expression of the C-terminally truncated DAZ1ΔEAR1,2 variant in a double mutant *daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$ null background gave rise to siliques containing a few

viable seeds and more aborted seeds. This raised the question of whether DAZ1 has any post-fertilisation role or whether failed fertilisation events arising from partial DAZ1 function may be involved in ovule abortion.

The major objective, documented in Chapter 3 was to explore the functional importance of the DAZ1 BR and DAZ1 CR. *In planta* complementation analysis of *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ mutants with DAZ1 mutant proteins affecting BR and CR sequences were used to understand their importance. In addition, transient expression assays were performed in tobacco to study DAZ1 protein localisation and function. The localisation of DAZ1, DAZ2 and DAZ1 variants were studied in germ cells and in tobacco leaf cells and compared with DAZ1 orthologs from several species. Previous work had concluded that ZnF1 has a critical roles in DAZ1 function (Rutley, 2014). A further objective reported in Chapter 3 was to validate and extend evidence for the relative importance of the different classes of ZnF domains for DAZ1 function.

The second objective, presented in Chapter 4, was to study the seed abortion phenomenon observed in *daz1 daz2* mutants expressing DAZ1ΔEAR1,2-mCherry. This was analysed by examining ovule, seed and silique development following controlled crosses. A further aim was to investigate whether the seed abortion occurred in other *daz1 daz2* mutants expressing DAZ1 mutants (muZnF1, muZnF2, muZnF1,2, muZnF2,3, and smEAR1,2) affecting different domains for protein localisation. Due to seed abortion in mutant DAZ1 domains and motifs also arise another question whether DAZ1 has significant role after germ cell division in fertilisation process?

The third objective, presented in Chapter 5 was to screen for genetic suppressors to find novel factors which may interact with DAZ1 or which may regulate DAZ1-dependent germ cell development. Chemically mutagenised plants which express DAZ1ΔEAR1,2 variant in a *daz1 daz2* null background showing highly reduced fertility were screened for fertility restoration.

Chapter 2

Materials and methods

2.1 Seed sowing and Growth condition

Arabidopsis thaliana plants of ecotype Columbia (Col-0) were grown on soil composed of 3:1:1 compost (Levington), vermiculite and sand. Seeds were sown in soil sub irrigated with herbicide basta (200 µl/L) diluted in water. Trays were incubated in a 4 ºC cold room for 2 days and then transferred to greenhouse with 16 hours light regime at 22ºC. Prior to plating, seeds were sterilized with 70 % (v/v) ethanol for 5 minutes and dried on sterile filter paper for 20-25 minutes. Seeds were placed on MSO growth media with or without antibiotic selection. Plates were incubated at 4 ºC for 2 days prior to moving to growth rooms under continuous illumination.

2.2 Preparation of media for plant tissue culture

Arabidopsis thaliana seeds sown on Murashige and Skoog media. Approximately 4.3 g of MS salts dissolved in 950 ml of deionised water with no added sucrose (MS0) or with 1-3% (w/v) sucrose. The pH of the media was adjusted to 7.0 with 5N NaOH and autoclaved for 20-25 minutes at 120 ºC and 15 psi on liquid cycle. To make MS agar for sowing seeds, 0.6% (w/v) of phytoagar added prior to autoclaving.

2.3 Antibiotics for selection of transgenic plants

Antibiotics	Working concentration
Gentamycin (Melford Lab, UK)	100 µg/ml
Kanamycin (Melford Lab, UK)	50 µg/ml
Glufosinate ammonium (BASTA) (Buyer)	30 µg/ml
Hygromycin (CalBioChem)	20 µg/ml

To make working stocks, antibiotics were dissolved in deionised water and stored at -20ºC.

2.4 Bacterial culture and storage

2.4.1 Bacterial strains

Escherichia coli (E. coli)	*Agrobacterium tumefaciens*
α-select (Bioline)	GV3101 (pMP90)

2.4.2 Long-term bacterial strain storage

A single bacterial colony was used to inoculate a falcon tube with 5 ml liquid LB medium with selected antibiotics and transferred to a shaker at 220 rpm set at 37ºC overnight. The overnight bacterial culture (700 µl) was added to a cryogenic storage tube and missed with 300 µl sterile 50 % (v/v) glycerol. Prior to storage at -80ºC, the tube was vortexed briefly and the tube flash frozen in liquid nitrogen. As per requirement, bacterial strains were recovered by scraping a small portion of the frozen culture with a wire loop and placing on agar medium containing selected antibiotics for subsequent growth.

2.4.3 Antibiotics for bacterial selection

Antibiotics	***E. coli***	***A. tumefaciens***
Gentamycin	na	50 µg/ml
Rifampicin	na	50 µg/ml
Spectinomycin	100 µg/ml	100 µg/ml
Kanamycin	50 µg/ml	50 µg/ml

To make the working stocks shown above, antibiotics were dissolved in deionised water and stored at -20ºC.

2.4.4 Streaking out bacterial strains on plates

Bacterial glycerol stock tubes were kept on dry ice from -80ºC. A platimun wire loop was flame sterilised and allowed to cool down before touching the bacterial culture. The loop was touched onto the plate and used to streak out bacterial strains on a plate in several directions, sterilising between each direction to obtain single colonies. Plates were sealed with parafilm and incubated upside down in an incubator overnight (*E. coli* - 37ºC) or for 2-3 days (*Agrobacterium* – 28ºC).

2.4.5 Culture of bacteria

Agrobacterium tumefaciens and *E. coli* were grown in Luria-Bertani medium (LB) containing 1% (w/v) of tryptone, 1% (w/v) NaCl and 0.5 % (w/v) yeast extract. For plate selection, 1.5 % (w/v) bactoagar was added to the media before autoclaving. The pH was adjusted to 7.0 with 5N NaOH and autoclaved for 20 minutes at 120ºC at 15 psi on a liquid cycle. After inoculation cultures were placed in an incubator or shaker at 220 rpm set at 37ºC for *E. coli* and 28ºC for *A. tumefaciens*.

2.4.6 Transformation of *E. coli* with plasmid DNA

For transformation of *E. coli*, a 200 µl aliquot of competent cells was removed from -80ºC freezer and thawed on ice for 10 minutes. 50 µl of the competent cells were added to labelled eppendorf tubes. 1 µl (~125 ng) of plasmid or half of the recombination reaction was added into a 50 µl aliquot, and after gently flicking tubes incubated for 30 minutes on ice. The mixture was heat shocked in a water bath at 42ºC for 45 seconds and put back on ice for 2 minutes. 500 µl of LB media was added and the culture was incubated at 37ºC for 1 hr. 100 µl of cell suspension were placed on LB agar plate containing antibiotics and plates incubated at 37ºC overnight. Colonies were screened by colony PCR.

2.4.7 Preparation of chemically competent *A. tumefaciens*

The *Agrobacterium* strain (GV3101) containing the T_i plasmid pMP90 was grown in 5 ml of LB media with selected antibiotics (Rifampicin – 50 μg/ml, Gentamycin – 100 μg/ml) transferred into shaker for overnight at 28 ºC – 200 rpm. A 2 µl aliquot of the overnight culture was added to 50 ml LB medium in a 250 µl flask. The culture was grown at 28ºC with vigorous shaking until the OD_{600} reached 0.5 –1.0 The cell suspension was then cooled on ice and centrifuged at 3000g for 10 minutes at 4ºC. The supernatant was discarded and cells were resuspended in 1 ml of 20 mM $CaCl_2$ solution. A 100µl aliquot of the cells was rapidly dispensed into pre-chilled microfuge tubes, flash frozen in liquid nitrogen and stored at -80ºC.

2.4.8 Transformation of *A. tumefaciens* with plasmid DNA

For the transformation of *Agrobacterium*, 0.5-1µg of plasmid DNA was added into the frozen aliquot and incubated at 37ºC for 5 minutes to heat shock the cells. 1 ml of LB medium was added to the mixture, which was incubated with gentle shaking at 200 rpm and 28ºC for 3 hours. The culture was centrifuged for 30 seconds at 3000g and the supernatant discarded. Transformed cells again resuspended with the remaining solution and spread on an LB agar plate containing selected antibiotics. Plates were incubated at 28ºC for 2 days and colonies were screened by colony PCR.

2.5 Nucleic acid analysis

2.5.1 Purification of high quality genomic DNA

For the purification of genomic DNA, about 1-2 rosette leaves from selected plants were collected into a labelled 1.5 ml microcentrifuge tube containing 200 glass beads (Sigma-Aldrich) with a size of 425-600 microns. Leaf tissues were frozen in liquid nitrogen and ground for 12 second by using silamat amalgam mixer. 250 µl of DNA extraction buffer [1.4 M NaCl, 3 % (w/v) CTAB, 20m MEDTA, 100 mM Tris-

HCL pH 8.0] was added to the tube. Microcentrifuge tubes were briefly vortexed and incubated for 15-20 minutes at room temperature. An equal volume (~250 µl) of chloroform: IAA (24:1) was added and the tube centrifuged for 10 minutes at 13,000g. The upper aqueous layer (~200 µl) was transferred to a fresh microcentrifuge tube containing 0.7 volume (~140 µl) of isopropanol. The mixture was incubated for 5 minutes at room temperature and centrifuged for 7 minutes at 13,000g. The supernatant was discarded and the pellet washed with 1 ml of 70 % (v/v) ethanol before centrifugation for 5 minutes at 13,000g. After removal of ethanol from the tube the pellet was vacuum dried for 10 minutes. The dried pellet was dissolved with 100 µl of 10 mM Tris 1m MEDTA pH 8.0 (TE) and incubated at 55ºC for 5 minutes before storage at -20ºC.

2.5.2 Plasmid DNA isolation from bacteria

Isolation of plasmid DNA from *E. coli* and *A. tumefaciens* was carried out using a GenElute Plasmid Miniprep Kit (Qiagen, UK) according to the manufacturer's instructions. A culture was grown overnight from a single colony in 5 ml LB media containing selected antibiotics. 1.5 ml of the culture was transferred into a microcentrifuge tube and centrifuged for 1 minute at 13,000g. For *E. coli* this step was repeated twice and for *A. tumefaciens* four times to achieve high yields of plasmid DNA. The supernatant was discarded and the pelleted cells resuspended in 200 µl resuspension solution by brief vortexing until homogenous. 200 µl of lysis solution was then added to lyse the cells and tubes were gently inverted 6-8 times until a clear and viscous mixture formed. Immediately, 350 µl of neutralization/binding solution was added to precipitate cell debris from the mixture by gently inverting 4-6 times and the lysed bacterial cells centrifuged for 10 minutes at 13,000g. Meanwhile, a binding column was prepared by adding 500 µl of column preparation solution to a genelute miniprep binding column, followed by centrifugation at 13,000g for 1 minute. The supernatant was discarded and the upper clear solution transferred to a prepared spin column and centrifuged for 1 minute at 13,000g. The supernatant was discarded and 500 µl of optimal wash solution was added to the spin column followed by centrifugation for 1 minute at 13,000g. The

column was then washed with 750 µl of diluted wash solution containing 70 % (v/v) ethanol and centrifuged at 13,000g for 1 minute. After discarding the supernatant, the column was centrifuged for 2 minutes at maximum speed and any remaining ethanol removed. The column was then transferred to a clean labelled collection tube and 50 µl of 10mM Tris-HCl, pH8.0 added to the centre of the column. The column was incubated at room temperature for 1 min before centrifugation for 1 min at 13,000g to elute the plasmid DNA.

2.6 DNA sequencing, cloning and modification

2.6.1 Gateway cloning

Cloning was performed by using Gateway recombination technology (Invitrogen, UK). DNA fragments were first amplified by PCR using primers that add recombination sites (attB1 and attB2). The PCR product was purified either by gel extraction or directly from the reaction. The concentration of purified products together with the vector was analysed by agarose gel electrophoresis by using a standard DNA marker ladder. The PCR products were cloned into expression vectors in two steps.

Step – 1 BP reaction: - First step, 125 ng of PCR product and 75 ng of donor vector were used in a BP reaction to generate entry clones. The mixture was incubated overnight at room temperature and the reaction was terminated by adding 0.5 µl of proteinase K and incubation at 37ºC for 5 minutes. 2.5 µl of the total reaction was used for transformation of *E. coli* cells (α-select) of 10^8 efficiency (Bioline) and colonies were screened by colony PCR by using M13 forward and reverse primers. Positive colonies were used to set up 5 ml overnight cultures. Plasmid DNA was isolated from cultures and analysed using restriction enzyme digestion and sequencing before use in LR recombination reactions.

BP Reaction

PCR product	~125 ng
pDONR vector	~75 ng
BP Clonase II	1.0 µl
TE buffer	5.0 µl

Step -2 LR reaction: - Verified entry clones were used to recombine the insert into the preferred destination vector in an LR reaction. For single site LR reactions, the recombination of 125 ng of entry clone and 75 ng of destination vector was catalysed with the LR Clonase II+ enzyme. In multisite LR reactions, recombination of 5 *f*mol of each entry clone and 10 *f*mol of the destination vector with LR Clonase II+ enzyme was used. The formula below was used to calculate the amount (ng) of plasmid DNA required to achieve the desired fmol is as follows:

ng needed = desired *f*mol (bp) X size of vector (bp) X (660×10^{-6})

(Where bp is the size of DNA in bp)

Reactions were incubated overnight at room temperature and terminated by adding 0.5 µl of proteinase K and incubating at 37ºC for 5 minutes. 2.5 µl of the total reaction was used to transform *E. coli* cells (α-select) of 10^8 efficiency (Bioline). Colony PCR using gene specific primers was used to identify positive colonies which were used to set up 5 ml overnight cultures. Plasmid DNA was extracted from the cultures and analysed using restriction enzyme digestion. Plasmids were then transformed into *A. tumefaciens* (GV31010) for stable transformation of *Arabidopsis thaliana.*

Three-site LR reaction

Destination vector	10 *f*mol
pDONRP4-P1R entry clone	5 *f*mol
pDONR221 entry clone	5 *f*mol
pDONRP2R-P3 entry clone	5 *f*mol
LR Clonase II +	1.0 µl
TE buffer	5.0 µl

2.6.2 Primer design using MacVector

The main purpose of oligonucleotide primer design was to amplify genomic DNA, cDNA and plasmid DNA, in addition to mutant DNA sequences. Primers were designed manually by using MacVector software (Version 12.6.0). Generally, primers were designed between 18 – 25 bases in length with a Tm of 50 - 60ºC and when possible, with a G or C at the 3' end. Pairs of primers were designed to have a Tm difference of ≤ 5ºC. Primers were designed and used in the gateway recombination cloning of PCR amplification products through PCR where partial attB adaptors were added to the 5' end of products to allow recombination with pDONR vectors. The partial adapter primers are shown in (Table 2.1).

Table 2.1. Partial attB adapter primers and oligonucleotide sequences along with compatible pDONR vectors.

Partial attB adapter	Sequence (5' – 3')	pDONR vector
attB4F	TGTATAGAAAAGTTG	pDONRP4-P1R
attB1R	TTTTGTACAAACTTG	pDONRP4-P1R
attB1F	AAAAAGCAGGCT***NN****	pDONR221
attB2R	AGAAAGCTGGGT***N****	pDONR221
attB2F	TCTTGTACAAAGTGG***NN****	pDONRP2R-P3
attB3R	TGTATAATAAAGTTG***N****	pDONRP2R-P3

N* - nucleotide(s) added to keep sequence in frame.

Primers were ordered and purchased from Sigma-Aldrich Company. Lyophilised primers were dissolved in an appropriate volume of TE buffer (10 mM Tris – CL (Ph 7.5), 1 mM EDTA) to make a primer stock of 100 µM (100 pmol/µl). Working stocks were prepared at concentrations of 5 µM (10 µl primer stock + 190 µl nuclease free water). Aliquoted primers were stored -20ºC.

2.7 PCR for DNA amplification

2.7.1 General PCR set up and conditions

Amplification of genomic and plasmid DNA by general PCR with using Biotaq polymerase. A 20 µl master mix was prepared by using relevant buffers, primers and Biotaq enzyme. 1:100 dilutions were prepared for plasmid DNA. Prior to use, all buffers and primers were left for 5 -10 minutes to defrost and vortexed briefly. PCR conditions were set up typically as 96ºC – 5 minutes, 96ºC – 30 seconds, 55ºC – 30 seconds, 72ºC – 30 seconds and 72ºC – 5 minutes, 36 – 40 cycles.

2.7.2 Phusion PCR

For Gateway cloning, PCR samples were amplified by using DNA polymerase phusion to create point mutations in the full length DAZ1 coding sequence (CDS). 1st and 2nd PCR amplification steps introduced by overlap PCR using forward and reverse primers pairs which contain 22 bp matching the ends of the DAZ1 CDS. The 1st PCR was performed to amplify the product with partial attB primers and the 2nd PCR was performed using 1st PCR product as template to extend the attB sites. Master mixes were prepared by using buffer, PCR water, primers, dNTPs and phusion enzyme. Prior to use, all buffers and primers were left for 5 -10 minutes to defrost and vortexed briefly. 1:100 dilutions were prepared for the plasmid DNA samples for the 1st PCR. The 1st phusion PCR cycle was set up as follows: - 98ºC – 2minutes, 98ºC – 20 seconds, 55ºC – 30 seconds, 72ºC – 30 seconds and 72ºC – 5 minutes, 10 cycles. The 2nd phusion PCR cycle was set up as follows: 98ºC –

2 minutes, 98ºC – 20 seconds, 55ºC – 30 seconds, 72ºC – 30 seconds and 72ºC – 5 minutes, 20 cycles.

2.8 Gel electrophoresis analysis for high quality DNA

Agarose gel electrophoresis was carried out for PCR products size analysis, purifications of PCR products and analysis of restriction enzyme fragments. Agarose gels were made from 1XTAE (40 mM Tris base, 20 mM glacial acetic acid, 1 mM EDTA) with or without 0.2 µl/ml ethidium bromide. The agarose concentration (1.5 %) was prepared according to the size of DNA fragments being analysed. 4 µl - 8µl of PCR samples were pipetted into gel wells with 5X orange G loading dye and a DNA ladder added to a control lane to allow monitoring of DNA fragment size. Electrophoresis was carried out at 140 volts for 30-60 minutes. DNA bands were visualised with a UV trans-illuminator and the size and quantity of DNA fragments compared with the DNA ladder control.

2.8.1 Qiagen gel extraction

DNA fragments were purified from agarose gels using a Qiagen gel extraction kit (Sigma). First, the gel slice was minimised by removing extra agarose and weighed in a clean 1.5 ml eppendorf tube, to which 3 volumes of buffer were added to 1 volume (or weight) of gel, where 100 mg ~ 100 µl. The tube was incubated for 10 minutes at 50ºC, and vortexed every 2-3 minutes. The mixture turned yellow in colour once the gel slice was completely dissolved in the solution (if the colour of the mixture was orange or violet, 10 µl of 3 M sodium acetate, pH 5.0 was added and mixed to reduce the pH to <7.5 to turn it yellow. 1 gel volume of isopropanol was added to increase the DNA yield (between 500 bp and 4 kb) followed by transfer to a QIA quick spin column in a 2 ml collection tube. The ~800 µl mixture on the column was spin down and the flow through removed and this process repeated. 750 µl buffer PE was added to the column and the sample centrifuged for 1 minute. The flow through was discarded and the column centrifuged again for 1 minute at 10,000g X , then after discarding the flow through the column was transferred to a

clean 1.5 ml microcentrifuge tube. Finally, 50 µl of elute DNA solution was added to the centre of the column membrane and the sample centrifuged for 1 minute at maximum speed. The column was discarded the 1.5 ml microcentrifuge tube containing the purified DNA was stored at -20ºC.

2.9 Plant transformation

2.9.1 Stable transformation of *Arabidopsis thaliana*

Arabidopsis thaliana plants were transformed by the floral dip technique with minor modifications (Clough and Bent, 1998). Plants were grown under 24-hour incandescent light and were left to grow for ~4-5 weeks, then fully mature inflorescences were clipped and 1 week later, plants had well developed axillary inflorescences with unopened floral buds. Before transformation siliques and fully open flowers were clipped off the plant with scissors.

A single colony of a selected Agrobacterium strain was used to inoculate 5ml LB culture media with appropriate antibiotics. Cells were grown to saturation overnight at 28ºC with vigorous shaking at 220 rpm. A 400 µl aliquot of the culture was sub cultured into a 1 l conical flask containing 400 ml of fresh LB with selected antibiotics and incubated overnight at 28ºC with shaking at 220rpm. Cells were centrifuged at 3,500g in a Sorvall benchtop centrifuge for 20 minutes and resuspended in 1 l infiltration medium (2.17 g/l half strength MS salts, 3.16 g/l full strength Gamborg B5 vitamins, 0.5 g/l MES, 50 g/l sucrose, 10 µg/l 6-benzylaminopurine). Prior to dipping, 300 µl of Silwet L-77 was added per litre of resuspended culture. The inflorescences were dipped into *Agrobacterium* solution for 30-35 seconds with gentle agitation. Dipped plants were kept in a clear plastic autoclave bag for 24 hours to maintain high humidity. Plants watered normally after removal of the bag and when siliques were matured watering was stopped plants kept in a greenhouse for at least 2 weeks completely dry. Seeds were harvested after 2 weeks and sown on selection plates after sterilisation or on soil with basta to identify primary transformants.

2.9.2 Agro-infiltration of leaves

Agrobacterium-mediated transient transformation (or agro-infiltration) of *Nicotiana tabacum* and *N. benthamiana* leaves in used methods described by (Sparkes *et al.* 2006). A single colony from a preferred *Agrobacterium tumefaciens* strain GV3101 (pMP90) was used to inoculate a 5 ml LB culture media with selected antibiotics and incubated overnight growth at 28ºC with shaking at 220 rpm. Cells were harvested from a 1.5 ml aliquot of overnight culture by centrifugation at 1,000g and resuspended in 1 ml of infiltration media (280mM D-glucose, 50 mM MES, 2 mM $Na_3PO_4.12H_2O$, 0.1 mM acetosyringone), then cells were washed with another 1 ml infiltration media to remove traces of antibiotics and the OD_{600} measured with a Beckman Coulter, DU 730 – UV/Vis spectrophotometer. Four to six -6 week old greenhouse-grown plants were adapted to 24 h white fluorescent illumination for up to 1 week before infiltration. The 3rd and 4th leaves of the plants from the apical meristem were chosen infiltrated, which was performed three times on each leaf using the lamina both sides of the midrib. The *Agrobacterium* solution was taken up in a 1 ml syringe and the underside of the leaf prepared by gentle rubbing 0.5 cm^2 regions to remove the cuticle. The syringe tip was then placed against the rubbed region and *Agrobacterium* solution carefully infiltrated. The infiltrated region was outlined with a marker pen and plants were maintained in a growth room at 25ºC for 2 days before tissue harvest or examination.

2.10 Dual luciferase transient expression assay

Dual luciferase assays were carried out to measure the promoter trans-activation and protein localisation in tobacco leaf cells. Two days after agro-infiltration (section 2.10) of leaves 15 ml of PLB (Passive Lysis buffer) was prepared from 5X stock (Promega cat – E1941). 300 µl of PLB was added to each leaf sample for extraction. A 7mm diameter cork borer was used to cut out a leaf disc, which was using a pestle and mortar for 1-2 mins on ice. The ground sample was transferred to a 1.5ml Eppendorf tube on the ice, before centrifugation at ~16000 *g* for 5 minutes at 4ºC to

clear the extracts. For assay of luciferase activity two different reaction buffers were prepared for assay of Firefly and Renilla luciferases (Table 2.2 and 2.3).

Table 2.2. Firefly Luciferase Reaction Buffer

Final concentration	Stock solution concentration	Volume for 10 ml buffer
25 mM Glycylglycine	250 mM	1 ml
15 mM KPO_4 (pH 7.6)	1 M	150 µl
4 mM EGTA	100 mM	400 µl
2 mM ATP	20 mM	1 ml
1 mM DTT	100 mM	100 µl
15 mM $MgSO_4$	1 M	150 µl
0.1 mM CoA	10 mM	100 µl
75 µM Luciferin	0.5 mM	1.5 ml
dH_2O		5.6 Ml

Table 2.3. *Renilla* Luciferase Reaction Buffer

Final concentration	Stock solution concentration	Volume for 10 ml buffer
1.1 M NaCl	5 M	2.2 ml
2.2 mM Na_2EDTA	0.22 M	100 µl
0.22 M KPO_4 (pH 5.1)	1 M	2.2 ml
0.44 mg/ml BSA	100 mg/ml	44 µl
1.43 µM Coelenterazine	1 mM	14.3 µl
dH_2O		5.44 ml

The pH was adjusted to pH 8.0 with 1 M KOH for firefly luciferase reaction buffer and pH 5.0 for Renilla reaction buffer by using 1 M KPO_4. Then the substrates luciferin (at -20ºC) and coelenterazine (at -80ºC) were allowed to thaw on ice. For the measurement of luciferase 96-well assay plates were prepared, labelled and placed on ice to cool. Then 25 µl of extract was added to individual wells in duplicate.

Finally the substrates were added to the reaction buffers and loaded in to the FLUO star omega (BMG - LABTECH). Light emission after injection of substrates was measured and data extracted using Omega data analysis software provided.

2.11 Analysis of genetic transmission in *Arabidopsis*

2.11.1 Forward crosses with *ms1-1* plants

In forward crosses, *ms1-1* plants were used as the female partner and transgenic *Arabidopsis thaliana* plants as the male pollen donor. Flowers were crossed manually by viewing flowers under a Zeiss Stemi SV8 dissecting light microscope and transferring pollen from a fully open flower to the stigma. Approximately, 2-3 male donor flowers were used per *ms1-1* flower. A total 6-7 crosses were performed for each male donor. Following crosses inflorescence branches were kept separate from each other and new branches removed. Plants were maintained until siliques were fully grown, which were then were allowed to dry for approximately 1 week before harvesting into individual labelled eppendorf tubes.

2.11.2 Back crosses of mutant *Arabidopsis thaliana*

In back crosses, non-mutagenised parental line plants and mutant *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ *Arabidopsis thaliana* plants were used as females and EMS_F1 (DAZ1Trunc1-mCherry) lines were used as male donors. Approximately, 4-6 flowers per EMS_F1 lines gave 4-5 good siliques and these back crosses gave rise to F2_BC (BC – Back Crossed) generations. The procedure was similar to that used for forward crosses, but *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ *Arabidopsis thaliana* plants used instead of *ms1-1* plants and unopened flower buds were emasculated by using forceps to remove stamens and crossing carried out 2 days later when the pistil was mature. Donor and recipient plants were isolated to avoid unwanted cross-pollination. Once dried, siliques were harvested and kept in individual eppendorf tubes for further analysis.

2.11.3 Male transmission analysis

For the male transmission analysis *ms1-1* crossed F1 seeds were sterilised and plated on MSO agar plates with 10 µg/ml PPT for selection of transgenic plants. F1 seeds were treated with 70 % ethanol and 0.01 % triton-X and incubated for 5 minutes with shaking at 37˚C. After discarding the 70 % ethanol/triton mixture, 1 ml 100 % ethanol was added before transferring seeds to folded whatman no.1 filter paper, which was then unfolded and allowed to dry. Once, seeds were dry, they were scattered onto the labelled PPT selection plates and sealed with parafilm tape to prevent contamination. Plates were wrapped in aluminium foil and kept for 2 days at 4˚C to stratify, then unwrapped and incubated under the fluorescent light for growth. After 2 weeks, seeding were scored for PPT resistance and sensitivity.

2.12 Dissecting siliques for analysis of ovules developmental stages

To investigate different stages of seed development, mature pistils from open flowers of *ms1-1* flowers were first pollinated with pollen from either wild type Col-0, from *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ or from *daz1-1*$^{-/-}$ *daz2-1*$^{+/+}$ mutant lines expressing DAZ1 variants. Ovule development was checked every 3 days until day 15. Two to three days following pollination siliques were clearly elongating and were measured. Clearing and observation of developing ovules microscopy was performed as described by Park et al. (2004). By using syringe-needles (0.4×12 mm) and forceps siliques were cut open and ovule numbers counted in siliques under a Zeiss STEMI SV8 dissecting microscope. Younger stage ovules were dissected out and transferred to a 1.5 ml eppendorf tube containing 300 µl of clearing solution (240 g chloral hydrate, 30 g glycerol, 90 ml water) for 30 min at room temperature and incubated for 30 minutes at room temperature. Larger ovules were cleared for several hours to overnight. Ovules were transferred to glass slides, a cover slip added and the sample sealed with nail varnish for observation by differential interference contrast (DIC) microscopy on an EclipseTE2000-E (Nikon, Japan) microscope.

2.13 Phenotype analysis of *A. thaliana*

2.13.1 DAPI staining for pollen nuclei

Pollen nuclear morphology was visualized by staining with 4', 6-diamidino-2phenylidole (DAPI) as described (Park *et al.,* 1998). Four to five mature open flowers were collected in a1.5 ml an eppendorf tube that containing 300 µl of DAPI staining solution (0.1 M sodium phosphate pH 7.0, 1 mM EDTA, 0.1 % (v/v) Triton X-100, 0.8 µg/ml DAPI). Pollen grains were released from the flowers into DAPI solution by briefly vortexing. Pollen was pelleted by centrifuging for 30 seconds, 3.5 µl of the pollen pellet was transferred to a microscope slide and allowed to settle for 30 seconds and an 18mm X 18mm cover slip added, which was sealed with nail varnish. Prepared slides were observed under an upright fluorescence microscope (Eclipse 80i, Nikon, Japan) or an inverted fluorescence microscope (EclipseTE2000-E, Nikon, Japan) with oil immersion on 40X and 60X objectives. For high throughput screening of the pollen phenotype, 1-2 mature fully open flowers were collected into the wells of a 96-well microtiter plate containing 100 µl of DAPI solution. Pollen grains was released by gentle tapping of the microtiter plate and sample were observed with an inverted fluorescence microscope (ZEISS Axiovert 100, Germany).

2.13.2 Fluorescent microscopy for marker analysis and image capture

All images of pollen were taken by using different cameras and objectives, which was dependent on the microscope. Using the EclipseTE-2000-E microscope, a mercury lamp was used as an excitation source and all images were taken at 40X and 60 X objectives with oil immersion. The Eclipse 80i used an LED-based excitation source (CoolLED) was used together with 40X and 60X objectives with oil immersion. Fluorescence images were captured with a Hamamatsu ORCA ER camera or DS-QiMc cooled CCD camera using NIS-Elements Basic Research v3.0 software (Nikon, Japan) in JPEG2000 format.

2.13.3 Fluorescent protein quantification in pollen

To measure the levels of fluorescent proteins in pollen images were captured on a DS-QiMc cooled CCD camera at 60X objective using oil immersion. During image capture standard exposure times were used under conditions avoiding over saturation of images for all mutant or transgenic lines. NIS-Elements BR v3.0 software (Nikon, Japan) was used to process the captured images and to analyse the total pixel intensity (TPI) of manually defined regions of interest (ROI) encompassing germ cell and sperm cell nuclei. The region of interest (ROI) was defined by using auto-detect feature within the measurements panels in the software. This ROI was then duplicated and used to measure the TPI of the cytoplasmic background within the same pollen grain. Real fluorescence of the germ cell and sperm cell nucleus was obtained by subtracting the cytoplasmic background TPI from the nuclear TPI. Each batch of images corresponded to a germ cell division mutant or a specific transgenic line and data was exported to clipboard and transferred into an excel file using the software.

2.14 Phylogenetic analysis for *A. thaliana* orthologous

For the phylogenetic analysis of *A. thaliana DAZ1* orthologous several different online domains were used. Plants.ensembl.org and phytozome.jgi.doe.gov were used to download DNA and protein sequences of *DAZ1* orthologs from bryophytes, monocots and eudicot plants. Protein sequences were later used to build phylogenetic trees using MacVector software along with sequence alignments. ClustalW was first used to create protein sequence alignments and advances options allowed these to be used to build phylogenetic trees based on similarity. MEGA 7.0.21 was also used to making alignments and build phylogenetic trees.

2.15 Statistical data analysis

The parametric student's T-test and non-parametric Mann-Whitney U test were performed for data analysis. Pairwise comparison of transient expression assay results was performed on the means of the technical replicates by using two tailed

t-tests assuming equal variance. The two-tailed t-test was also used to evaluate whether average RFP fluorescence intensity measurements of mutant germ cell and sperm cell nuclei were significantly different. A *p* value less than a α of 0.05 was used to denote significance. The evaluation of genetic data was performed by using the chi-square (x^2) function in excel. This test was used to establish whether observed ratios of phenotypes were significantly different from their expected ratios and for the segregation of herbicide resistance and sensitivity in populations of seedlings after crossing. *p* values less than an α of 0.05 was used to denote significance.

Chapter 3

Functional role of DAZ1 protein domains

3.1 Abstract

Background: In DAZ1 and DAZ2, a short basic-rich sequence of amino acids near the N-terminal region, referred to as the basic region (BR), was identified as a putative nuclear localisation sequence (NLS). Sequence alignment of DAZ1 orthologs however revealed that the BR is conserved in dicots, but absent from monocots. A second shared region between AtDAZ1 and AtDAZ2, known as the conserved region (CR), is located between ZnF1 and ZnF2 and is characterised by the sequence motif CLLM. The CR is conserved in eudicot and monocot DAZ1 orthologues. DAZ1-mCherry fusion constructs were used to investigate the ability of modified DAZ1 variants to restore male germline development of *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants. Further constructs were made in which DAZ1 variants were fused to GFP to investigate protein localisation in tobacco leaves and in promoter *trans*-activation assays.

Results: Complementation analysis with DAZ1-BR and DAZ1-CR variants showed partial rescue of generative cell division and fertility in T1 generation *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants. Genetic analysis revealed that DAZ1-BR and DAZ1-CR variants both showed reduced male transmission. In promoter *trans*-activation assays DAZ1-BR and DAZ1-CR variants failed to activate the *GEX1* promoter, indicating that both regions are important for DAZ1 function. Results showed that neither the BR nor the CR were required for the nuclear-enriched localisation of DAZ1 in sperm cells. In contrast, in tobacco leaf cells the DAZ1 BR, but not the CR, was shown to be essential for tight nuclear localisation. The analysis of additional DAZ1 ZnF variants and DAZ1 orthologues isolated from several different species revealed variability in their localisation in sperm cells.

Conclusions: The work presented shows that the DAZ1 BR and DAZ1 CR both make a significant contribution to the role of DAZ1 in generative cell division and to the production of fertile gametes. The DAZ1 BR has a greater role, while the CR has a less critical role in the germline, but both significantly influence promoter *trans*-

activation. In contrast, neither region is important for the translocation of DAZ1 to sperm cell nuclei.

3.2 Introduction

This Chapter describes the functional analysis of the DUO1-ACTIVATED ZINC FINGER 1 (DAZ1) transcription factor and its conserved domains. The functional role of DAZ1 and DAZ2 was established by investigating the pollen phenotype of plants with T-DNA insertions in both genes. This work showed that *daz1 daz2* mutant germ cells fail to undergo division and are unable to complete fertilisation (Borg *et al.,* 2014). Thus, DAZ1 and DAZ2 redundantly control germ cell division and sperm cell differentiation in *Arabidopsis thaliana* (Borg *et al.* 2014).

The DAZ1 and DAZ2 proteins belong to the C_2H_2-type zinc finger (ZnF) protein family, the most abundant family of transcription factors in *A. thaliana* (Englbrecht *et al.* 2004). DAZ1 and DAZ2 belong to the C1-3iC subgroup of proteins which contain three zinc finger domains. DAZ1/DAZ2 genes have a basic amino acid stretch at the N-terminus which has been proposed to act as a nuclear localisation signal (Borg *et al.* 2014). In this chapter, the function of this basic region (BR) and a second conserved region (CR), located between ZnF1 and ZF2 motifs and containing the shared CLLM motif (Figure 3.1), were explored. DAZ1/DAZ2 genes are found throughout flowering plants and complementation of mutant *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants showed that DAZ1 activity is partially conserved between *A. thaliana* and *Solanum lycopersicum* (Sari, 2014). Furthermore, the BR and CR are not always found together in eudicot and monocot DAZ1 orthologues. The BR is absent from DAZ1 orthologues in monocot species, but is conserved in the basal angiosperm *Amborella trichopoda,* however the conserved region (CLLM) is present throughout angiosperms.

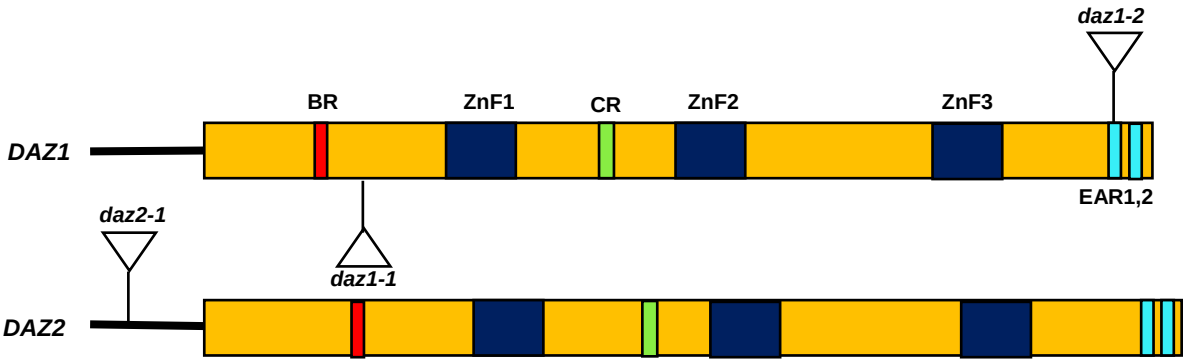


Figure 3.1 DAZ1 and DAZ2 protein domain organisation.

DAZ1 and DAZ2 form subgroup C1-3iC within the C_2H_2-type family that are characterised by three dispersed zinc finger domains and a conserved region around the CLLM motif located between ZnF1 and ZnF2. The positions of the two EAR-motifs and T-DNA insertion alleles are indicated (adapted from Borg et al., 2014).

DAZ1 and DAZ2 contain two motifs at their C-termini which are recognised as ethylene-responsive element binding-factor-associated amphiphilic repression (EAR) motifs (Figure 3.1). EAR-mediated transcriptional repression is emerging as one of the principal mechanisms of plant gene regulation and the EAR motif was the first active repression motif reported in plants. It was originally identified almost a decade ago in a subset of class II ERFs and TFIIIA- type zinc finger proteins as a small motif with a consensus of L/FDLNL/F(x) P (Ohta *et al.* 2001). Recent discoveries of co-repressors interacting with EAR motifs, such as TOPLESS (TPL) and AtSAP18, have begun to unravel the mechanisms of EAR motif-mediated repression (Kagale and Rozwadowski 2011).

The C_2H_2 zinc finger protein family is one of the largest transcription families in *A. thaliana* which contains 176 genes (Englbrecht *et al.* 2004). Previous work has shown that the zinc finger proteins have diverse biological roles from stress responses to the control of floral organ identity. In *Arabidopsis thaliana* DAZ1/DAZ2

genes contain three zinc finger domains and previous work (Rutley 2015) uncovered their functional role in male germline development. ZF1 is a rare K-type zinc finger while ZF2 and ZF3 are plant-specific Q-type zinc fingers (Englbrecht *et al.* 2004).

In recent work, the expression of DAZ1 variants with mutated ZnF1 (H96N; DAZ1muZf1) or ZnF1 and ZnF2 (H96N, H182N; DAZ1muZf1,2) in *daz1 daz2* pollen, identified an essential role for ZnF1 (Rutley 2015). Neither DAZ1muZf1 nor DAZ1muZf1,2 variants complemented *daz1 daz2* pollen phenotypes. However, DAZ1muZf2 and DAZ1muZf3 partially rescued germ cell division of *daz1 daz2* pollen indicating that ZnF2 and ZnF3 contribute to DAZ1 function, but are less important than ZnF1.

Functional analysis of the EAR motifs in DAZ1/DAZ2 further revealed that both EAR motifs are involved in transcriptional repression (Borg *et al.* 2014). Constructs in which EAR1 (mEAR-1; 247 – 251, DLNPP to AAAAA), or EAR2 (mEAR-2; 266 – 270, LRLGL to ARAGA) motifs were mutated showed that both led to reduced *daz1 daz2* transmission efficiency of 63.0 % and 37.5 % respectively (Borg *et al.* 2014). Mutation or truncation of both EAR motifs had a more severe effect, with DAZ1mEAR-1, 2 and DAZ1ΔEAR-1,2 showing only 8.9 % and 21.7 % male transmission respectively. Thus, the EAR motifs have a key role in the function of DAZ1 in generative cell division and sperm fertility, but are not essential for DAZ1 function. This is further supported by the limited ability of the homozygous DAZ1ΔEAR-1,2 constructs to rescue germ cell division, as approximately 15 % tricellular pollen are produced in a homozygous *daz1-1 daz2-1* mutant background (Borg *et al.* 2014, Rutley 2015).

The functional roles of DAZ1 zinc finger domains and EAR motifs uncovered by previous research provided insight into their relative importance for DAZ1 function in male gametophyte development (Borg *et al.* 2014, Rutley 2015). The main objective of the work presented was to extend these studies to investigate the

significance of the DAZ1 BR and CR. Figure 3.2 and Figure 3.3 show sequence alignments of the BR and CR in DAZ1 and DAZ2.

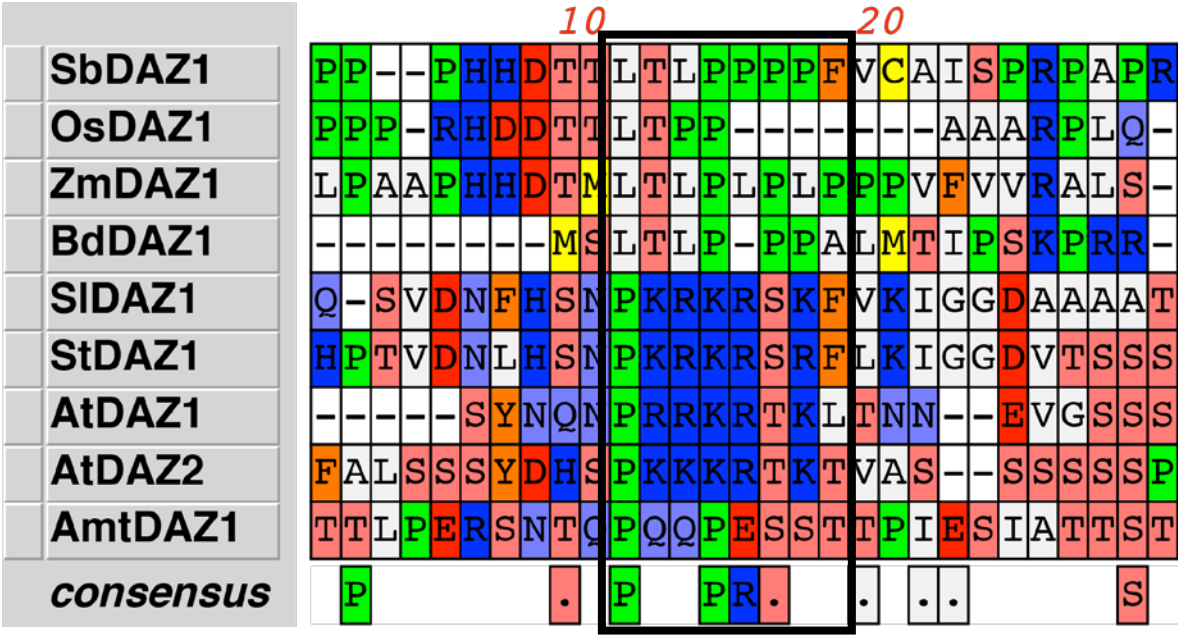


Figure 3.2. DAZ1 BR conservation in eudicot and monocot species. ClustalW alignment was carried out using DAZ1 sequences from selected species. Abbreviations: SbDAZ1 – *Sorghum bicolor,* OsDAZ1 – *Oryza sativa,* ZmDAZ1A – *Zea mays,* BdDAZ1A – *Brachypodium distachyon,* StDAZ1A – *Solanum tuberosum,* SlDAZ1A – *Solanum lycopersicum,* AtDAZ1 – *Arabidopsis thaliana,* AmtDAZ1A – *Amborella trichopoda*.

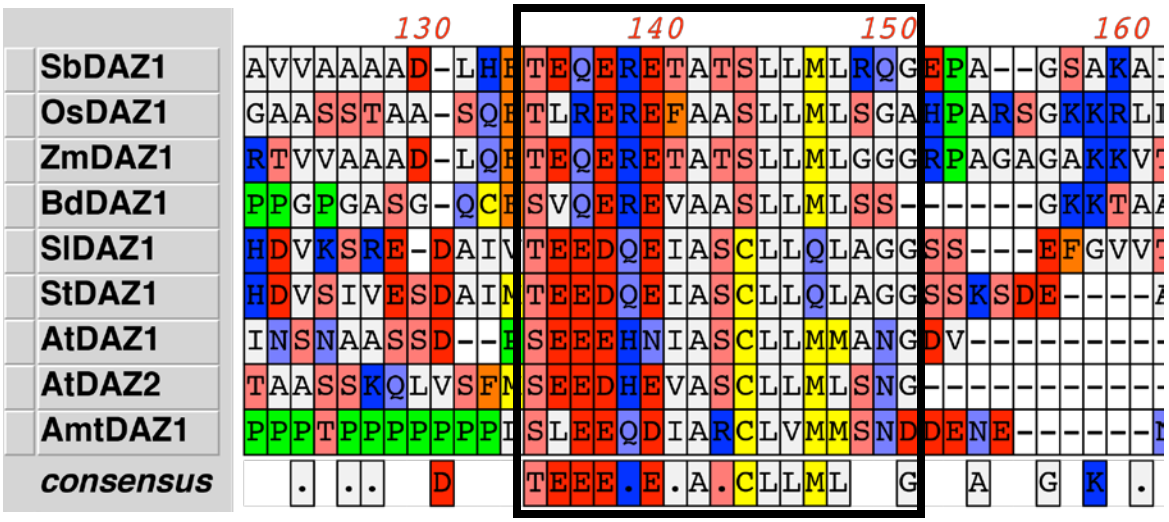


Figure 3.3. DAZ1 CR (CLLM/ SLLM motif) in eudicot and monocot species. ClustalW alignment was carried out using DAZ1 sequences from selected species. Abbreviations: SbDAZ1 – *Sorghum bicolor,* OsDAZ1 – *Oryza sativa,* ZmDAZ1A – *Zea mays,* BdDAZ1A – *Brachypodium distachyon,* StDAZ1A – *Solanum tuberosum,* SlDAZ1A – *Solanum lycopersicum,* AtDAZ1 – *Arabidopsis thaliana,* AmtDAZ1A – *Amborella trichopoda*.

3.2.1 Nuclear localisation signals

Nuclear localisation signals are classified as either monopartite or bipartite (Kalderon *et al.* 1984). Dingwall and Laskey (1982) first discovered NLS regions, and later the monopartite NLS sequence PKKKRKV was identified in the SV40 Large T-antigen. Mating type α2-like NLSs consist of short hydrophobic regions that contain one or more basic amino acids (KlPlK), (Hall *et al.* 1984). Bipartite NLSs are generally a combination of the two regions of amino acids, which are separated by almost 10 amino acids (Dingwall and Laskey 1991). NLSs are short sequences and contain one or two clusters of the positively charged basic residue of lysine and arginine. There is some variation in the sequences of NLSs and monopartite NLSs or bipartite NLSs have been identified in different animal and plant species (Dingwall and Laskey 1991, Garcia-Bustos *et al.* 1991). Mutation in positive charges is often

the simplest way to disrupt NLSs and nuclear import (Bonifaci *et al.* 1997). Further, deletion of an NLS motif interrupts nuclear import and, normally, a non-nuclear protein will be imported into the nucleus if fused to an NLS. Both facts have been used consistently to unravel NLS motif function (Tinland *et al.* 1992, Moede *et al.* 1999).

3.2.2 Mechanisms of nuclear protein import

The importin α/β complexes mediate the nuclear transportation or import of proteins with a classical NLS through nuclear pore complexes. In the cytoplasm, cargo proteins bind to the importin-β, via the importin-α adaptor. Binding to the importin complex facilitates passage through NPCs to the nucleus, where the import complex is dissociated by Ran-GTP, a Ras-family GTPase. (Weis 2003, Conti *et al.* 2006, Stewart 2006). Figure 3.4 shows a schematic of the nuclear protein transportation pathway.

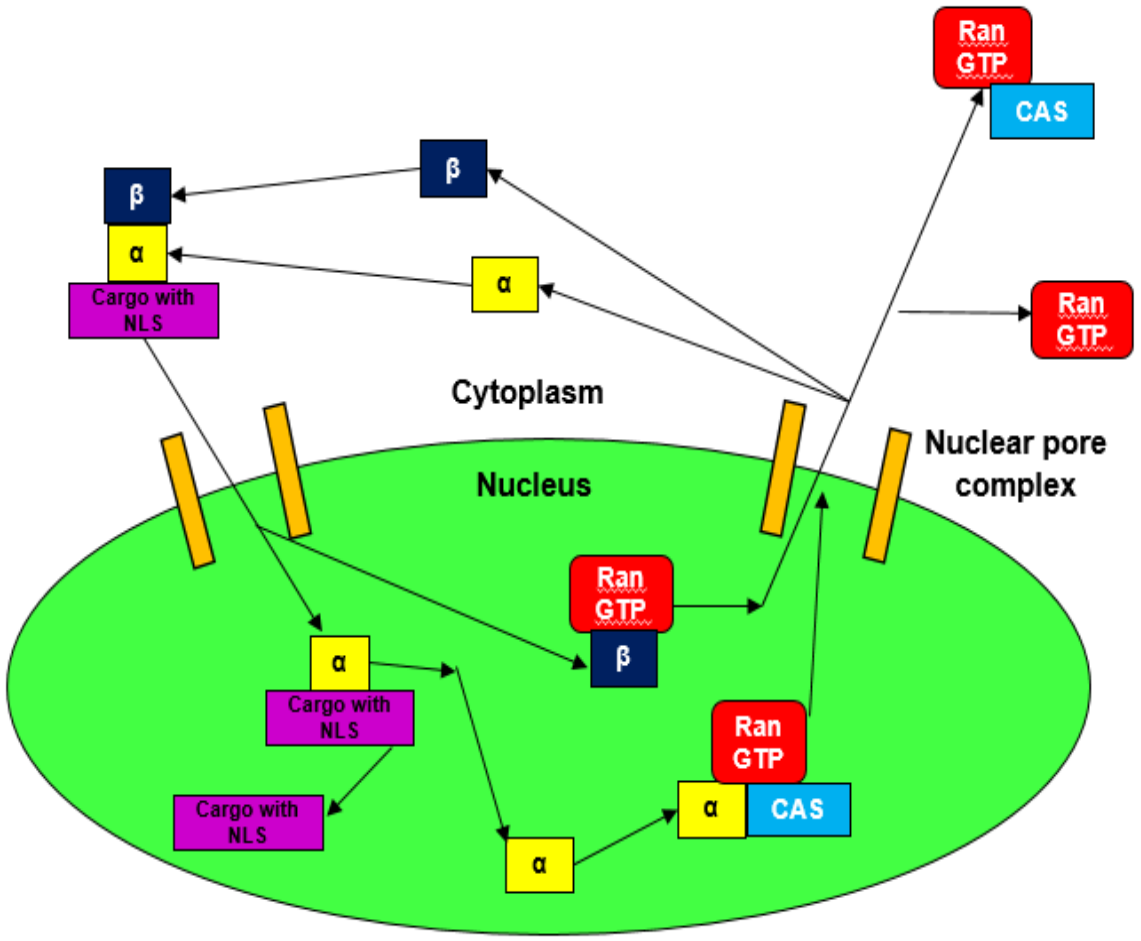


Figure 3.4. Nuclear protein transport pathways. The diagram illustrates the pathway of nuclear protein transport via nuclear pore complex with the association of importin-α / β and RanGTP to maintain import cycles. CAS (Cselp in yeast) and importin-α bind with RanGTP to exit cytoplasm and release importin-α.

The importins are recycled to the cytoplasm with importin-β complexed with Ran-GTP and importin-α complexed with both Ran-GTP and its nuclear export factor, CAS (Cselp in yeast), which is a member of the importin-β superfamily of karyopherins. Finally, cytoplasmic RanGAP (Ran-GAPase-activating protein; Rnalp in yeast) activates Ran-GAP hydrolysis, releasing the importin for further protein import cycles (Stewart 2006).

3.2.3 Evolution of the DAZ1 BR and CR

DAZ1-related genes were identified and recovered using the online portal phytozome v12.0 and their sequences aligned to identify conserved and variable domains (Figure 3.5, 3.6). The tomato, rice, sorghum, maize and *Brachypodium distachyon* genomes encode a single DAZ1 protein and show conservation of ZnF1 and two C-terminal EAR motifs. A clear difference in the domain structure of DAZ1 proteins in flowering plants is that DAZ1 proteins from Poaceae species lack ZnF2, which is conserved in DAZ1 proteins from eudicots. Monocot species also lack the BR at the N-terminus but possess the CR (CLLM/SLLM) between ZnF1 and ZnF2. Further, *Solanum lycopersicum* is unusual among eudicots because the single DAZ1 orthologue, SlDAZ1, has only two zinc finger domains and is missing ZnF3. Protein sequence alignment of a selection of DAZ1 homologues are shown in Figure 3.6.

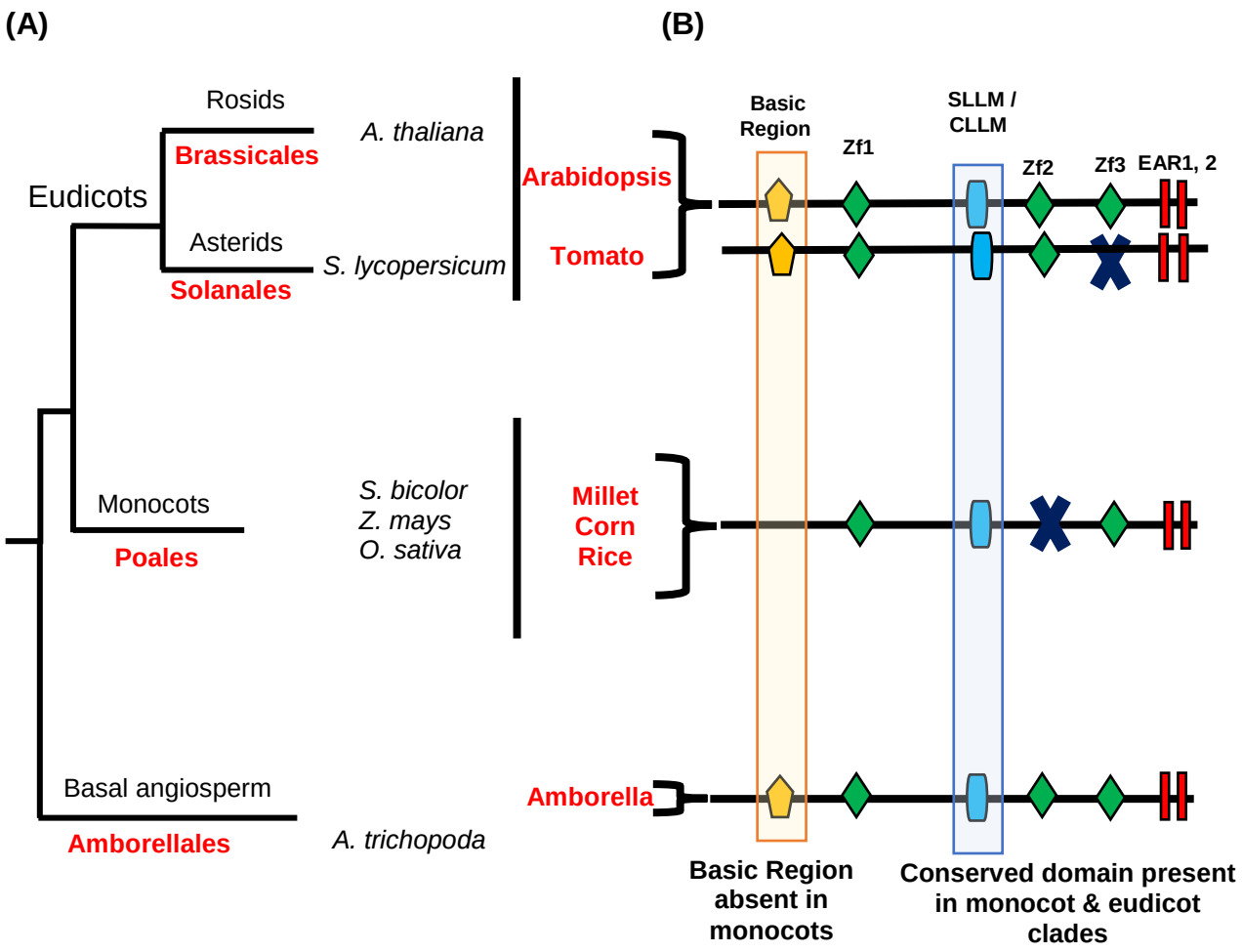


Figure 3.5. Evolution and conservation of DAZ1 in angiosperms. (A) Cladogram illustrating the evolutionary relationship between land plants. (B) Conservation of domains and motifs in DAZ1 orthologues in flowering plants.

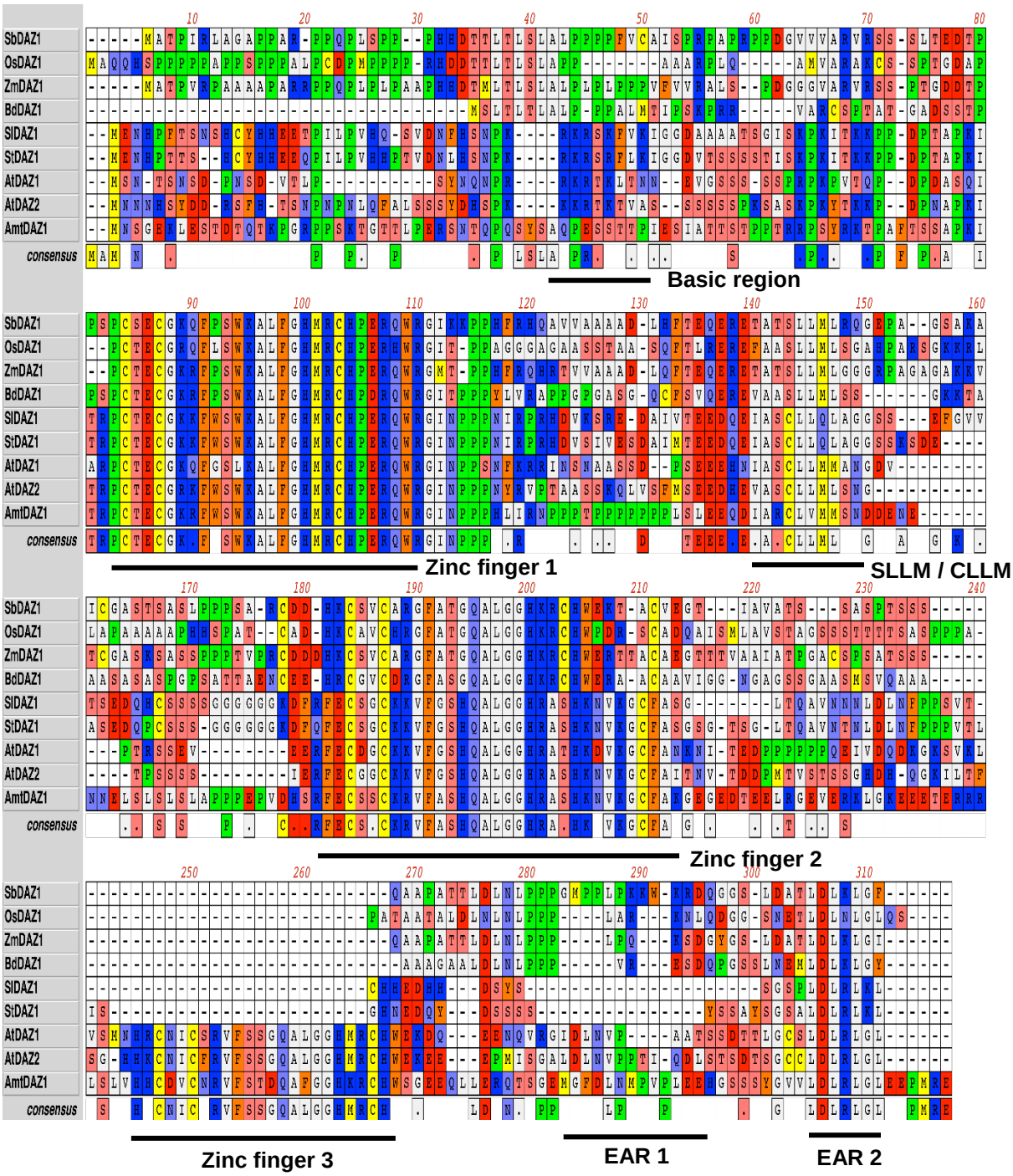


Figure 3.6. Alignment of DAZ1 from selected angiosperms. The amino acid sequences of selected DAZ1 orthologues were aligned using ClustalW.

Monocot species lack the BR at the N-terminus but possess the CR (CLLM/SLLM) between ZnF1 and ZnF2. Dicot species show conservation of BR and CR. Abbreviations: Sb – *Sorghum bicolor,* Os – *Oryza sativa,* Zm – *Zea mays,* Bd –

Brachypodium distachyon, Sl – *Solanum lycopersicum,* St – *Solanum tuberosum,* At – *Arabidopsis thaliana,* Amt – *Amborella trichopoda.*

3.2.4 Overview of the DAZ1 CR

In addition to the ZnF and EAR motifs in DAZ1, the CR represents a signature of DAZ1 proteins in dicot and monocot species (Figure 3.5, B; Figure 3.6). Given that non-functional sequences are expected to diverge faster than functional sequences (Haudry *et al.* 2013), the CR is proposed to play an important role in DAZ1 function. As the significance of this region has not been shown, this presented the opportunity to investigate its functional role in male gamete development. Protein alignment revealed that the DAZ1 CR is conserved in dicot species like *Arabidopsis thaliana,* and *Solanum lycopersicum* and present with some amino acid differences in monocot species such as *Sorghum bicolor, Oryza sativa* and *Zea mays* which possess the SLLM rather than the CLLM regions (Figure 3.6).

3.2.5 Strategy for functional analysis of the DAZ1 BR and CR

The requirement for the DAZ1 BR was studied by making two different constructs encoding mutated versions of DAZ1 fused with the fluorescent protein mCherry under the control of the DAZ1 promoter. First, a mutation was introduced to substitute the basic residues arginine (R) to asparagine (N) and lysine (K) to asparagine (N) and second, the entire basic region was removed by deletion (Figure 3.7).

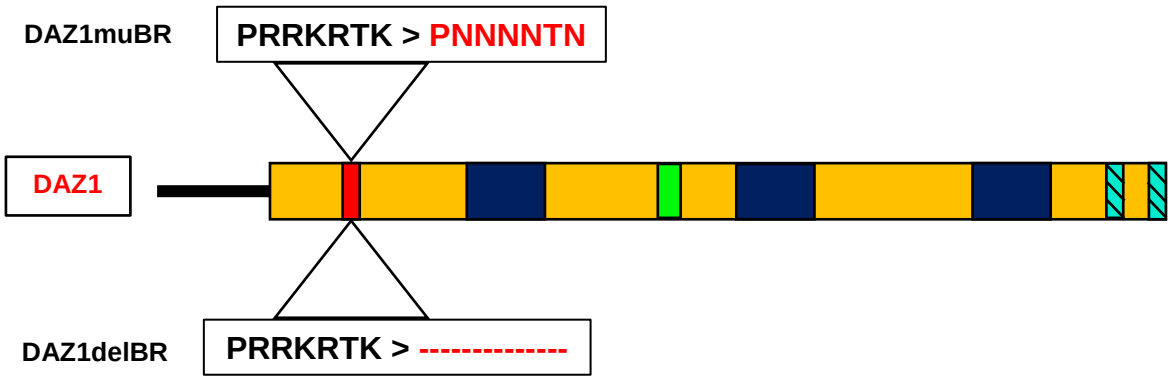


Figure 3.7. DAZ1 BR mutagenesis. Diagram illustrates substitution mutations and deletion of the BR. Mutant BR (muBR) changes basic residues (R>N and K>N) and deletion BR (ΔBR) completely removes the basic region.

In planta complementation was used to test the functional role of the basic region by introduction of DAZ1muBR and DAZ1delBR into *daz1-1*$^{-/-}$*; daz2-1*$^{+/-}$ plants by *Agrobacterium tumefaciens* based transformation (Clough and Bent 1998). In T1 plants it is expected that full complementation or rescue would lead to an increase in the percentage of tricellular pollen up to 75 % if the hemizygous transgene can fully rescue the germ cell division defect of *daz1-1 daz2-1* mutant germ cells (Table 3.1). If the BR is required to support the role of DAZ1 in germ cell division, DAZ1 BR variants will not rescue division and *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants hemizygous for the transgene are predicted to show 50 % TCP (tricellular pollen) and 50 % BCP (bicellular pollen).

(A) Full length DAZ1-mCherry variant.

(B) DAZ1 mutant and Δ BR-mCherry variants.

Figure 3.8. Schematic of DAZ1 BR variants. Substituted and deleted DAZ1 BR constructs fused to mCherry fusion under control of the DAZ1 promoter.

Table – 3.1. Punnet square showing predicted genotypes of progeny from selfed *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants heterozygous for a single locus DAZ1-mCherry transgene ($T^{+/-}$)

♀ \ ♂	**D1-; D2+; T+ TCP**	**D1-; D2+; T- TCP**	**D1-; D2-; T+ TCP**	**D1-; D2-; T- BCP**
D1-; D2+; T+	D1-; D2+; T+ D1-; D2+; T+	D1-; D2+; T+ D1-; D2+; T-	D1-; D2-; T+ D1-; D2+; T+	No rescue
	100 % TCP	**100 % TCP**	**100 % TCP**	
D1-; D2+; T-	D1-; D2+; T+ D1-; D2+; T-	D1-; D2+; T- D1-; D2+; T-	D1-; D2-; T+ D1-; D2+; T-	No rescue
	100 % TCP	**100 % TCP**	**75 % TCP**	
D1-; D2-; T+	D1-; D2+; T+ D1-; D2-; T+	D1-; D2+; T- D1-; D2-; T+	D1-; D2-; T+ D1-; D2-; T+	No rescue
	100 % TCP	**50 % TCP**	**100 % TCP**	
D1-; D2-; T-	D1-; D2+; T+ D1-; D2-; T-	D1-; D2+; T- D1-; D2-; T-	D1-; D2-; T+ D1-; D2-; T-	No rescue
	75 % TCP	**50 % TCP**	**50 % TCP**	

3.2.6 Cloning and generation of DAZ1 CR mutant transgenic lines

DAZ1-CLLM-mCherry variants were made by using fusion PCR to generate entry clones with mutant DAZ1 sequences. Gateway cloning (Chapter 2.7.1) was used in three-part recombination reactions, where mutant DAZ1 cDNA was fused with the mCherry driven by the DAZ1 promoter. Entry clones were confirmed by sequencing and constructs confirmed by restriction endonuclease digestion of isolated plasmids. The construct map of DAZ1LLAA variants with mCherry fusion protein are shown in Figures 3.9 and 3.10.

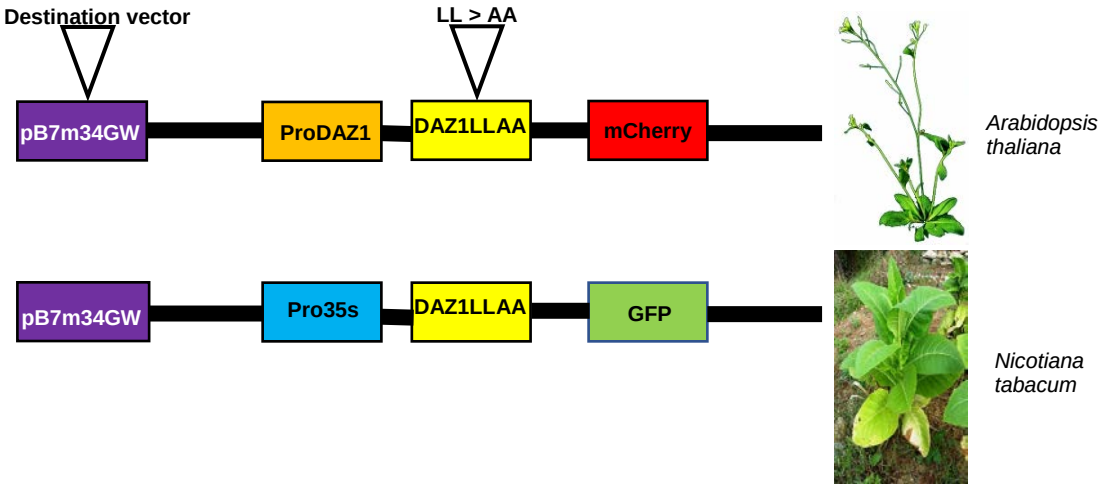


Figure 3.9. Structure of proDAZ1:DAZ1LLAA-CLLM-mCherry and pro35S:DAZ1LLAA-CLLM-GFP constructs. The *DAZ1* promoter construct was used for transformation of *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants to generate stable transgenic lines and the *CaMV 35S* promoter construct was used for transient expression assays in tobacco leaves. Amino acid substitution leucine to alanine LL > AA.

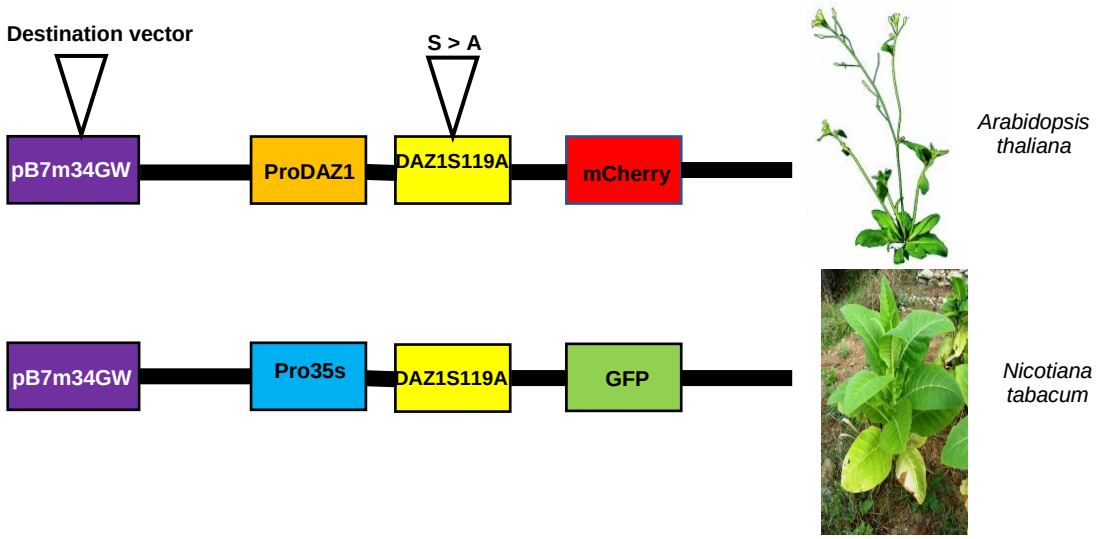


Figure 3.10. Structure of proDAZ1:DAZ1S119A-CLLM-mCherry and pro35S:DAZ1S119A-CLLM-GFP constructs. The *DAZ1* promoter construct was used for transformation of *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants to generate stable transgenic lines and the *CaMV 35S* promoter construct was used for transient expression assays in tobacco leaves. Amino acid substitution serine to alanine, S > A (S119A).

3.2.7 Sequence alignment of DAZ1 BR variants

The plasmid DNA of *DAZ1-muBR* and *DAZ1-ΔBR* entry clones were sequenced. Sequence data were compared with the *AtDAZ1* reference sequence to confirm that the introduced mutations were correctly incorporated and there were no other changes. Figure 3.11 shows sequence alignment of mutant DAZ1-BR variants.

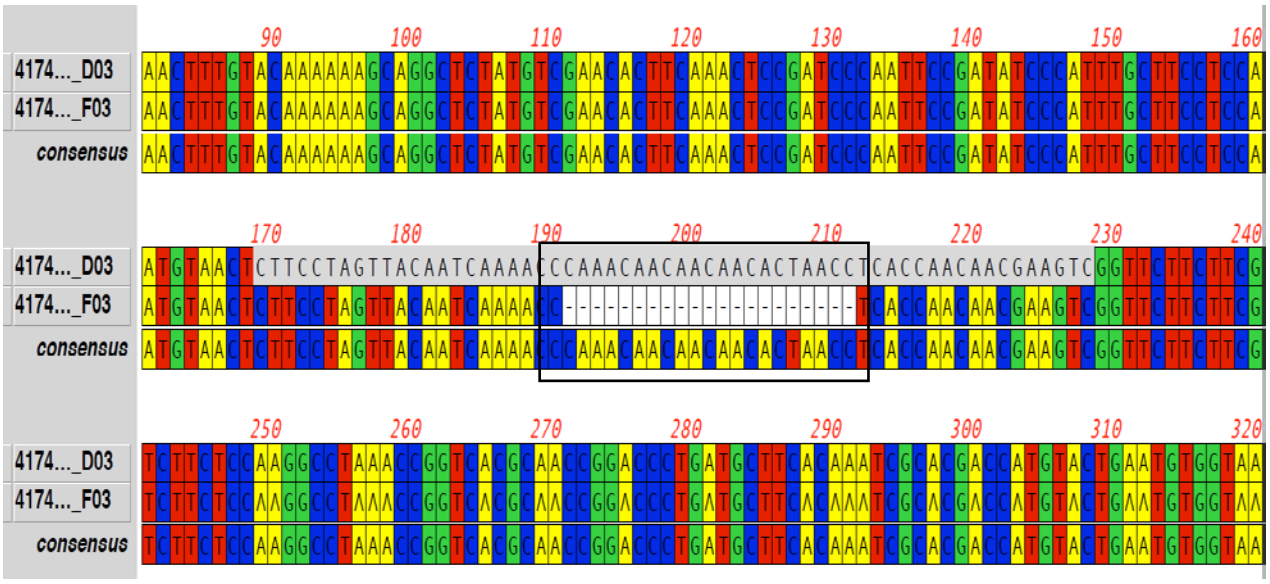
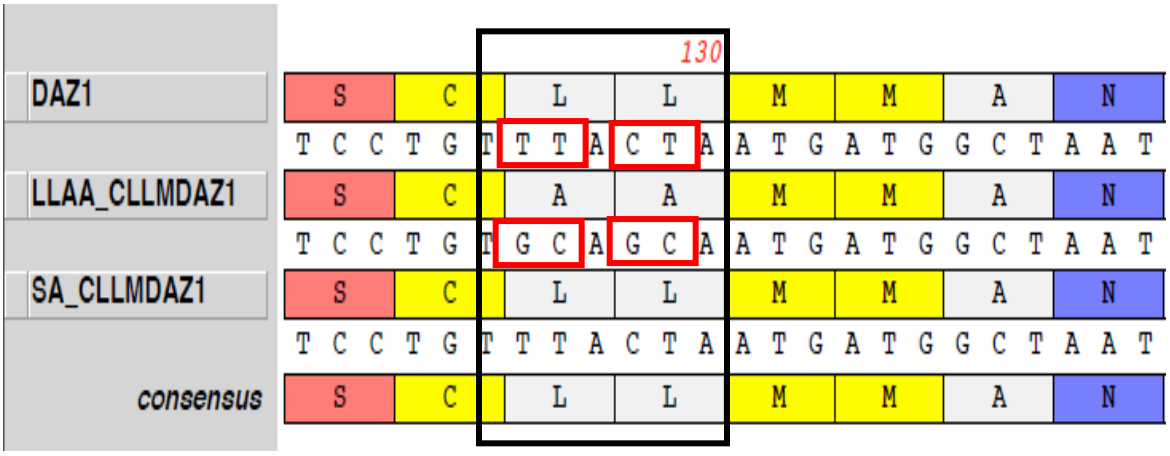


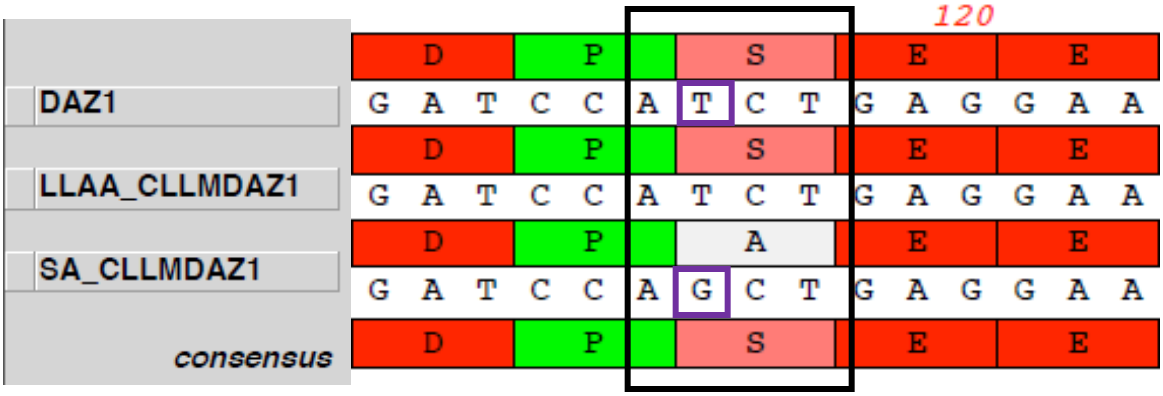
Figure 3.11. Sequence alignments *DAZ1 BR* variants. The figure verifies the 21 bp deletion in the BR region marked in grey.

3.2.8 Sequence alignment of DAZ1 CR variants

Sequencing results for the DAZ1-CR variants aligned with the *DAZ1* are shown in (Figure 3.12 A, B). Sequencing results confirmed the presence of the intended changes and that there were no additional errors in the *DAZ1* coding sequence.



(A) *DAZ1LLAA-CLLM*: leucines (LL) changed to alanines (AA)



(B) *DAZ1S119A*: Single amino acid change, serine (S) to alanine (A).

Figure 3.12. Sequence verification *DAZ1-CR* variants. Aligned sequences shown in (A) and (B) confirming the mutations LL > AA and S > A, the latter of which is a potential phosphorylation site in CR.

3.3 Results

3.3.1. DAZ1 BR is required for promoter *trans*-activation

In planta complementation assays of *daz1 daz2* double mutant pollen can provide insights into the biological consequences of disruption of the DAZ1 BR, but does not provide information on how these constructs might affect the regulation of individual genes within the DAZ1 gene network. Therefore, the promoter of *GEX1* (Engel *et al.* 2005) was used as a known target of DAZ1 (Rutley 2015), to assess the impact of DAZ1 BR modifications on promoter *trans*-activation. To test the role of the DAZ1 BR, transient expression assays were used to measure *trans*-activation of the *GEX1* promoter with the DAZ1-BR variants as effectors. *GEX1* promoter driven firefly luciferase (FLuc) activity was normalised to that of CaMV 35S driven renilla luciferase (RLuc) to control for variation in transformation efficiency (Sherf *et al.* 1996, Borg *et al.* 2011).

To perform DAZ1 *trans*-activation assays, effector constructs were made in binary T-DNA vectors under the control of the CaMV 35S promoter. These included the

destination vector plasmids pPro35S:DAZ1 as the positive control, pPro35S:DAZ1muBR and pPro35S:DAZ1ΔBR. The reporter vector was pProGEX1:LUC (Figure 3.14). *Agrobacterium* strains with the reporter alone and with different effector plasmids were infiltrated into tobacco leaves and examined after 48 hours for dual luciferase activity (Figure 3.13).

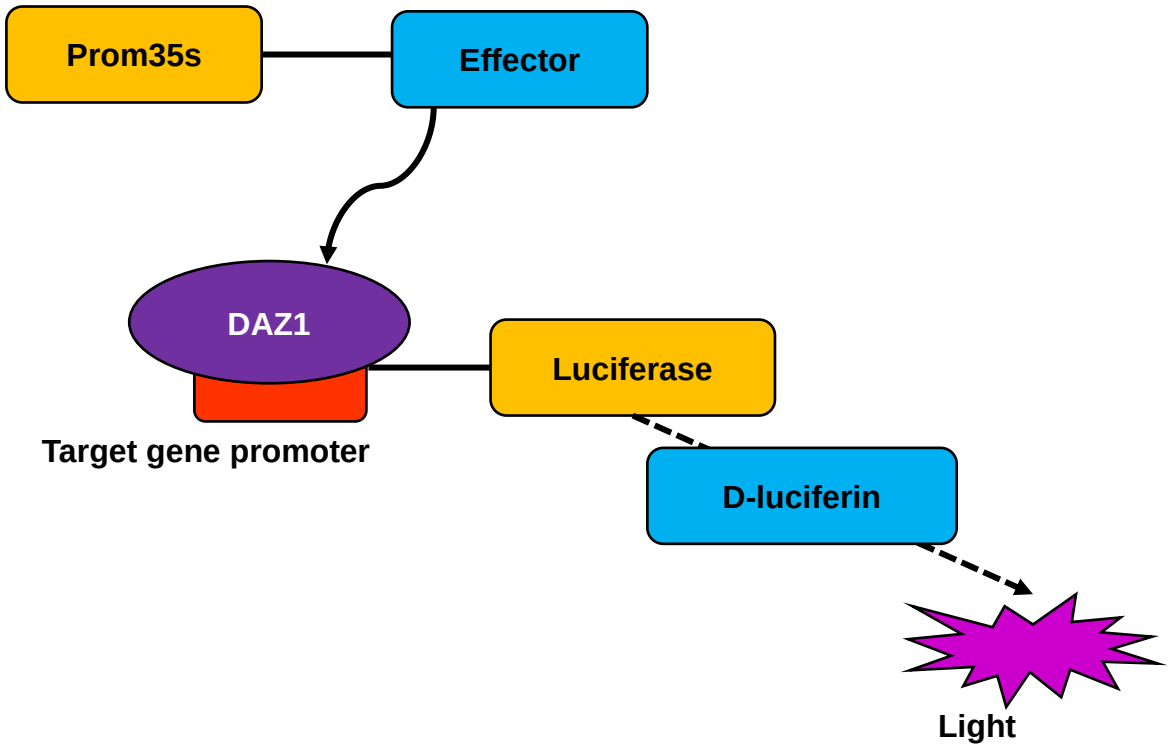


Figure 3.13. Schematic of luciferase *trans*-activation assay. DAZ1 effector constructs are expressed under control of the CaMV 35S promoter. DAZ1 should bind to a GEX1 to promote transcription and expression of luciferase. Luciferase will oxidize D-luciferin in the presence of ATP, O2 and Mg^{2+} to generate a quantifiable release of bioluminescence.

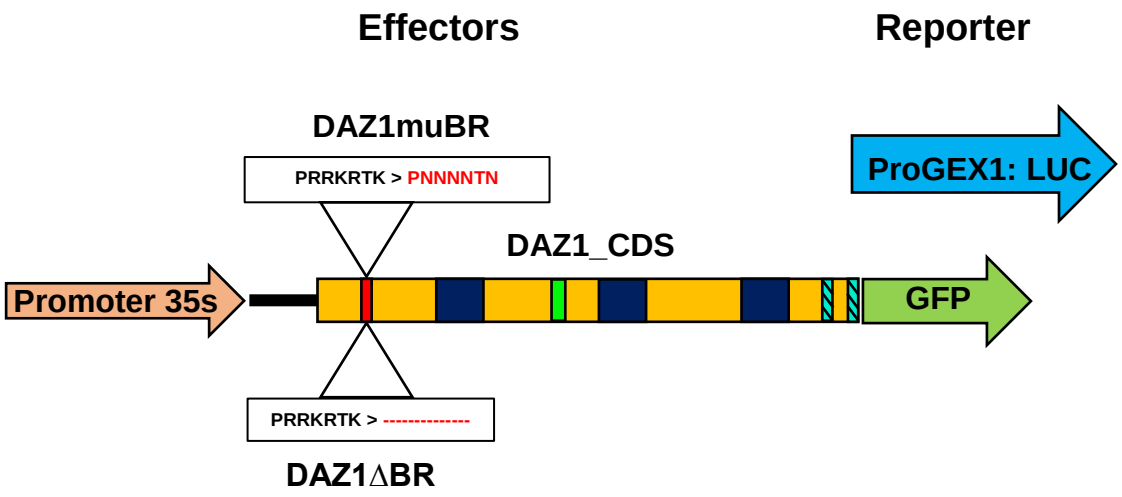


Figure 3.14. Diagram of the BR constructs used in *trans*-activation assays. Effector construct map shows the CaMV 35S promoter fused to DAZ1 variants DAZ1muBR and DAZ1ΔBR. ProGEX1:LUC was used as the reporter.

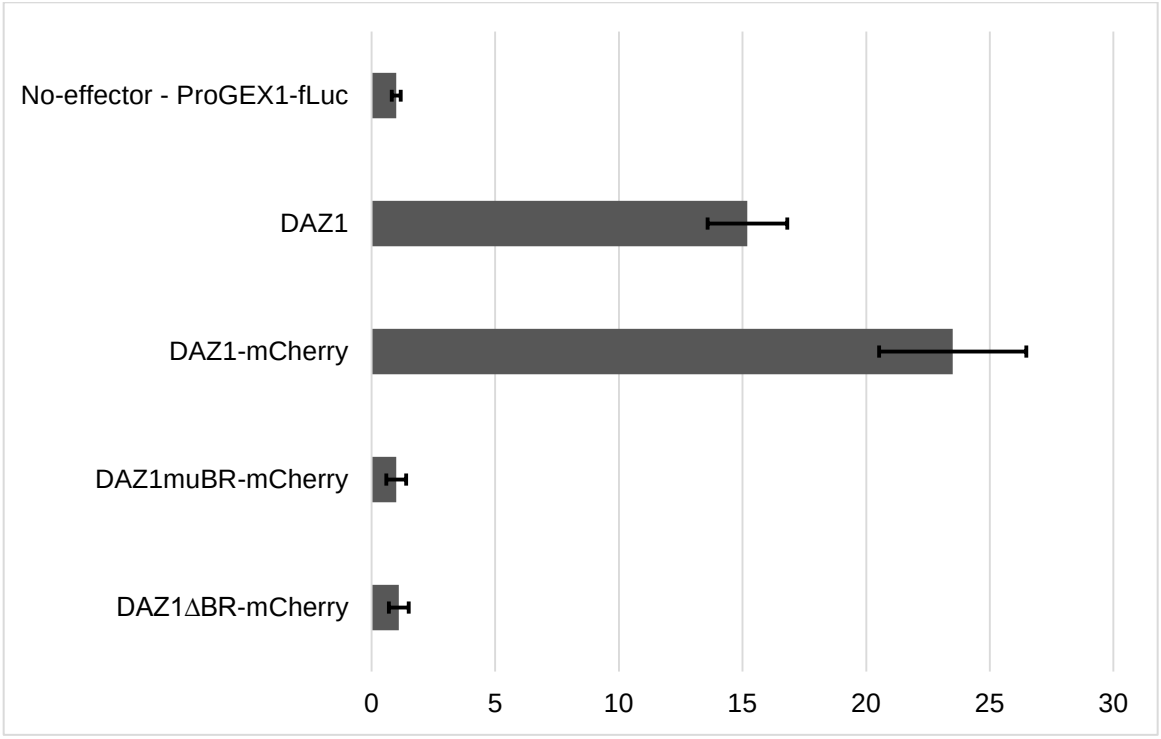


Figure 3.15. *Trans*-activation assays of DAZ1 BR variants. Mean relative fold-change in luciferase activity is shown reflecting *trans*-activation of the *GEX1*

promoter. No effector ProGEX1-fLUC, DAZ1 and DAZ1muBR/DAZ1delBR-mCherry effectors were used as controls. Error bars indicate SE (n = 4 replicate assays).

Trans-activation assay data show that infiltration with DAZ1 BR variants resulted in significantly reduced activity from the GEX1 promoter compared with unmodified DAZ1 (Figure 3.15). For DAZ1muBR and DAZ1ΔBR promoter transactivation was reduced approximately 20-fold compared to unmodified DAZ1, suggesting that the DAZ1 BR is required for the ability of DAZ1 to function as a *trans*-activator of the germline-specific GEX1 promoter.

3.3.2 *Trans*-activation assays of DAZ1 CR variants

Trans-activation assays with the GEX1 promoter were also used to test the activity of Pro35S:DAZ1S119A and Pro35S:DAZ1LLAA-CLLM constructs (Figure 3.17). *Agrobacterium* strains with the reporter alone and with Pro35S:DAZ1S119A and Pro35S:DAZ1LLAA-CLLM constructs were infiltrated into tobacco leaves along with the controls, DAZ1, DAZ1muBR and DAZ1ΔBR, which were fused with GFP. Assay of leaves infiltrated with Agrobacterium strains carrying modified DAZ1S119A and DAZ1LLAA-CLLM constructs showed an approximately 6-fold reduction in activity compared with the full-length DAZ1-GFP control and DAZ1stop version as negative control (Figure 3.16). The activity of DAZ1-CR variants was low and comparable with that observed for the DAZ1muBR and DAZ1ΔBR constructs and for the no effector control (Figure 3.13). This suggests that the DAZ1 conserved region is required for the ability of DAZ1 to function as a *trans*-activator of the germline-specific *GEX1* promoter.

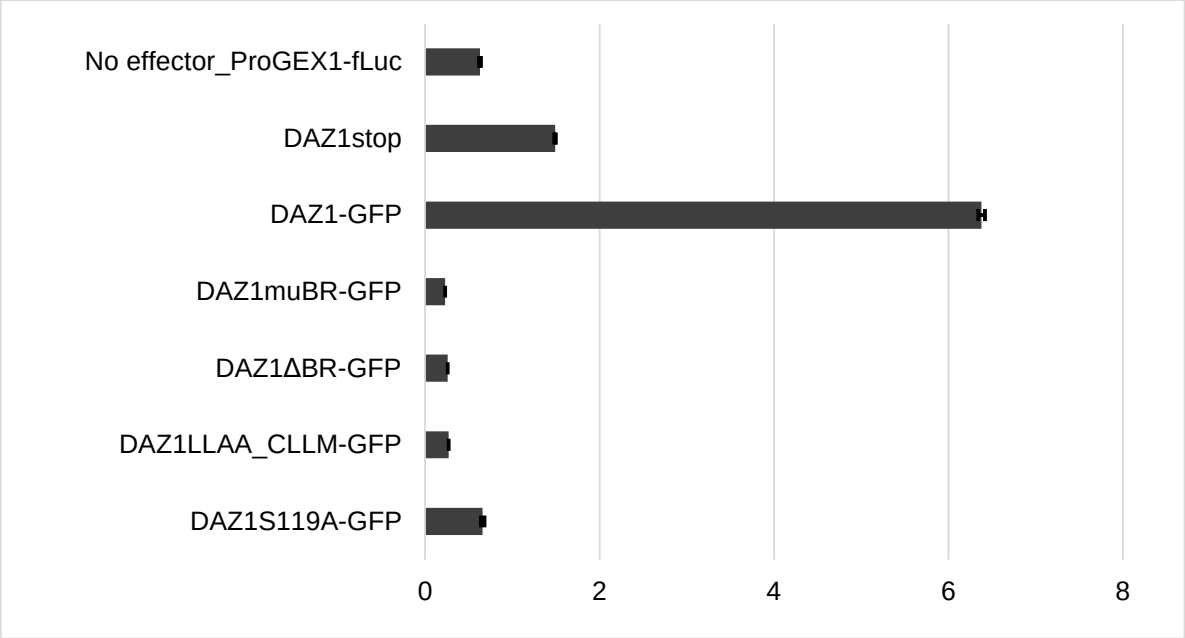


Figure 3.16. *Trans*-activation assay of DAZ1-CR variants. **Mean** relative fold-change in luciferase activity is shown reflecting *trans*-activation of the *GEX1* promoter. No effector ProGEX1-LUC, DAZ1stop and DAZ1muBR/DAZ1ΔBR-GFP effectors were used as controls.

3.3.3 DAZ1 protein localisation in tobacco leaf cells

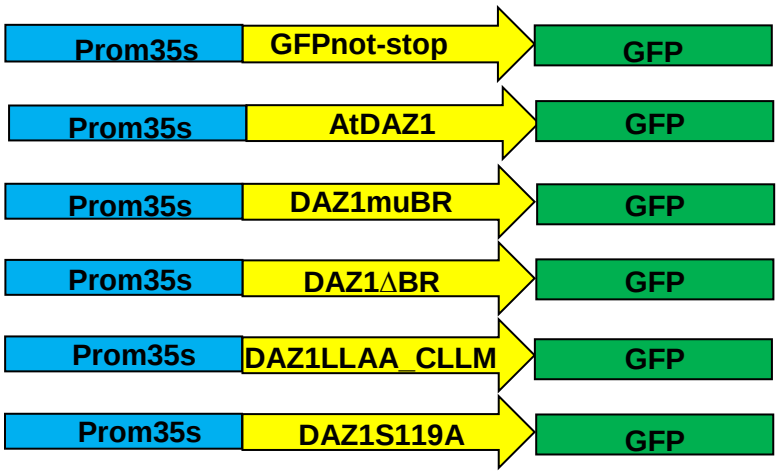


Figure 3.17. Structure of DAZ1-BR and DAZ1-CR constructs. Negative control GFPnot-stop, mutated *DAZ1* BR (muBR and ΔBR) and CR variants (LLAA_CLLM and S119A) were fused to GFP under control of the CaMV 35S promoter.

Strains carrying the constructs shown in Figure 3.17 were infiltrated into *Nicotiana benthamiana* leaves and observed after two days by fluorescence microscopy. It is predicted that if the DAZ1 BR is important for nuclear localisation then DAZ1-BR variants will delocalise. The full-length AtDAZ1-GFP shows tight nuclear localisation (Figure 3.18B). DAZ1muBR and DAZ1delBR-GFP show delocalisation compared to the full-length AtDAZ1-GFP (Figure 3.18B, E, and F). The 2XGFP (GFPnot-stop-GFP) negative control has a largely cytoplasmic GFP signal (Figure 3.18A). DAZ1LLAA-CLLM and DAZ1S119A-GFP variants were not predicted to affect localisation and show tight nuclear localisation similar to unmodified DAZ1-GFP (Figure 3.18C, D).

These results suggest that the DAZ1 CR CLLM motif is not important or required for the nuclear localisation of DAZ1 in leaf cells (Figure 3.18). Both DAZ1 BR variants showed delocalisation and cytoplasmic GFP signals in leaf cells, which showed that the N-terminal basic region of DAZ1 is required for the tight localisation to the nucleus.

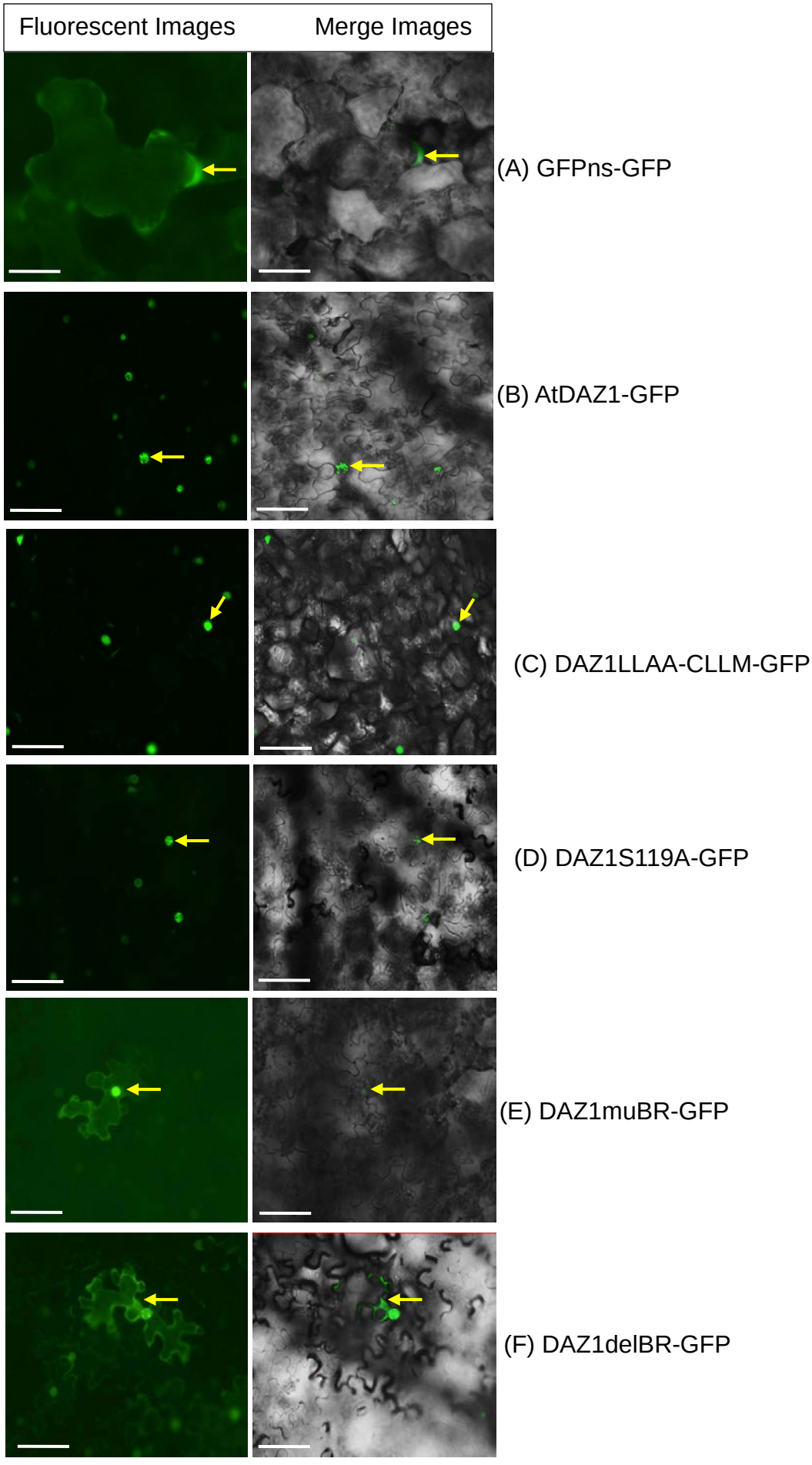

Fluorescent Images
Merge Images
(A) GFPns-GFP
(B) AtDAZ1-GFP
(C) DAZ1LLAA-CLLM-GFP
(D) DAZ1S119A-GFP
(E) DAZ1muBR-GFP
(F) DAZ1delBR-GFP

Figure 3.18. DAZ1-GFP protein localisation in leaf cells.

Localisation of DAZ1 and DAZ1 variants in *Nicotiana benthamiana* leaves. (A) GFPns-GFP (2XGFP) is a negative control and mostly cytoplasmic, (B) full-length AtDAZ1-GFP localises to the nucleus, (C), (D) DAZ1LLAA-CLLM and DAZ1S119A localise to the nucleus, (E), (F) DAZ1muBR/ΔBR proteins are present in both the cytoplasm and nucleus. Yellow arrows indicate GFP signal. Scale bar, 10µm.

3.3.4 Functional role of the DAZ1 BR in Arabidopsis germ cells

The function of DAZ1-BR variants was analysed *in planta* by expressing these under the control of the *DAZ1* promoter after their introduction into the *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ mutant background. Half of the pollen populations in a *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ mutant plant is tricellular and has the genotype *daz1-1*$^{-}$ *daz2-1*$^{+}$, while the other 50 % is bicellular and has the genotype *daz1-1*$^{-}$ *daz2-1*$^{-}$. These phenotypes can be distinguished by DAPI-staining (Borg *et al.* 2014). Rescue of generative cell division in double mutant *daz1-1 daz2-1* pollen will produce tricellular pollen.

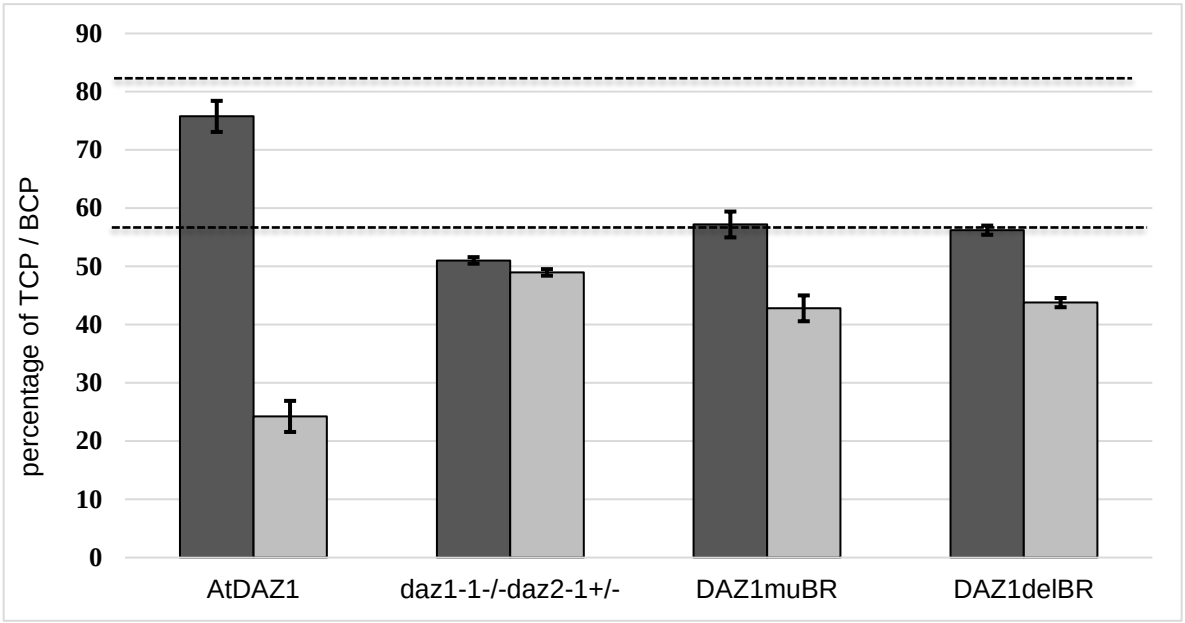


Figure 3.19. Ability of DAZ1-BR variants to rescue germ cell division of *daz1-1 daz2-1* pollen. DAZ1-BR variants show partial rescue of germ cell division in different single locus insertion T1 lines compared with full-length DAZ1-mCherry.

Whereas 75 % TCP represents full rescue, 50 % represents no rescue in mutant *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants. Dark bars show percentage TCP and light bars, BCP. Error bars indicate SE (n=5 for each construct).

A large number of T1 lines were screened for representative *daz1-1 daz2-1* mutant lines with single locus insertions for the transgene showing ~50 % DAZ1-mCherry signal by fluorescence microscopy. Rescue of cell division was scored by counting TCP and BCP after DAPI-staining. Plants showing single locus insertions and bright DAZ1-mCherry signals were further examined in detail. While full-length DAZ1-mCherry showed 75 % TCP, the negative controls with no transgene showed 50 % TCP. Lines expressing DAZ1muBR-mCherry (line B1 and D5) showed 57 % and 62 % TCP (Figure 3.19). This is derived from 50 % TCP *daz1*$^{-}$ *daz2-1*$^{+}$ pollen and partial rescue (3% and 12% TCP) from the *daz1*$^{-}$ *daz2-1*$^{-}$ double mutant class, representing a rescue efficiency of 12 % and 48 % respectively. DAZ1ΔBR (lines A2 and D4) lines showed similar results with 56 % and 60 % TCP, representing 24 % and 40 % division rescue (Figure 3.19). This represents partial rescue compared with full-length DAZ1-mCherry. These data strongly suggest that the DAZ1 BR is important for the role of DAZ1 in germ cell division.

3.3.5 Functional role of the DAZ1 CR *in planta*

DAPI-staining was used to identify *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ mutants among T1 lines expressing DAZ1LLAA-CLLM-mCherry or DAZ1S119A-Cherry and single locus lines expressing ~ 50 % mCherry with bright signals were selected for detailed analysis. DAZ1LLAA-CLLM lines showed partial rescue of germ cell division compared with full-length DAZ1 and *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants. Out of 48 T1 plants three plants were identified with a single locus insertion and bright mCherry expression. These lines showed 51.9 - 54.5 % TCP for DAZ1LLAA-CLLM, indicating that the DAZ1LLAA-CLLM variant showed limited ability to germ cell rescue division in *planta* (Figure 3.20). In contrast, the DAZ1S119A variant gave a total four single locus lines with 60.0 - 63.6% TCP giving a calculated rescue efficiency of 40 - 54.4

% (Figure 3.20). Phenotypic analysis for the two DAZ1-CR variants revealed no evidence to suggest that modification of the conserved region changes its localisation in pollen compared to full-length DAZ1-mCherry (Figure. 3.21 and 3.22). These data strongly suggest that the DAZ1 CR has an important role for DAZ1 function in germ cell division, but is not important for DAZ1 protein localisation.

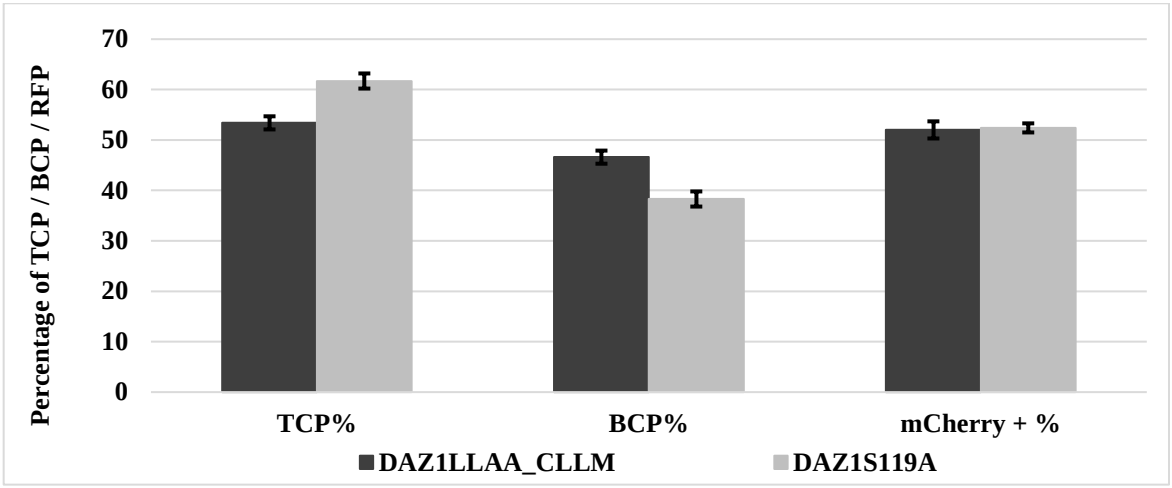


Figure 3.20. Complementation analysis with DAZ1-LLAA and DAZ1S119A variants. Phenotypic data for single locus DAZ1LLAA-CLLM and DAZ1S119A T1 lines are shown, indicting partial rescue of germ cell division in *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ splants. The mean percentages of TCP, BCP and mCherry positive (mCherry+) pollen are shown for each construct. Error bars indicate SE (n=5 counts for each construct).

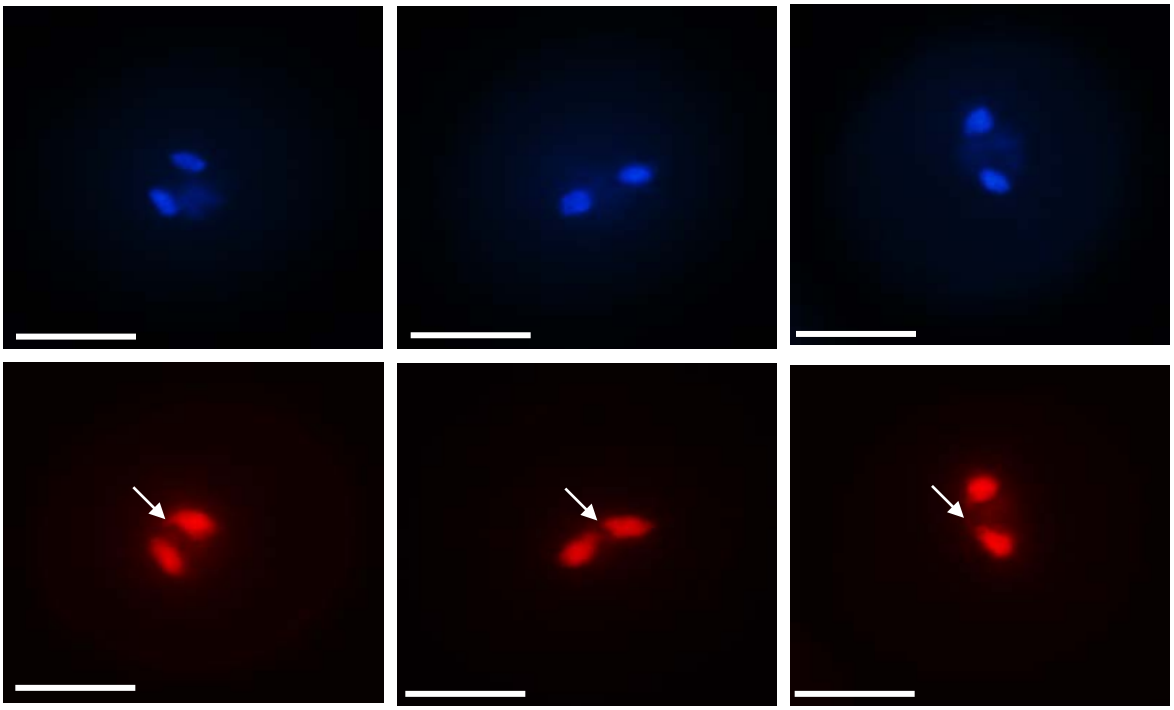


Figure 3.21. Fluorescence micrographs of pollen expressing DAZ1LLAA-CLLM-mCherry. White arrows indicate cytoplasmic and nuclear signals in sperm cells, mCherry signals extend over a larger area than nuclear DAPI signals. (Scale bar = 10 µm).

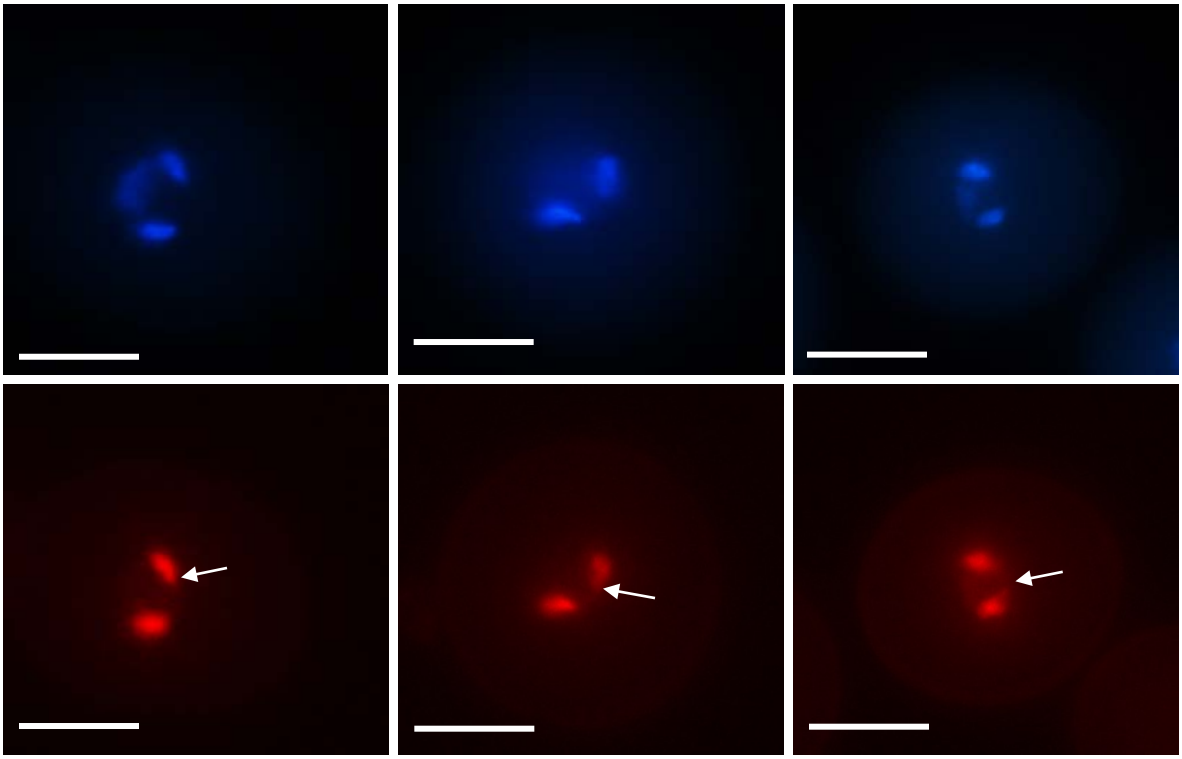


Figure 3.22. Fluorescence micrographs of pollen expressing DAZ1S119A-mCherry. White arrows indicate cytoplasmic and nuclear signals in sperm cells, mCherry signals extend over a larger area than nuclear DAPI signals (scale bar = 10µm).

3.3.6 Male transmission analysis for Basic region (BR) and Conserved region (CR) variants

The transmission of mutant *daz1 daz2* pollen from *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants expressing DAZ1 BR or CR variants was determined. Single locus T-DNA lines were chosen that were hemizygous for the transgene and crossed with *ms1-1*. It was predicted that for single locus lines half of the *daz1-1 daz2-1* pollen will receive the transgene and therefore have the potential to transmit *daz1-1 daz2-1,* while the remaining half will not be able to transmit *daz1-1 daz2-1* without the transgene. This results in 3 progeny classes in the F1 generation plants with a 2:1 ratio of pptR (resistant) to pptS (sensitive) seedlings as the transgene carries pptR and is transmitted with *daz1-1 DAZ2* and *daz1-1 daz2-2* pollen. Statistical analysis was

carried out by using chi-square analyses to determine differences from 2:1 (full rescue) and 1:1 (no rescue) pptR:pptS (Table 3.2).

Table 3.2. Male transmission efficiency of DAZ1-BR and DAZ1-CR variants from outcrosses. 2:1 and 1:1 ratios of PPT^R to PPT^S seedlings were predicted for full rescue and no rescue respectively. ns – not significant P>0.05 *, P<0.01 **, P<0.001 ***.

					Based on 1:1 ratio		Based on 2:1 ratio	
Construct	**Lines**	**PPT^R**	**PPT^S**	**R:S**	**χ^2**	**Significance**	**χ^2**	**Significance**
AtDAZ1	C5	76	40	1.9	5.7	*	0.1	ns
DAZ1muBR	D5_D4.1	50	43	1.2	0.5	ns	7.0	**
DAZ1ΔBR	D4_B4.2	60	45	1.3	2.1	ns	4.3	*
DAZ1S119A	A4	102	75	1.4	4.1	*	6.5	**
DAZ1S119A	C3	135	70	1.9	20.6	***	0.1	ns
DAZ1LLAA_CLLM	A2	72	63	1.1	0.6	ns	10.8	***
DAZ1LLAA_CLLM	A5	85	63	1.4	3.3	ns	5.7	**

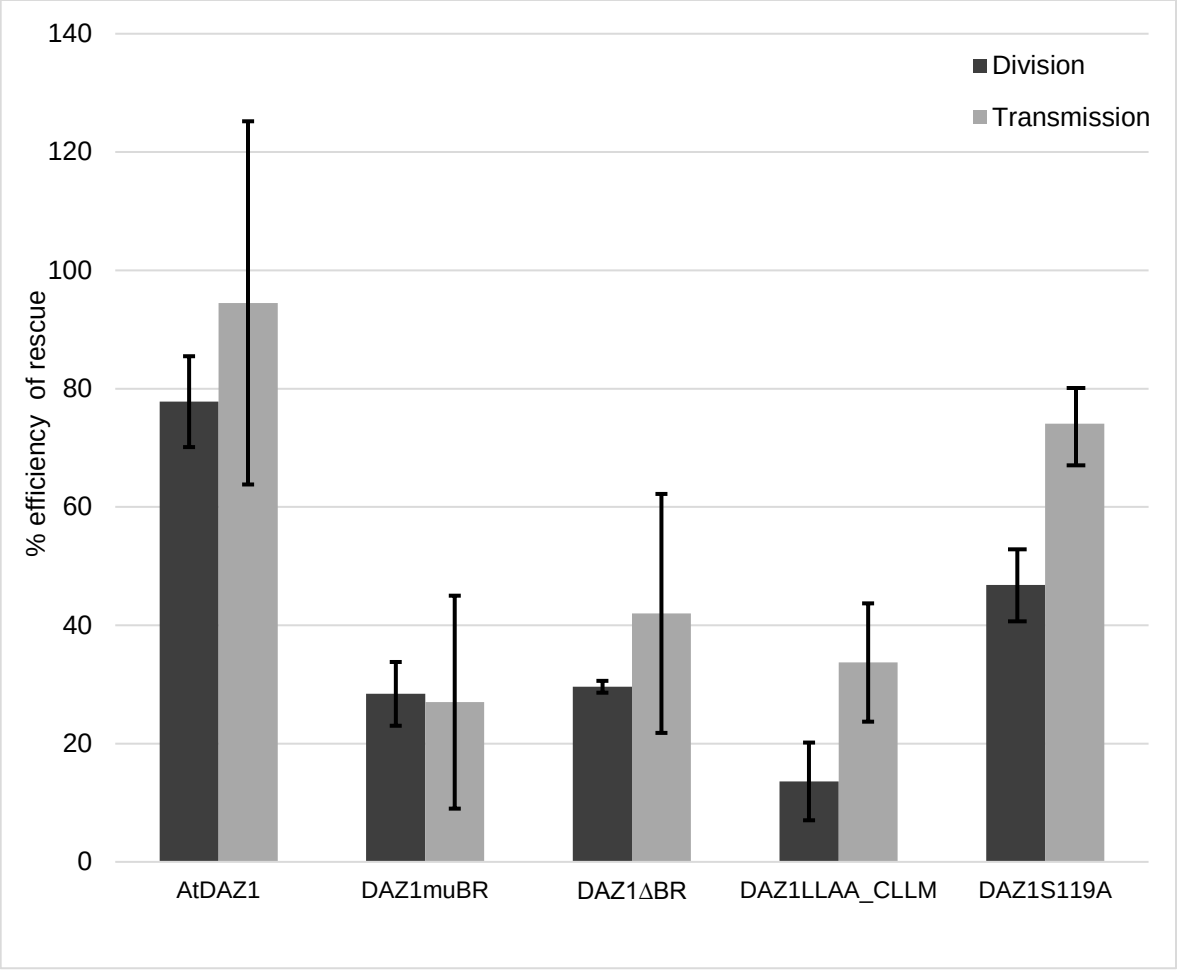


Figure 3.23. Functional analysis of DAZ1 domain variants *in planta*. DAZ1 domain variants were introduced into *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ mutant plants. The efficiencies of rescue are calculated relative to maximum values (100 % division rescue = 75 % TCP; 100 % transmission = 2:1 pptR:pptS seedlings). Error bars indicate SE (n, data see Table 3.2)

The analysis of male transmission efficiency indicates that DAZ1-BR proteins show partial rescue. The mean percentage of pptR F1 seedlings was greater than 50 %, which is expected for no transmission. DAZ1muBR gave 54.5 % pptR seedlings and DAZ1ΔBR 57 %, giving calculated rescue efficiencies of 27 % and 42 % respectively (Figure 3.23, Table 3.2). DAZ1LLAA-CLLM gave 55.6 % pptR F1 seedlings and DAZ1S119A 62.3 %, giving calculated rescue efficiencies of 33.7 % and 74.1 % respectively (Figure 3.23, Table 3.2). The transmission data for

DAZ1LLAA-CLLM deviated significantly from 2:1 pptR : pptS ratio for full rescue (Table 3.2 and Figure 3.23). The combined data for two DAZ1S119A lines (A4, C3) did not differ significantly from a 2:1 ratio (X^2, 3.17 $p<0.05$), but when tested individually line A4 show deviation from 2:1 while C3 did not (Table 3.2).

3.3.7 Localisation of DAZ1 in germ cells

Experimental data from previous results suggested that full-length DAZ1-mCherry and DAZ2-mCherry behave slightly differently with DAZ2 being strictly nuclear, while DAZ1 was nuclear-enriched, but with some cytoplasmic signal in approximately 35 % of pollen (Borg *et al.* 2014). Mature pollen was examined to determine the localisation of DAZ1-mCherry and DAZ2-mCherry proteins and mutant variants. The data presented in Figures 3.24 and 3.25 shows the percentage of pollen expressing DAZ1muBR or DAZ1ΔBR-mCherry variants with nuclear and nuclear/cytoplasmic RFP signals.

(1) DAZ1muBR-mCherry

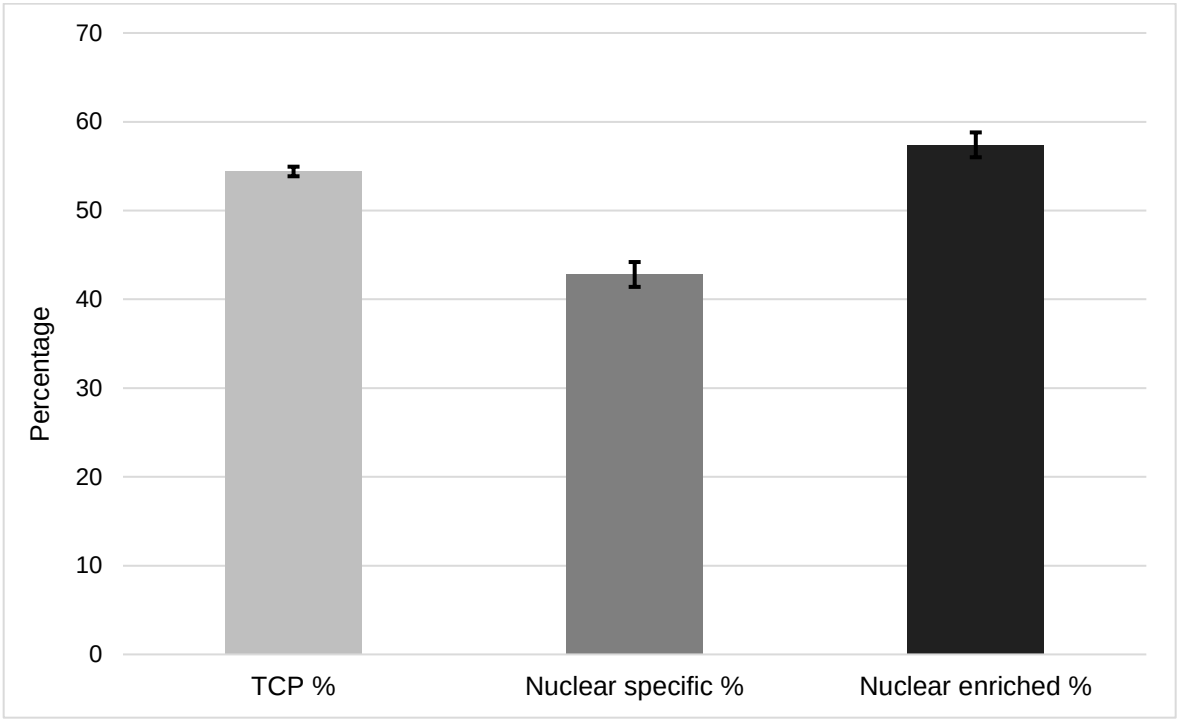


Figure 3.24. Localisation of DAZ1muBR-mCherry in pollen. The percentage of pollen which show nuclear-specific or nuclear-enriched localisation is plotted, along

with the percentage of TCP pollen in T2 lines in the *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ background. Error bars indicate SE.

(2) DAZ1-DAZ1ΔBR-mCherry

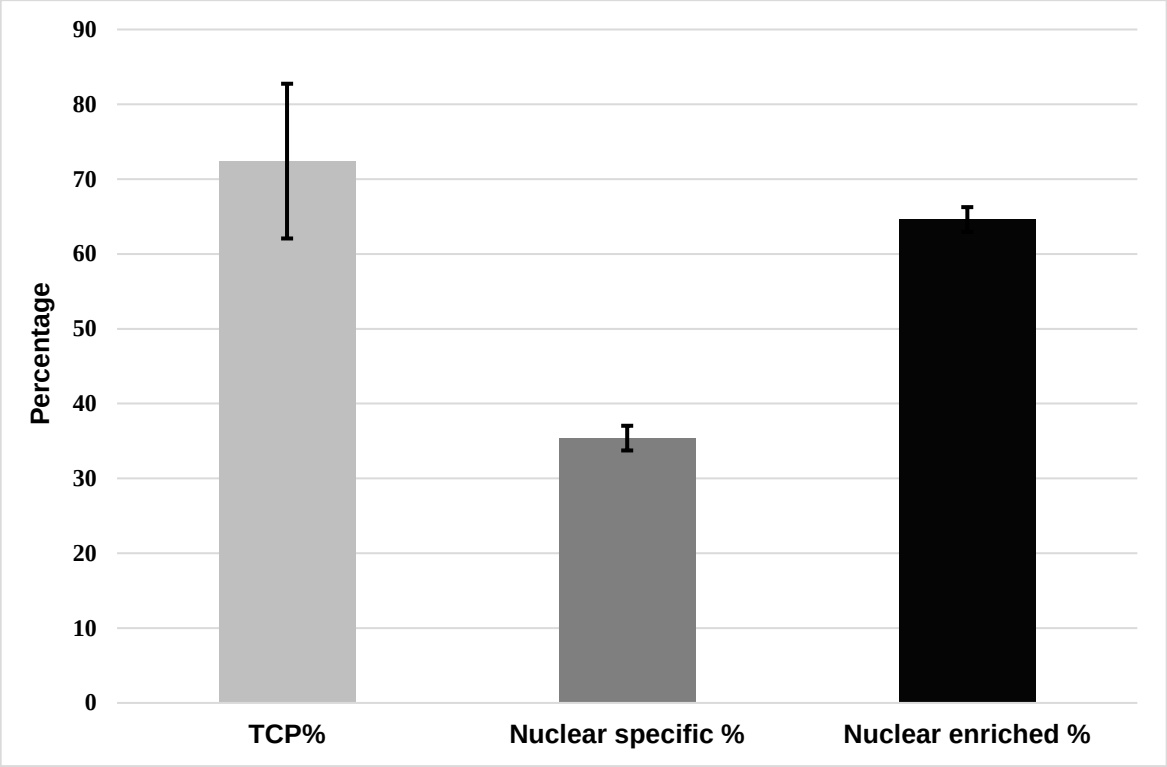
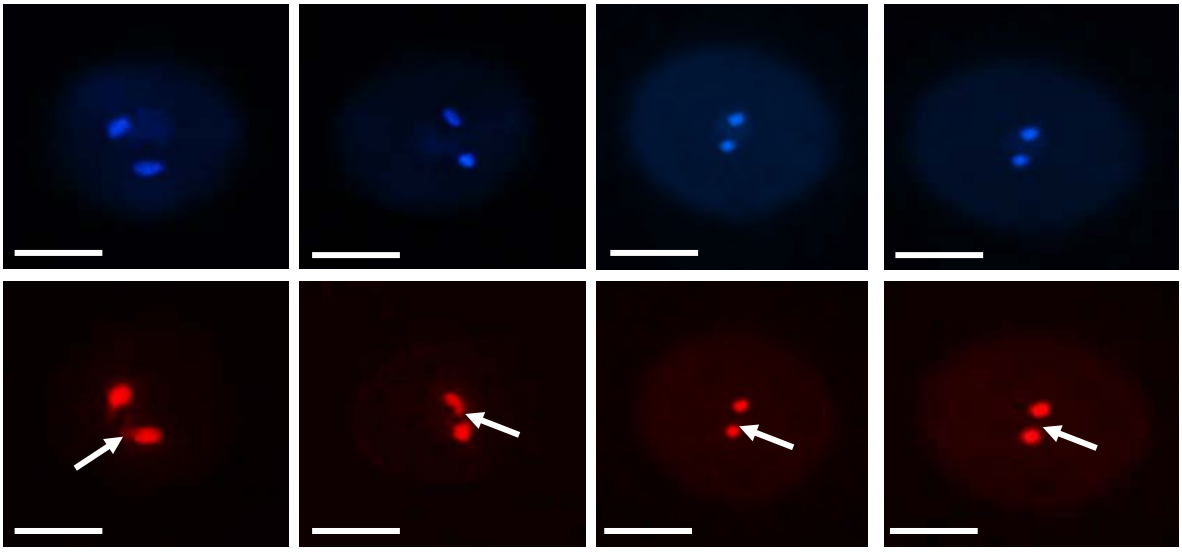
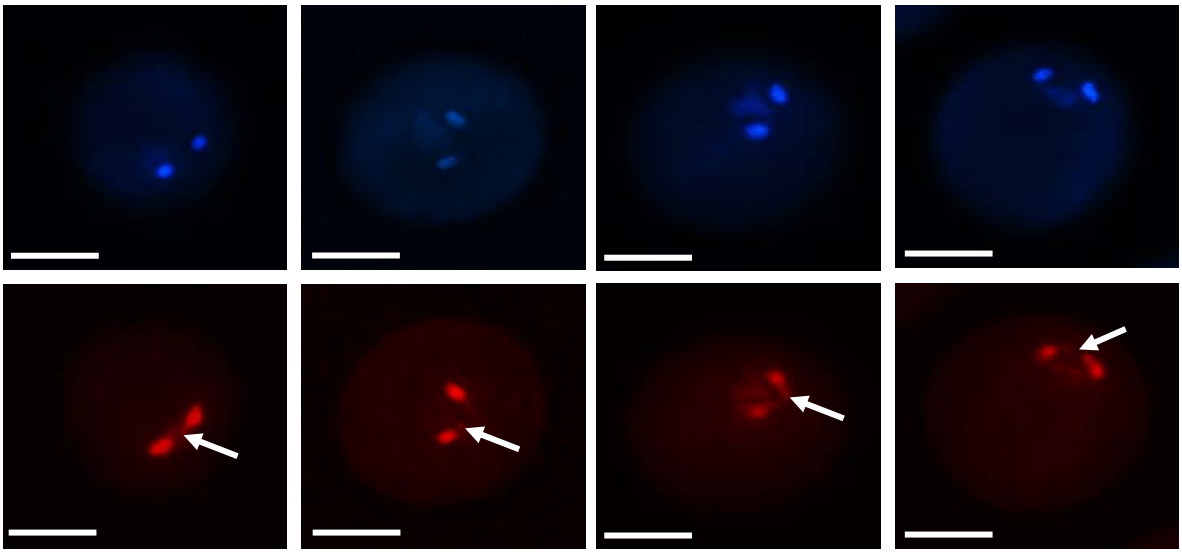


Figure 3.25. Localisation of DAZ1ΔBR-mCherry in pollen. The percentage of pollen which show nuclear-specific or nuclear-enriched localisation is plotted, along with the percentage TCP in a T1 line of *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ background. Error bars indicate SE.

Examination of DAZ1muBR and DAZ1ΔBR-mCherry localisation in pollen showed that both had nuclear-enriched signals with lower cytoplasmic signal in ~ 60 % - 64 % of pollen, similar to full-length DAZ1 with ~ 63 %. (Figure 3.24, 3.25; 3.26 A – D). These data show that the DAZ1 BR does not have important role in the distribution of DAZ1 in germ cells.



(A) AtDAZ1-mCherry (B) AtDAZ2-mCherry



(C) AtDAZ1muBR-mCherry (D) AtDAZ1BRΔ-mCherry

Figure 3.26. Localisation of **DAZ1/DAZ2-mCherry and DAZ1mu/ΔBR-mCherry in sperm cells.** Epifluorescence microscopy for DAPI (blue) and RFP (red) to show the protein location for (A) DAZ1-mCherry, (B) DAZ2-mCherry, (C) DAZ1muBR-mCherry, and (D) DAZ1ΔBR-mCherry. White arrows indicate cytoplasmic RFP signals compared with nuclear DAPI signals in sperm (scale bars =10 µm).

3.3.8 Analysis of localisation of *DAZ1* variants

Based on previous analysis AtDAZ1 is nuclear-enriched but has cytoplasmic signals in 63.1 % of sperm, while DAZ2, is almost exclusively nuclear-specific (Borg *et al.* 2014). DAZ1-BR mutants showed a similar localisation to full-length AtDAZ1 with 64 % nuclear-enriched protein signals (Figure 3.27). Further DAZ1 variant proteins were also examined to try and understand factors and other domains that may affect protein localisation.

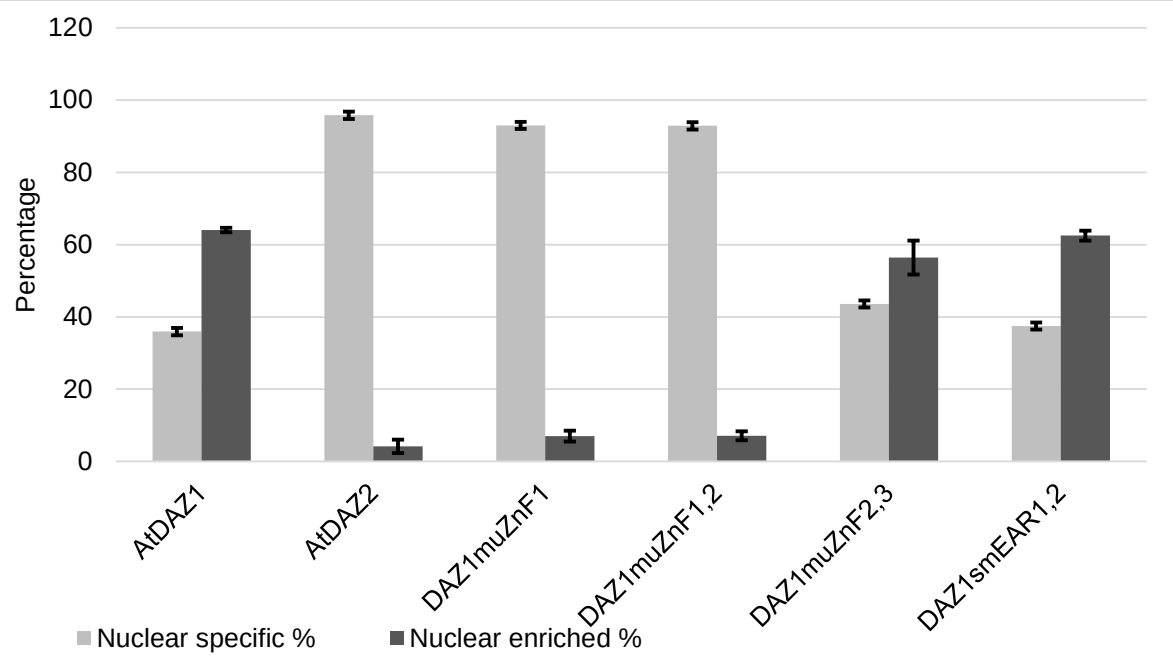
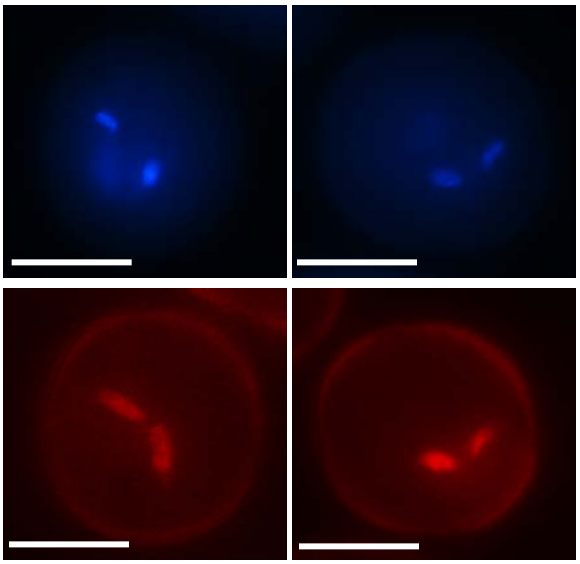


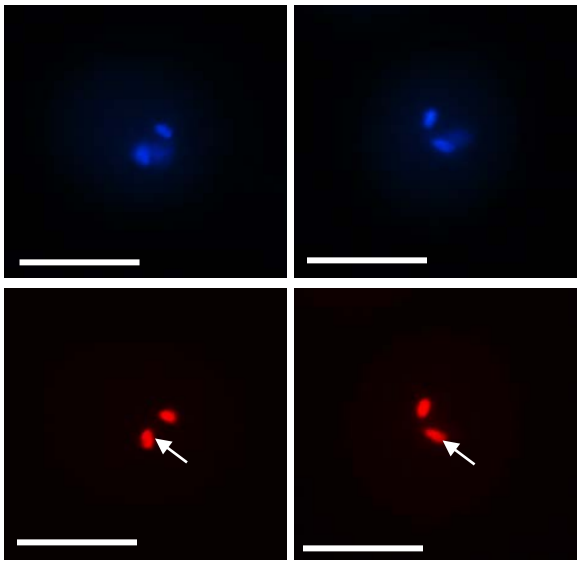
Figure 3.27. Localisation of DAZ1 ZnF and EAR variants. The percentage of pollen which show nuclear-specific or nuclear-enriched (nuc/cyt) localisation for DAZ1-mCherry variants with mutations in ZFs and in the EAR motifs. Error bars indicate SE.

DAZ1ZnF2,3-mCherry and DAZ1smEAR1,2-mCherry showed a similar distribution of 56.4 % to 64.1 % nuclear-enrichment (Figure 3.27). In contrast AtDAZ2-mCherry shows nuclear-specific localisation in 95.8 % of pollen. These results indicate that DAZ1 protein is localised in a similar way even after mutation, which suggests that DAZ1 ZnF2 and ZnF3 do not have a role in nuclear localisation (Figure 3.28).

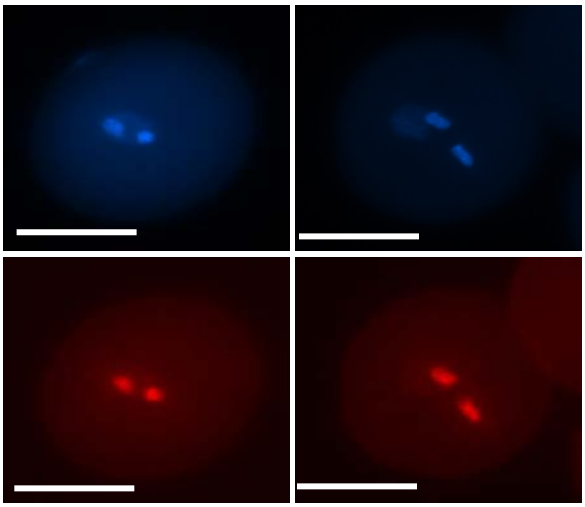
However, DAZ1muZnF1-mCherry and DAZ1muZnF1,2-mCherry variants shows different localisation which was not expected. Both proteins show more nuclear-specific signal compared with other DAZ1 molecules. DAZ1muZnF1 and DAZ1muZnF1,2 show 92.9 % and 93 % nuclear-specific and 7 % to 7.1 % nuclear-enriched signals (Figures 3.27 and 3.28).



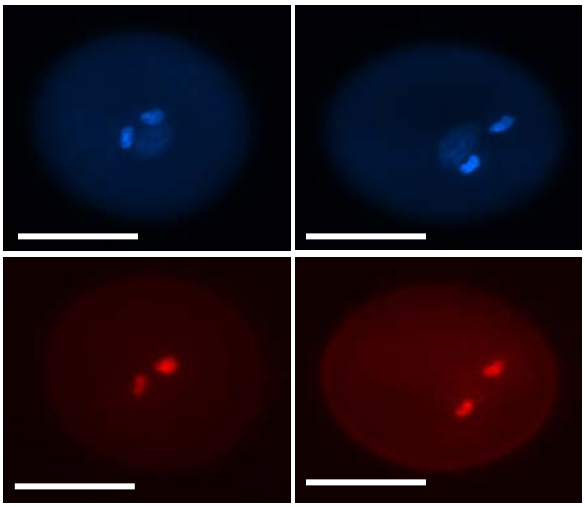
(A) AtDAZ1



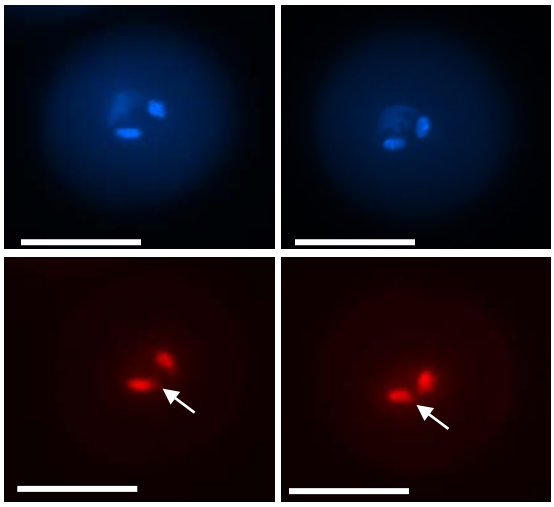
(B) AtDAZ2



(C) DAZ1muZnF1



(D) DAZ1muZnF1,2



(E) DAZ1muZnF2,3

Figure 3.28. Mutant DAZ1 protein localisation. (A) AtDAZ1, (B) AtDAZ2, (C) AtDAZ1muZnF1, (D) AtDAZ1muZnF1,2, and (E) AtDAZ1muZnF2,3. White arrows indicate cytoplasmic RFP signals compared with nuclear DAPI signals in sperm (scale bars =10 µm).

3.3.9 Localisation of DAZ1 protein orthologues

As described earlier DAZ1 is conserved in angiosperms and DAZ1 sequence alignments show that the BR is not conserved in monocot species (eg. rice, maize) (Figure 3.2) and ZnF3 is absent in tomato (Figure 3.5B). Thus, the localisation of DAZ1 orthologues was investigated to examine if their localisation was similar to that of AtDAZ1 or AtDAZ2. ZmDAZ1A/B, OsDAZ1, SlDAZ1 and BrDAZ1 were examined together with DAZ1/DAZ2-mCherry as controls.

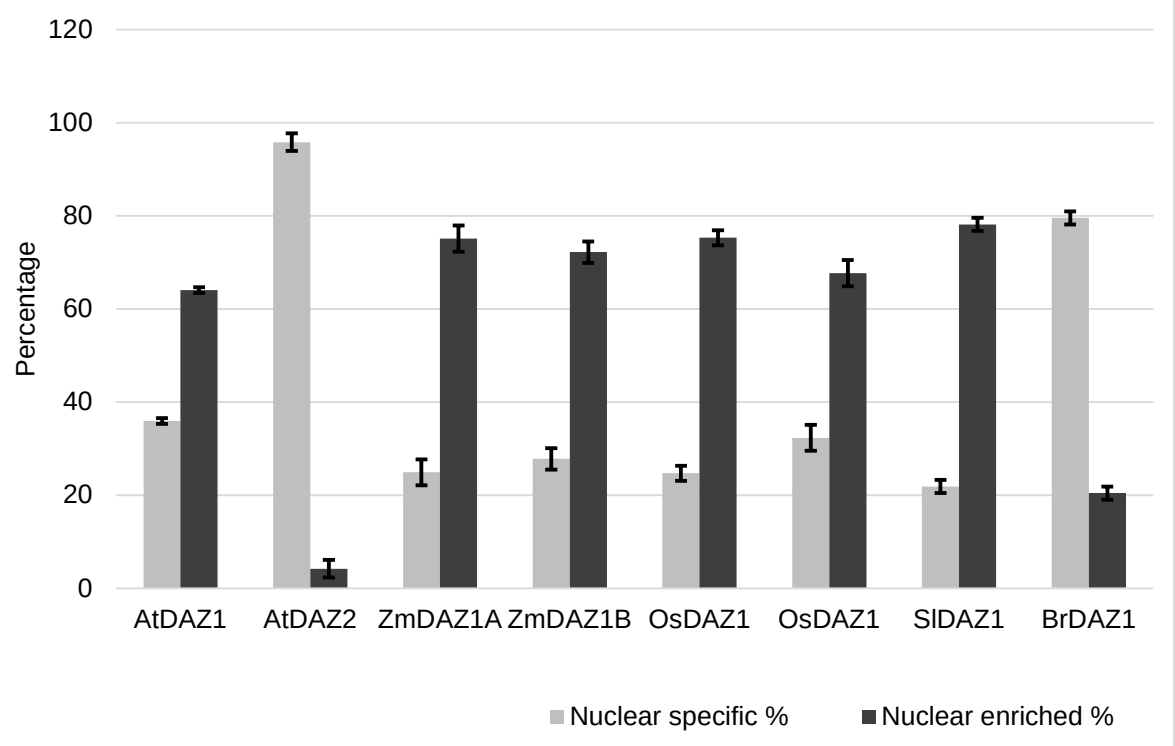


Figure 3.29. Localisation of DAZ1 orthologs fused to mCherry. The chart shows the percentage of pollen-expressing DAZ1 orthologues with nuclear-specific or nuclear-enriched (nuclear/cytoplasmic) localisation. Abbreviations: Zm - *Zea mays,*

Os - *Oryza sativa,* Sl - *Solanum lycopersicum,* and Br - *Brassica rapa* along with full length AtDAZ1 and AtDAZ2 as controls. Error bars indicate SE (n = 5).

DAZ1 orthsologues showed similar patterns of localisation to AtDAZ1-mCherry. ZmDAZ1, OsDAZ1 and SlDAZ1-mCherry protein signals were mostly nuclear-enriched (67.7 % to 78.2%) with 21.9 % to 32.3 % of pollen showing nuclear-specific signals (Figure 3.29 and Figure 3.30D - G). BrapaDAZ1 showed completely opposite results with 79.6 % of pollen showing a nuclear-specific signals similar to AtDAZ2 (Figure 3.29 and Figure 3.30C). With the exception of BrapaDAZ1, the localisation pattern of DAZ1 is conserved among DAZ1 orthologues from diverse taxa of flowering plants.

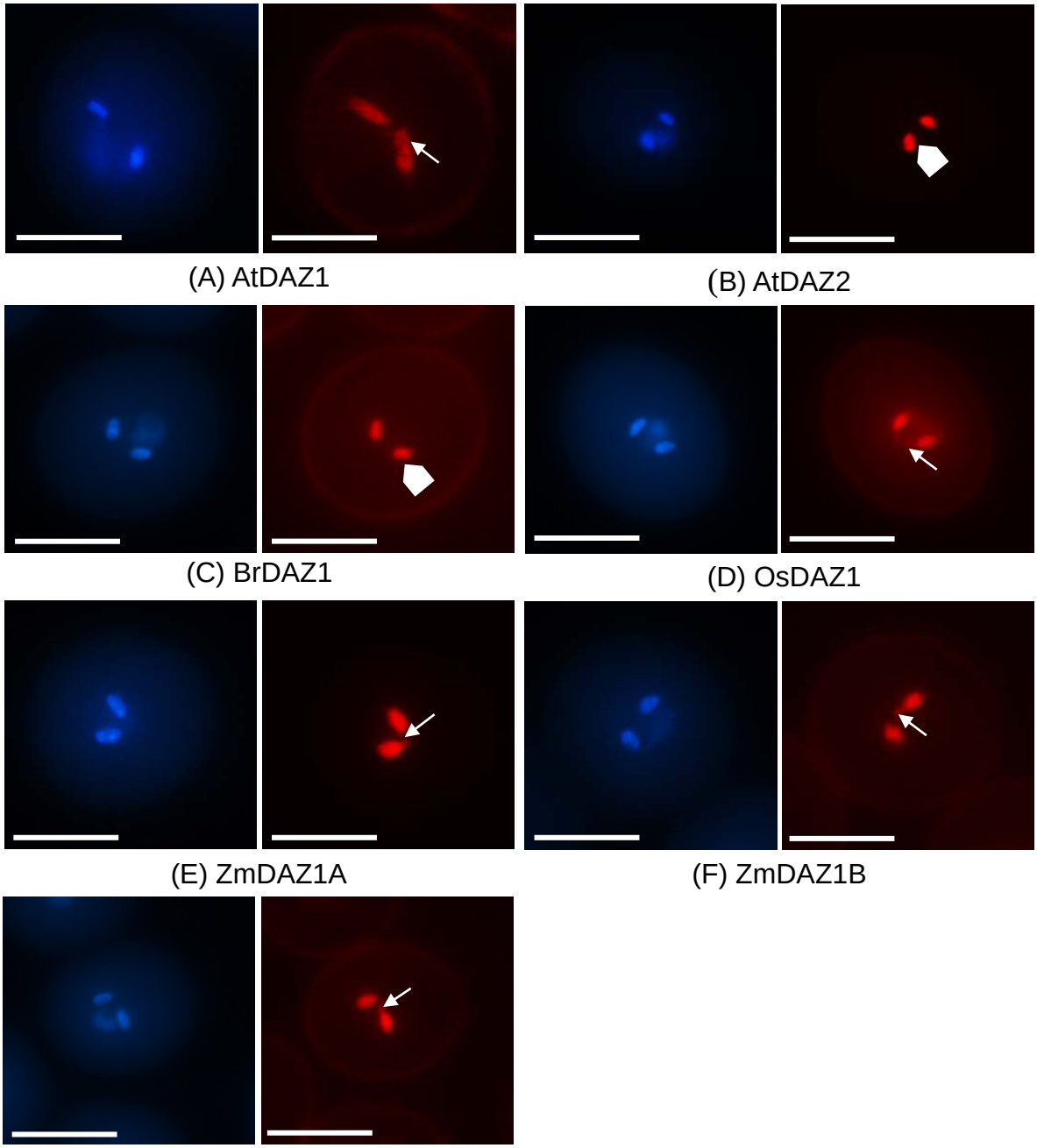


Figure 3.30. DAZ1 orthologs fused with RFP show nuclear localisation in sperm cells. (A) AtDAZ1, (B) AtDAZ2, (C) BrDAZ1, (D) OsDAZ1, (E) ZmDAZ1A, (F) ZmDAZ1B and (G) SlDAZ1. White arrows on RFP images indicate signal is present in the nucleus and the cytoplasm in A, D, E, F and G (compare RFP and DAPI images). White arrowheads indicate nuclear-specific signals in B and C.

Abbreviations: At – *Arabidopsis thaliana,* Br - *Brassica rapa,* Os - *Oryza sativa,* Zm – *Zea mays,* Sl – *Solanum lycopersicum.* Scale bar = 10µm.

3.4 Analysis of DAZ1 ZnF domains

Previous experimental research data suggest that ZnF1 is fundamentally important for DAZ1 function and that ZnF2 and ZnF3 are important but not critical for DAZ1 function (Rutley, 2015). It was also shown that SlDAZ1 which lacks ZnF3 was only able to rescue germ cell division partially (56% - 58% TCP or 25-32 % efficiency) respectively in *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ background *A. thaliana* plants (Sari, 2014). The current investigation aimed to validate and extend previous findings work to include analysis of the function of a new DAZ1 variant with both ZnF2 and ZnF3 mutated, which was not previously investigated.

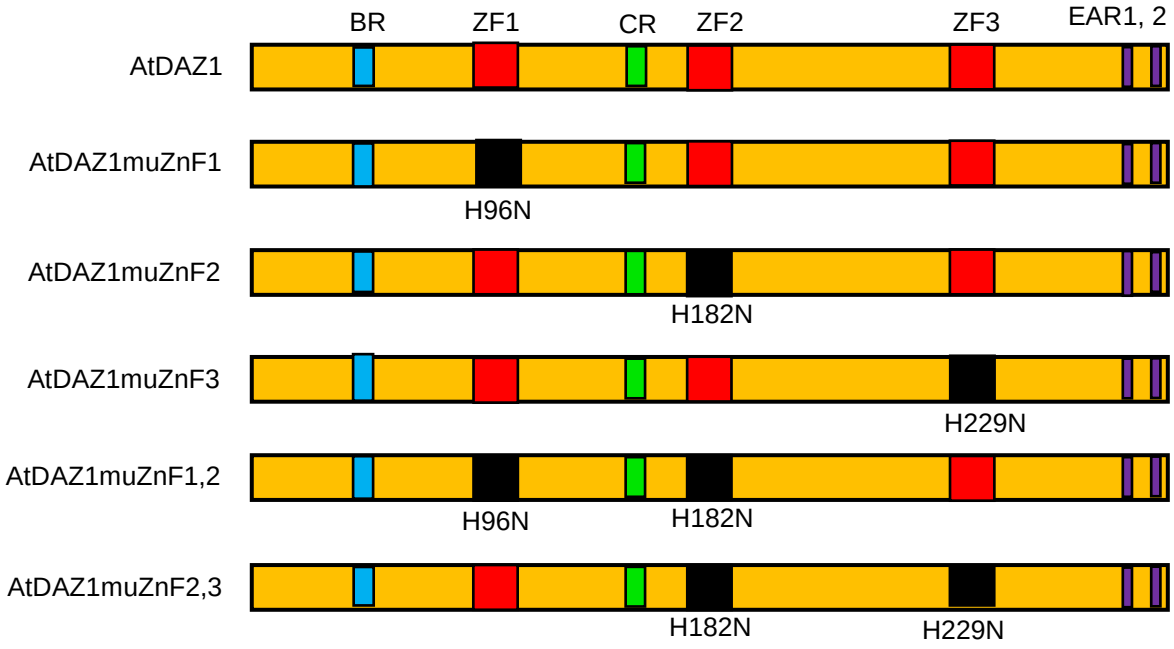


Figure 3.31. DAZ1 ZnF domain variants. AtDAZ1 and AtDAZ1 ZnF mutant variants are illustrated. Red and black boxes represent ZnF domains and black boxes the substitution mutants (histidine to asparagine) corresponding to positions:

DAZ1muZnF1 (H96N), DAZ1muZnF2 (H182N), DAZ1muZnF3 (H229N), DAZ1muZnF1,2 (H96N and H182N) and DAZ1muZnF2,3 (H182N and H229N).

The strategy used was described in Rutley (2015), in which site-specific mutation in ZnF domains were introduced by overlap-extension PCR, resulting in a substitution of the codon corresponding to the zinc ion-coordinating histidine residue at position +7 of the α-helix with an asparagine codon (Yim, 2013, Fisher, 2014), (Rutley 2015). Total five zinc finger mutation was created for DAZ1muZnF variants. DAZ1muZnF1 (H96N), DAZ1muZnF2 (H182N), DAZ1muZnF3 (H229N), DAZ1muZnF1,2 (uniting H96N and H182N) and last DAZ1muZnF2,3 (combining H182N and H229N) (Figure 3.31). All mutant DAZ1muZnF CDS parts were recombined into the destination vector pB7m34GW, downstream of a DAZ1 promoter fragment and upstream and in-frame of the coding sequence of the fluorescent protein mCherry (Rutley 2015).

3.4.1 Rescue of germ cell division by mutant DAZ1 ZnF domain variants

To investigate the functional importance of combinations of mutations in DAZ1 ZnF2 and ZnF3, T1 generation plants expressing DAZ1muZnF2,3-mCherry were analysed by scoring TCP and BCP from plants in the *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ background. Previous results obtained for DAZ1muZnF1-mCherry, DAZ1muZnF2-mCherry, and DAZ1muZnF3-mCherry constructs were also validated by analysing T2 generation plants.

The activities of DAZ1muZnF1, DAZ1muZnF2 and DAZ1muZnF1,2 proteins in T2 generation plants confirmed that ZnF1 is essential for DAZ1 function. DAZ1muZnF1 and DAZ1muZnF1,2 gave mean TCP values of 44 % and 43 % respectively, indicating failure to rescue germ cell division (Figure 3.32). DAZ1muZnF2 gave 61 % TCP (44 % rescue efficiency), compared with 75 % TCP for the DAZ1 contro1 (100 % rescue efficiency). Moreover, DAZ1muZnF2 showed comparable rescue of germ cell division in T1 (57 % TCP) and T2 (61 % TCP) generations (Figure 3.33).

These data validate the previous findings that ZnF1 is essential for DAZ1 function, and suggest that DAZ1muZnF2 although not critical has an important role.

In addition, T2 generation DAZ1muZnF3 plants were examined along with T1 DAZ1muZnF2,3 plants. T2 DAZ1muZnF3 plants gave 63.4 % TCP while T1 DAZ1muZnF3 plants gave 65.2 % TCP. These values were comparable with those of DAZ1muZnF2 (57-61 % TCP). The overall analyses of single and double mutant DAZ1 ZnF domains indicates that only ZnF1 is critical for DAZ1 function, whereas, ZnF2 and ZnF3 have a non-essential but significant functional role in the germline.

In summary, the mutation of different ZnF domains had variable effects on the rescue of germ cell division in *daz1-1 daz2-1* pollen (Table 3.3, Figure 3.34). DAZ1muZnF1 and DAZ1muZnF1,2 showed no rescue, with no significant deviation from the percentage TCP for no complementation ($P>0.001$), whereas, DAZ1muZnF2, DAZ1muZnF3 and DAZ1muZnF2,3 showed a significant rescue of germ cell division. However, division rescue in DAZ1muZnF2, DAZ1muZnF3 and DAZ1muZnF2,3 were lower and significantly different from that of AtDAZ1 ($P> 0.001$) (Table 3.3, Figure 3.34).

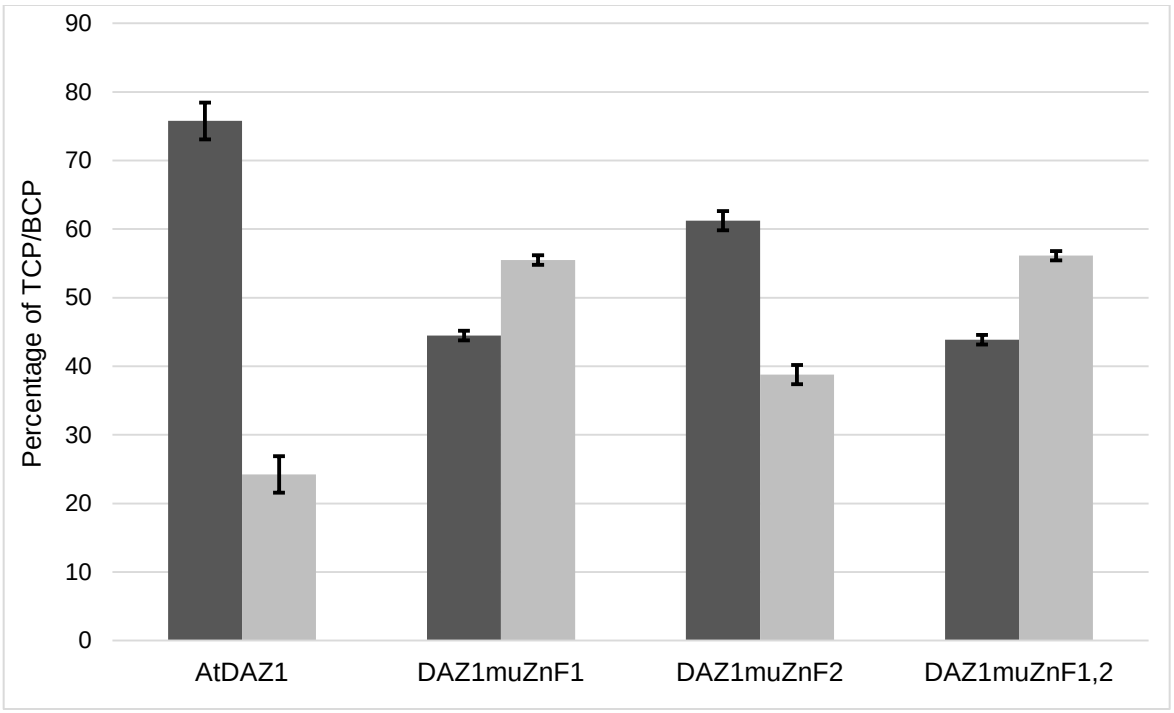


Figure 3.32. Ability of DAZ1 mutant ZnF (ZnF1, ZnF2) variants to rescue germ cell division. The chart shows mean values for five T2 plants. Error bars indicate the standard error of the mean (n =5). Tricellular pollen (dark bars), bicellular pollen (light bars). Error bars indicate SE.

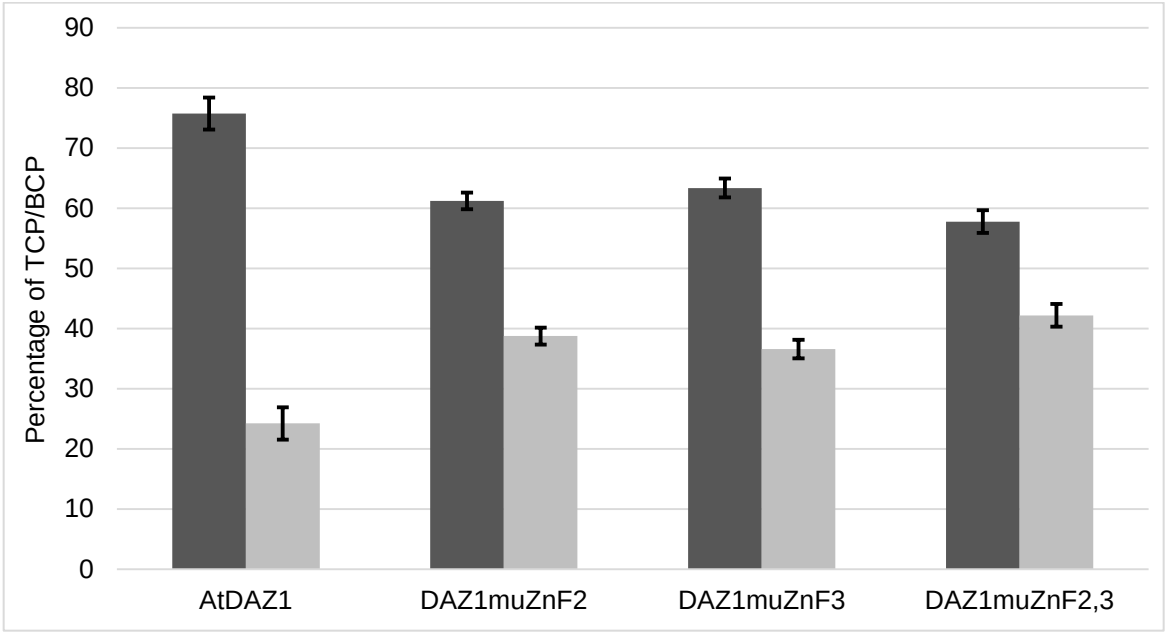


Figure 3.33. Ability of DAZ1 mutant ZnF (ZnF2, ZnF3) variants to rescue germ cell division. The chart shows mean values for five T2 plants. Tricellular pollen (dark bars), bicellular pollen (light bars). Error bars indicate SE (n = 5).

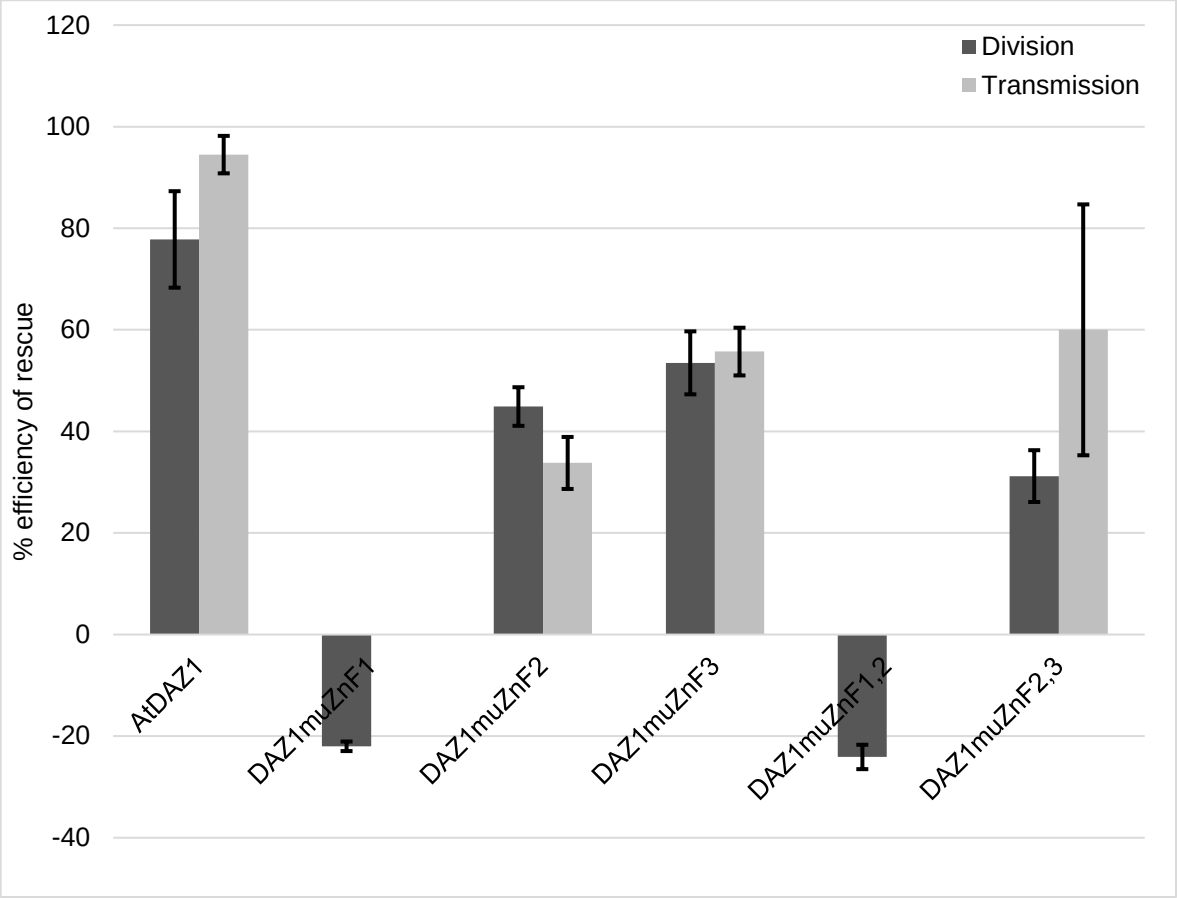


Figure 3.34. Functional analysis of DAZ1 domain variants *in planta*. DAZ1 mutant ZnF domain variants were introduced into *daz1-1*$^{-/-}$*; daz2-1*$^{+/-}$ mutants. The efficiencies of rescue are calculated relative to maximum values (100 % division rescue = 75 % TCP; 100 % transmission = 2:1 pptR:pptS seedlings). Error bars indicate SE.

Table 3.3. Germ cell division rescue for DAZ1muZnF variants in *daz1-1$^{-/-}$; daz2-1$^{+/-}$* plant. Chi-square tests were calculated by using expected values for no complementation (1:1 ratio of TCP to BCP) and the rescue efficiency of AtDAZ1 (2.27:1 ratio). ns, not significant P<0.05 *, P<0.01 **, and P<0.001***.

						Based on no rescue 1:1		Based on AtDAZ1 rescue 2.27:1	
Construct	**No. of lines**	**TCP**	**BCP**	**SUM**	**TCP: BCP**	**χ2**	**Signifi-cance**	**χ2**	**Signifi-cance**
AtDAZ1	6	834	367	1201	2.27	94.3	***	0.05	ns
DAZ1muZnF1	10	1101	1311	2412	0.84	9.2	**	268.6	***
DAZ1muZnF2	10	1394	888	2282	1.57	56.8	***	31.6	***
DAZ1muZnF3	6	711	424	1135	1.68	36.9	***	10.34	*
DAZ1muZnF1,2	10	1122	1516	2638	0.74	29.5	***	374.3	***
DAZ1mZuZnF2,3	5	733	536	1269	1.37	15.4	***	34.7	***

3.4.2 Male transmission efficiency of DAZ1muZnF2,3

The transmission efficiency of DAZ1muZnF2,3 was determined by scoring among F1 progeny pptR to pptS seedlings on selection plates. Heterozygous DAZ1 muZnF lines in a *daz1$^{-/-}$ daz2$^{+/-}$* background were crossed onto *ms1-1* pistils. The data in Table 3.4 show that full-length DAZ1 deviated significantly from expected values for no male transmission (1:1 ratio of pptR to pptS seedlings) but two DAZ1muZnF2,3 lines did not. AtDAZ1 and DAZ1muZnF2,3 lines all deviated significantly from full transmission (2:1 ratio).

Table 3.4. Male transmission efficiency analysis for mutant DAZ1-BR and DAZ1-CR variants. Chi-square tests were calculated by using expected values for no transmission (1:1 ratio of pptR to pptS seedlings) and full transmission (2:1 ratio). ns, not significant P<0.05 *, P<0.01 **, and P<0.001***.

					Based on 1:1		Based on 2:1	
Construct	**Line ID**	**PPT^R**	**PPT^S**	**R:S**	**χ^2**	**Significance**	**χ^2**	**Significance**
AtDAZ1	C5	30	8	3.8	6.9	**	1.6	ns
DAZ1muZnF2,3	Line_2_C1.1	72	48	1.5	2.4	ns	1.1	ns
DAZ1muZnF2,3	Line_5_B5.1	57	37	1.5	2.2	ns	0.6	ns

3.5 Discussion

3.5.1 The DAZ1 basic region is essential for promoter *trans*-activation

To investigate the functional significance of the DAZ1 BR, this region was mutated and deleted and the ability of these modified proteins to activate the *GEX1* promoter tested. The results showed that the activity of DAZ1 BR variants (DAZ1muBR / DAZ1ΔBR) was highly reduced or inactive compared with full-length DAZ1 (Figure 3.15). These results suggest that the DAZ1 BR is an important protein domain required for promoter *trans*-activation by DAZ1.

3.5.2 Protein localisation in *N. benthamiana* leaf cells

The DAZ1 basic region is conserved in eudicots and is present in DAZ1 from the most basal extant angiosperm *Amborella trichopoda.* However, in monocot species the BR is absent, so it is a mystery why eudicot species retain the BR while monocots have lost this sequence motif. A speculation is that unknown sequences in DAZ1 replaced BR function, or that partner proteins were recruited to substitute for BR function in monocot DAZ1, or simply the requirement for BR function was lost.

Previous research shows that short basic amino acid sequences can support nuclear protein import, for example the monopartite NLS as discovered in SV40 large T-antigen (Dingwall *et al.* 1982). To understand the potential role of the BR in nuclear transport, expression in *N. benthamiana* leaf cells showed evidence that DAZ1-GFP BR variants were delocalised and appeared in the cytoplasm, in contrast full-length DAZ1 which was localised strictly to the nucleus (Figure 3.18). These data strongly suggest that the DAZ1 BR can function as an NLS and is required for tight localisation of DAZ1 to the nucleus. Since small stretches of basic amino acids are known to bind importin α / β and support transport though the nuclear pore complex, it may be predicted that the BR may bind to α-importin. There remains the need to further investigate the divergence between eudicot and monocot DAZ1 proteins in relation to the localisation and functional role for DAZ1.

3.5.3 The DAZ1 basic region is important for germ cell division *in planta*

The DAZ1 BR is conserved in DAZ1 proteins from eudicots but has been lost in monocot species like rice, sorghum and maize. DAZ1 also contains different domains (Zinc finger 1, 2, 3) and two EAR motifs and their involvement in DAZ1 function was studied. Deletion (DAZ1ΔEAR1, 2) and mutation (DAZ1muEAR1, 2) variants show severely reduced DAZ1 function *in planta*, likely due to loss of interaction with TOPLESS (Borg *et al.* 2014). However, deletion and mutation of the DAZ1 BR demonstrated the important role for the BR in germ cell division (Figure 3.19). These data demonstrate a previously unknown functional role for the DAZ1 BR in Arabidopsis germ cell development, which supports twin sperm cell production and male fertility.

The potential role of the DAZ1 conserved region (CR) has not been previously remained, but the results from *in planta* complementation assays shows that the DAZ1 proteins show reduced function when the CR region is mutated. Compared with full-length DAZ1, DAZ1LLAA-CLLM, DAZ1S119A showed reduced rescue of

male germline division (Figure 3.20), which suggests that the DAZ1 CR has important functions or interactions.

3.5.4 DAZ1 protein localisation in *A. thaliana* germ cells

The observed difference in DAZ1-mCherry protein localisation in germ cells (nuclear and cytoplasmic) compared with leaf cells (nuclear) may arise because there may be differences in DAZ1 protein interactions and/or components of the nuclear import pathways machinery. Alternatively, there may be differences between Arabidopsis germ cells and tobacco leaf cells used for transient expression assays. Interestingly DAZ1 and DAZ2 proteins possess the same BR sequence but do not show the same localisation in germ cells. DAZ1-mCherry and BR variants are nuclear-enriched, but also present in the germ cell cytoplasm (Figure 3.24, Figure 3.26.C, D). However, DAZ2-mCherry proteins are mostly tightly associated with the nucleus (Figure 3.26.B) consistent with the observation of Borg et al (2014).

In *Arabidopsis* CRY1 (Crypto-chrome 1) and CRY2 (Crypto-chrome 2) genes identified and well conserved within the N-terminal region (amino acid positions 1-500), sharing 54% identity (Kleiner *et al.* 1999). Whereas, CRY1 was found to be a soluble protein (Lin *et al.* 1996), but CRY1 was also found to be enriched in the membrane fraction (Ahmad *et al.* 1998). CRY2 are indicated as blue light photoreceptor which involved in measuring daylight. Experimental data shows that CRY2 NLS is very similar to NLS of proteins for the nuclear localisation signal (Görlich and Mattaj 1996, Kobayashi *et al.* 1998, Thresher *et al.* 1998, Kleiner *et al.* 1999). Additionally, data shown in 1999 from Cashmore lab, fusion GFP protein consisting of *Arabidopsis* CRY1 and CRY2 localised into the nucleus on transient expression in onion epidermal cells, indicating that *Arabidopsis* CRY1 is a nuclear protein (Cashmore *et al.* 1999). Those previous results support current research in *Arabidopsis* DAZ1 basic region (BR) required for the protein localisation into the nucleus in leaf cells. Furthermore, in DAZ1 and DAZ2 protein location shows different nuclear protein phenotype in germ cells might be indicated that nuclear

protein transportation also involve alternative or additional protein localisation machinery that allows DAZ2 localise protein only in nucleus because, nuclear localisation of CRY2 is evolutionary conserved in plants and animals (Kleiner *et al.* 1999).

The basic amino acids lysine and arginine in NLS are important for the transport of nuclear localised proteins in the nucleus (Görlich and Kutay 1999, Huang *et al.* 2016) Demonstrated that MSA 1 localised in the cytosol could not complement mutant *msa1-1* and therefore this results showed that nuclear localisation is essential for MAS1 function in the regulation of S-homeostasis (Huang *et al.* 2016). In 2007, Roig-villanova lab demonstrated that PAR1 protein sequence does not contain any obvious canonical NLS but, fusion with GFP protein (PAR1-GFP) was shown to be nuclear and finding suggest that PAR1 protein contained a yet uncharacterised non-canonical NLS, therefore, precluding its recognition by the available subcellular localisation prediction tools (Galstyan *et al.* 2011).

However, in several DNA binding proteins, NLS motifs are embedded within the DNA-binding domains. For example, in bZIP (Basic-leucine zipper) proteins TGA-1A and TGA-2B was showed as a NLS in plants (van der Krol and Chua 1991) it suggest that might be in DAZ1/DAZ2 proteins basic region NLS might be binding with another basic-leucine zipper (bZIP) proteins within DNA binding domains in germ cells. Additionally, several *Arabidopsis thaliana* bHLH proteins also expected to comprise canonical NLS motifs within their DNA-binding basic domains (Galstyan *et al.* 2011) and another BIM1 protein, also contains a short stretch of basic residues that resemble a monopartite NLS shown to be nuclear signals (Chandler *et al.* 2009) those results suggesting that other sequence are playing hidden functional role in *Arabidopsis thaliana* orthologues, monocot species and mutant lines as NLS motifs, region or domains which substitute or replaces basic region function that allows DAZ1 less tight nuclear signal and in DAZ2 nuclear specific signals compared with other DAZ1 BrDAZ1, OsDAZ1, ZmDAZ1, SlDAZ1, muZnF1, muZnF2) monocot species and mutant lines. This results suggested that presence of the alternative

mechanisms to drive the mCherry fusion proteins to the nucleus. Furthermore, results support our results for the alternative mechanism involved in nuclear protein transportation i.e. PAR1 analysis data showed that truncation in PAR1 proteins lacking the N-terminus (AHC) were still able to localise in the nucleus and recruit its biological function, while AH version were shows cytoplasmic signal due to their abolish dimerization ability (Galstyan *et al.* 2011).

3.5.5 Localisation of DAZ1 variants and DAZ1 orthologs in sperm cells

Previous work demonstrated that mutations in DAZ1 ZnF1, result in loss of function in germ cell division, but mutation of ZnF 2 and ZnF3 are less severe and rescue germ cell division at 60-80 % efficiency of unmodified DAZ1. However, when these DAZ1 variants were tested for nuclear localisation *in planta,* they showed a surprising result. DAZ1muZnF1 and DAZ1muZnF1,2 show predominantly nuclear-specific localisation similar to the pattern of localisation of AtDAZ2-mCherry.The question arises how DAZ1muZnF1 can show an altered and tighter nuclear localisation. One possibility is that mutation of ZnF1 uncovers a cryptic nuclear localisation sequence present in DAZ1 which enhances nuclear localisation or allows binding to another component in the nucleus. In contrast, DAZ1muZnf2, 3 behaves similarly to AtDAZ1 and is able to rescue germ cell division as earlier mentioned indicating different role for ZF1 compared with ZF2 and ZF3. DAZ1 proteins with mutant EAR1,2 motifs show a pattern of localisation similar to AtDAZ1 (Figure 3.27).

DAZ1 proteins from monocot species OsDAZ1, ZmDAZ1A and ZmDAZ1B do not possess the typical BR present in AtDAZ1 or NLS-like sequences elsewhere in their sequences, but when they are expressed in Arabidopsis the majority of pollen show nuclear-enrichment with the remainder of 24.7 % - 32.3 % showing nuclear-specific signals (Figure 3.29, Figure 3.30). Other transcription factors have been shown to show conserved localisation between species. For example, the maize trans-acting factor, O2, regulates the expression of 22-KD ZEIN genes (Schmidt *et al.* 1992) and

has been shown to localise to the nucleus in maize endosperm tissue and in transformed tobacco plants, indicating that the protein nuclear import machinery operates similarly in monocot and dicots (Varagona *et al.* 1992).

3.5.6 Zinc finger domains are required for DAZ1 activity in *A. thaliana*

The analysis of the role of DAZ1 ZnF domains presented in this chapter suggest that all three ZnF domains are important and required to support the function of DAZ1 in generative cell division and sperm differentiation. ZnF1 was confirmed to be essential for DAZ1 function in *planta*. The severe effect of disrupting ZnF1 may arise because DAZ1 DNA binding ability is disturbed; since DAZ1 has been shown to bind *in vitro* to a consensus with the core sequence AGCT (Franco-Zorrilla *et al.* 2014). Interestingly, DAZ1muZnF1 shows nuclear-specific localisation in sperm cells (Figure 3.28.B, C), however it remains unclear how mutation of ZnF1 allows tighter nuclear localisation of DAZ1 similar to DAZ2-mCherry. ZnF2 and ZnF3 domains were shown to influence DAZ1 activity but were not critical for function (Figure 3.28.A, D, E). Similarly, it was shown that the *S. lycopersicum* DAZ1 orthologue lacking ZnF3 also showed reduced function *in planta* compared with AtDAZ1 (Sari, 2014). Overall, the analysis of a complete set of ZnF domains highlights their relative importance and indicates their impact on DAZ1 localisation in sperm cells.

Chapter 4

Analysis of seed abortion in *daz1 daz2* null mutants expressing DAZ1 variants

4.1 Abstract

Background: In flowering plants haploid generations develops within the sporophytic tissues of the ovule. After fertilisation the maternal seed coat develops in a coordinated manner with establishment of the embryo and endosperm (Brownfield *et al.* 2009). In *Arabidopsis thaliana* DAZ1/ DAZ2 are required for germ cell division, sperm cell differentiation and fertility. However, *daz1 daz2* null plants expressing DAZ1ΔEAR1,2 show reduced fertility and seed abortion, which was not predicted (Rutley, 2015). Therefore, this raised the question of whether DAZ1 has any functional role after fertilisation, or whether failed pre- or post-fertilisation events arising from partial DAZ1 function might lead to seed abortion?.

Results: The reproducibility of the seed abortion phenotype was examined by screening for *daz1 daz2* null lines expressing DAZ1smEAR1,2 (mutant EAR1,2 motif), DAZ1ΔEAR1, 2 or DAZ1muZnF2. Several independent *daz1 daz2* null lines were isolated in T2 populations for all three DAZ1 variant proteins. Compared with wild type Col-0 plants which showed uniform and full seed set, *daz1 daz2* null lines expressing DAZ1smEAR1,2, DAZ1ΔEAR1,2 or DAZ1muZnF2 showed fewer normal seeds, and more aborted seeds and failed ovules in siliques. In *daz1 daz2* mutants expressing DAZ1 variants some ovules enlarge but abort and were found not to develop embryos or endosperm.

Conclusion: The seed or ovule abortion phenomenon was shown to be not specific to particular DAZ1 protein domains and rather reflects impaired DAZ1 protein function and possibly sperm cell differentiation. It is speculated that incomplete sperm cell differentiation may result in early post-fertilisation defects resulting in failure of zygote and endosperm division, but stimulation of ovule enlargement.

4.2 Introduction

Seed and fruit are the key yield components for living organisms and for food security. Seeds and fruit develop soon after fertilisation and this process determines seed and fruit numbers, their final size hence, yield potential. Double fertilisation is very sensitive to abiotic and biotic stresses, which can affect mitotic phases as essential genes are required for gametogenesis, seed development, or post-embryonic growth and development in flowering plants (Tzafrir *et al.* 2004). The seed is the unit of dispersal in angiosperms and arises from a double fertilisation event, which gives rise to the triploid endosperm and zygotic embryo (Zuber *et al.* 2012). As the embryo and the endosperm develop, the ovule enlarges including the maternal tissues of the inner and outer integuments surrounding the embryo sac which form the seed coat (Chaudhury *et al.* 1997). In past experimental results shows that pollen tube overgrowth has also been observed in the absence of the LORELEI function (Capron *et al.* 2008) in self-fertilised absence of mutual consent mutants (Boisson-Dernier *et al.* 2008) and in scylla (syl) mutant embryo sacs (Rotman *et al.* 2008). However, interestingly at low frequency syl/SYL heterozygous plants shows proliferation of the central cell nucleus in the absence of fertilisation, indicating that the pollen tube overgrowth phenotype may also depend on the some central cell functions (Rotman *et al.* 2008, Matias-Hernandez *et al.* 2010). Additionally, haploid cells are not only responsible for the production and secretion of a signal that guides the pollen tubes towards the embryo sac (Higashiyama *et al.* 2001, Okuda *et al.* 2009, Tsukamoto *et al.* 2010) but, also mediate pollen tube reception (Huck *et al.* 2003, Rotman *et al.* 2003, Escobar-Restrepo *et al.* 2007). Once pollen tube correctly reach the micropyle, synergid-specific expression of the FERONIA (FER) receptor-like kinase is required for the tube growth arrest, rupture and sperm cell discharge (Huck *et al.* 2003, Escobar-Restrepo *et al.* 2007). Furthermore, example in fern mutants, the pollen tubes fails to arrest and keep growing within the embryo sac, leading to pollen tube overgrowth (Huck *et al.* 2003).

The process of transition from ovule and ovary to seed and fruit, respectively, are characterised by extensive cell division and coordinated development of the maternal and filial tissues (Barnabás *et al.* 2008, Ruan *et al.* 2010, Ruan *et al.* 2012). Newly developed seeds and fruits undergo cell expansion and accumulation of storage products, such as proteins, starch and oils, which are typical features of growth and maturation stages (Weber *et al.* 2005, Ruan *et al.* 2012). The molecular mechanisms activating reproductive development are largely unknown and the interactions between the various seed tissues remain complex and unresolved (Grossniklaus *et al.* 2001). In 2003 Claudia's lab demonstrate that phenotype of the *fis* mutant is their ability to initiate endosperm development in the absence of fertilisation (Köhler *et al.* 2003). Furthermore, the *FIS-class* genes are all expressed before fertilisation, and most likely genes involving in endosperm development (Vielle-Calzada *et al.* 1999, Luo *et al.* 2000, Spillane *et al.* 2000).

Previous work has shown that truncation of DAZ1 (DAZ1ΔEAR1,2) leads to reduced fertility and has a severe effect on male transmission (Borg et al., 2014). Moreover, the expression of the C-terminally truncated DAZ1ΔEAR1,2 variant in a *daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$ null background gave rise to siliques containing a few viable seeds and more aborted seeds. This raised the question of whether DAZ1 has any post-fertilisation role or whether failed fertilisation events arising from partial DAZ1 function may be involved in ovule abortion. Mutant *DAZ1* proteins are investigated here to explore the origin of seed abortion phenotypes arising due to defective or partial DAZ1 protein function.

4.3 Results

4.3.1 Screening for homozygous DAZ1 variant plants in the T3 generation

T2 progeny of *daz1-1*$^{-/-}$, *daz2-1*$^{+/-}$ plants expressing the DAZ1ΔEAR1,2-mCherry; DAZ1smEAR1,2-mCherry or DAZ1muZnF2-mCherry were screened for reduced fertility based on silique length. Individuals that were homozygous for *daz1-1 daz2-1* were identified that showed high frequencies of BCP. Two independent T3 lines were selected to study seed abortion phenomena (Table 4.1). DAZ1ΔEAR1,2, and DAZ1smEAR1,2 lines showed a low percentage of TCP of 14 % and 21% respectively (Figure 4.2). However, DAZ1muZnF2-mCherry showed greater rescue of germ cell division with 66 % to 68 % TCP and 31 % to 33 % BCP.

Table 4.1. DAZ1 variant lines selected in T2 and T3 generations.

DAZ1 variants	T2 generation	T3 generation
AtDAZ1-mCherry	D2	D5.A3
DAZ1 ΔEAR1,2-mCherry	C1/C2	AA5
DAZ1 ΔEAR1,2-mCherry	C1/C2	AB1
DAZ1smEAR1,2-mCherry	A8	AA2
DAZ1smEAR1,2-mCherry	A8	CA2
DAZ1muZnF2-mCherry	A6	AA1
DAZ1muZnF2-mCherry	A6	BD2

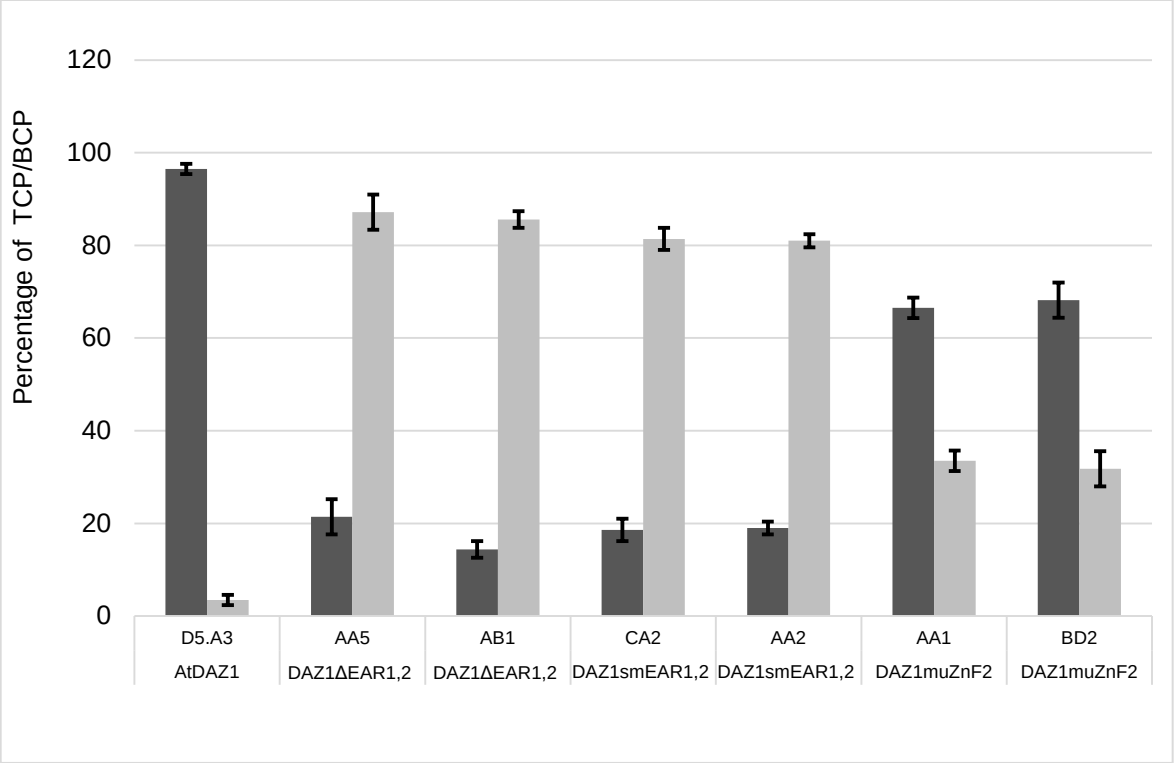


Figure 4.1. Pollen phenotypes of T3 *daz1 daz2* null plants expressing DAZ1ΔEAR1,2, DAZ1smEAR1,2, and DAZ1muZnF2. The mean percentage TCP and BCP are shown (n = 4 plants).

4.3.2 Silique measurements for T3 generation DAZ1 variant plants

Examination of siliques for T3 generation plants revealed that all DAZ1 variants showed shorter siliques compared with WT Col-0 plants. A total of 5 siliques were measured for WT and for each variant. Wild-type plants had siliques had an average length of 17 mm, while the data show similar average silique lengths of 8-9 mm for plants expressing any the three DAZ1 variants. The similar silique length of DAZ1muZnF2 plants was not predicted because lines with this DAZ1 variant shows a higher percentage of TCP than the other two variants (Figure 4.2). However, this observation is consistent with the low seed number per silique for DAZ1muZnF2 lines (Figure 4.4), which further suggests that *daz1-1 daz2-1* sperm cells expressing DAZ1muZnF2 are more likely to fail to produce viable seed.

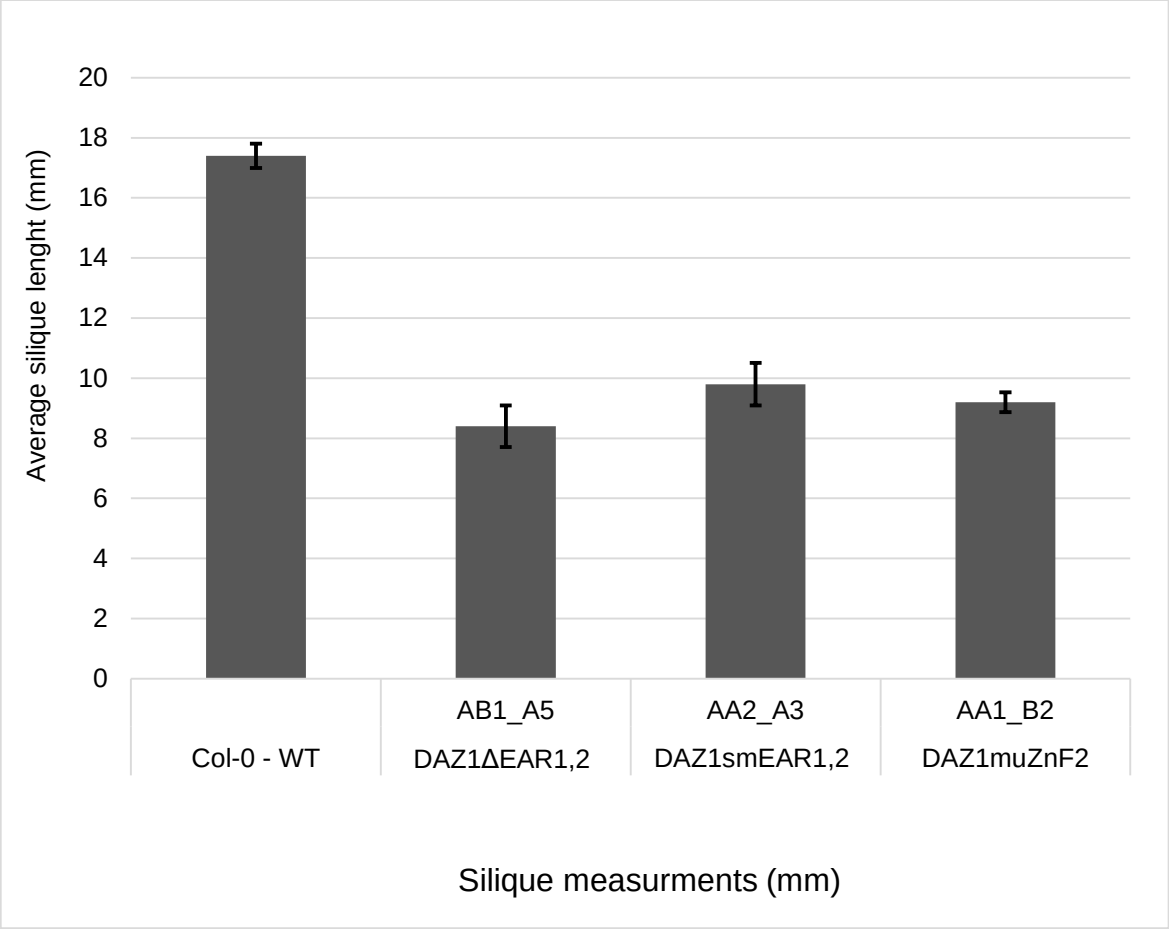


Figure 4.2. Silique length for WT, DAZ1ΔEAR1,2, DAZ1smEAR1,2, and DAZ1muZnF2. Average silique length is shown in mm (n = 5 siliques).

4.3.3 Aborted seed in T3 generation DAZ1 variant plants

To understand the origin or cause of seed abortion phenotypes, mature siliques were opened and seeds were categorised into (1) normal seed, (2) aborted seed, and (3) failed ovules (white stubs). The data in Figure 4.3 show that the average number of normal seeds is low for all three DAZ1 variants, ranging from 4 to 7 seeds per silique. However, all DAZ1 variant lines showed higher numbers of aborted seed, ranging from 8 to 15 per silique. Col-0 showed an average of 56 normal seed per silique with rare cases (<1%) of aborted seeds.

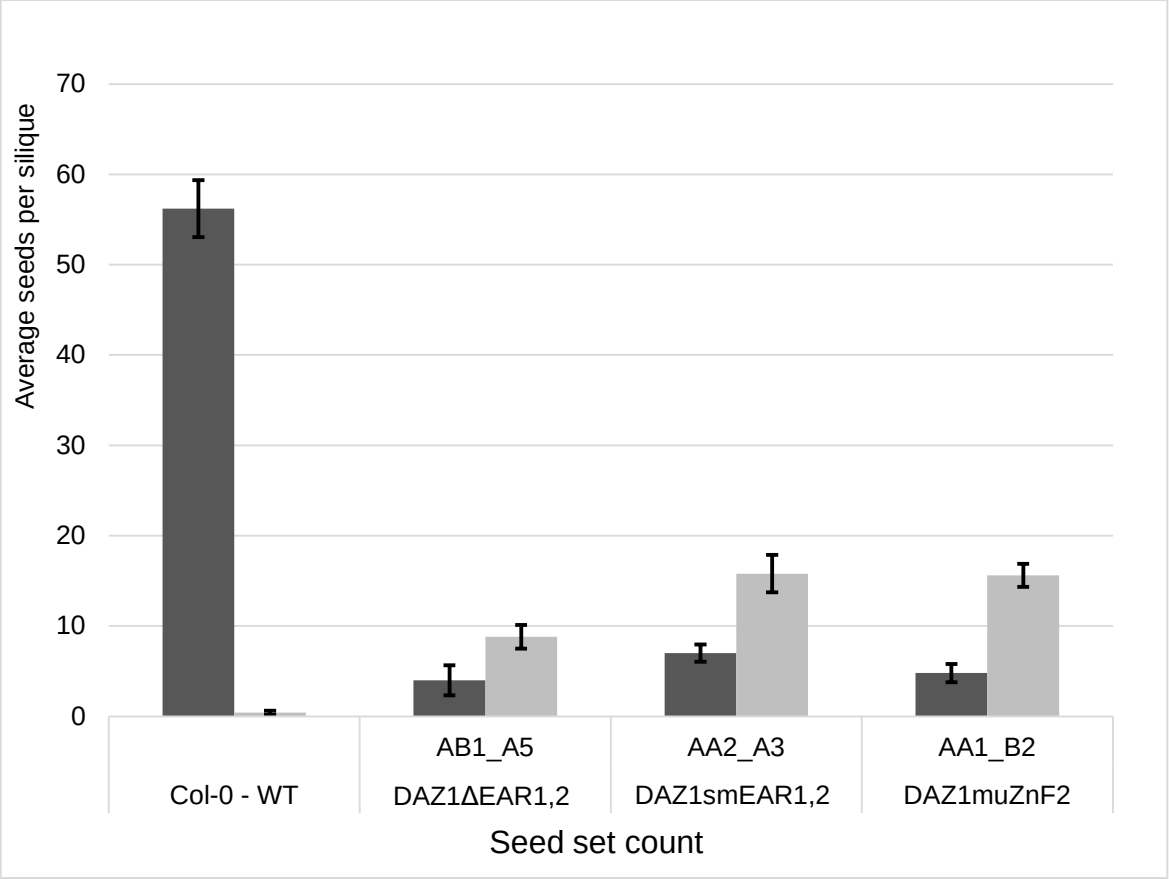
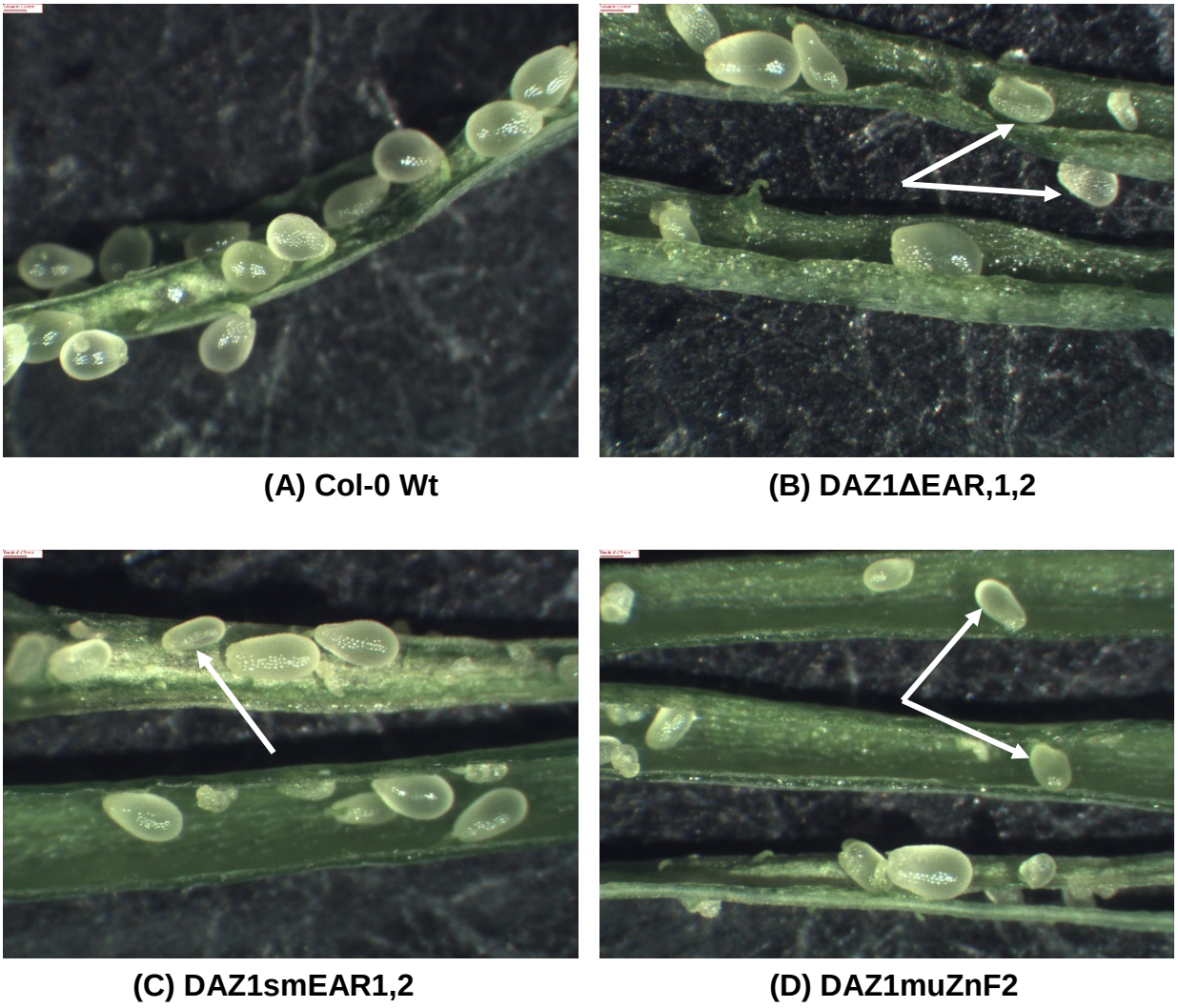


Figure 4.3. Seed set for WT and DAZ1ΔEAR1,2, DAZ1smEAR1,2, and DAZ1muZnF2. The average number of seeds in mature siliques is shown for each DAZ1 variant (Dark bar = Normal seeds, Light bar = aborted seeds), (n = 5 siliques).

4.3.4 Seed development in WT and DAZ1 variant plants

Observations were made of seeds in developing siliques of Col-0 WT and *daz1-1 daz2-1* mutants expressing DAZ1ΔEAR1,2, DAZ1smEAR1,2, or DAZ1muZnF2. In Col-0 developing seeds are uniform and large compared with DAZ1 mutant lines, which contained many smaller under developed seeds (Figure 4.5 A-D). Siliques from DAZ1 variant lines also showed very few seeds of normal size, consistent with observations of shorter siliques and reduced seed number in mature siliques (Figures 4.3 and 4.4).

Overall, the study of ovule development in *daz1 daz2* null lines expressing DAZ1 variants indicates that the seed abortion phenomena initially observed in DAZ1ΔEAR1,2 lines (Rutley, 2015) was not particular to the DAZ1 EAR domains and may therefore reflect general loss of DAZ1 function, though disturbing different protein domains.



(A) Col-0 Wt **(B) DAZ1ΔEAR,1,2**

(C) DAZ1smEAR1,2 **(D) DAZ1muZnF2**

Figure 4.4. Seed development for Col-0 WT and DAZ1 mutant variants. Images show developing seeds for (A) Col-0 WT and (B) DAZ1ΔEAR1,2, and (C) DAZ1smEAR1, 2 (D) DAZ1muZnF2. White arrows show smaller seeds (Scale bar = 0.22 mm).

4.3.5 Ovule developmental stage in DAZ1 mutant variants

Examination of siliques of T3 and T4 generations of DAZ1ΔEAR1,2, DAZ1smEAR1, 2, and DAZ1muZnF2 lines revealed aborted seeds in mature siliques (Figure 4.5B, C, D). Earlier stages of seed development were examined after dissecting out developing seed from siliques, followed by clearing and observation by differential interference contrast (DIC) microscopy. Stages from zygote to mature embryo were recorded on multiple days after pollination (DAP) by examining ovules after controlled pollination of *ms1-1* pistils (Table 4.2).

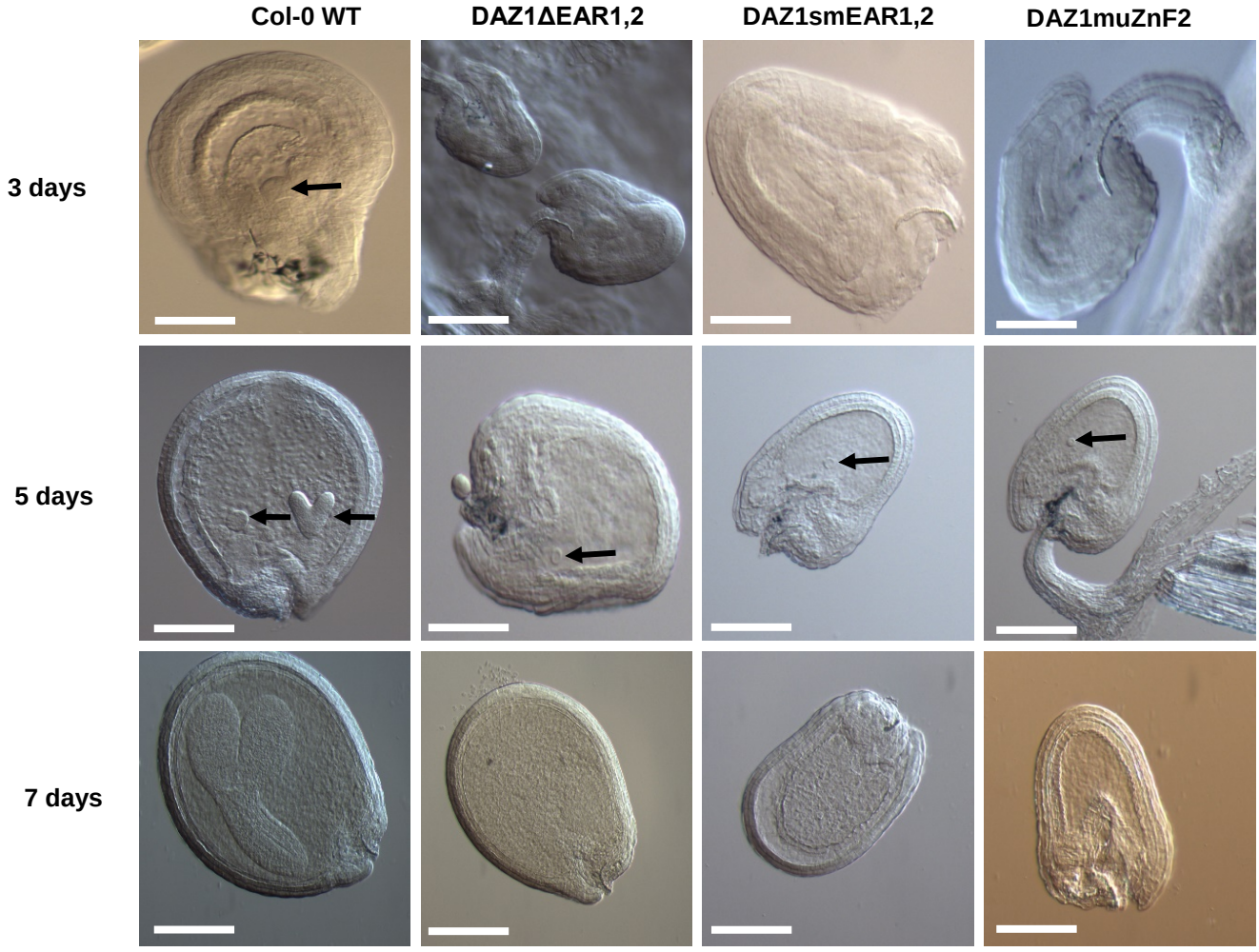
Table 4.2. **Seed development in WT and DAZ1 variant plants.** Pollen from *daz1-1 daz2-1* DAZ1 variant plants were used to pollinate *ms1-1* pistils. Ovules/developing seeds were observed at different time intervals after pollination (DAP) (n = 5 siliques).

DAP (DAYS)	Line ID	Ovule like	Zygote	Pre-globular	Globular	Tri-angular	Heart	Torpedo	Linear cotyledon	Bent cotyledon	Mature green
3	Col -0 WT	80		36							
5	Col -0 WT	105					45				
7	Col -0 WT	147							29		
9	Col -0 WT	165								97	
11	Col -0 WT	177									68
13	Col -0 WT	210									75
3	DAZ1ΔEAR1	25		4							
5	DAZ1ΔEAR1	22	3								
7	DAZ1ΔEAR1	28									
9	DAZ1ΔEAR1	20									3
11	DAZ1ΔEAR1	15									
13	DAZ1ΔEAR1	13									
3	DAZ1smEAR1,2	20		2							
5	DAZ1smEAR1,2	26					1				
7	DAZ1smEAR1,2	26		1						1	
9	DAZ1smEAR1,2	22									4
11	DAZ1smEAR1,2	20									
13	DAZ1smEAR1,2	17									

Three different class of developing seeds were observed (1) normal ovules containing bent cotyledon stage embryo, (2) abnormal ovules that show enlargement but no sign of embryo and endosperm development (3) aborted ovules which do not show any initial development of either embryo or endosperm both are completely absent in DAZ1 mutant variants.

Observations of ovules resulting from pollination of *ms1-1* pistils with DAZ1 variant plants showed significant numbers of aborted and apparently unfertilised ovules

compared with the Col-0 WT control (Table 4.2). Inspection was carried out of ovule development 3, 5, 7, 9, 11, and 13 days after pollination of *ms1-1*.



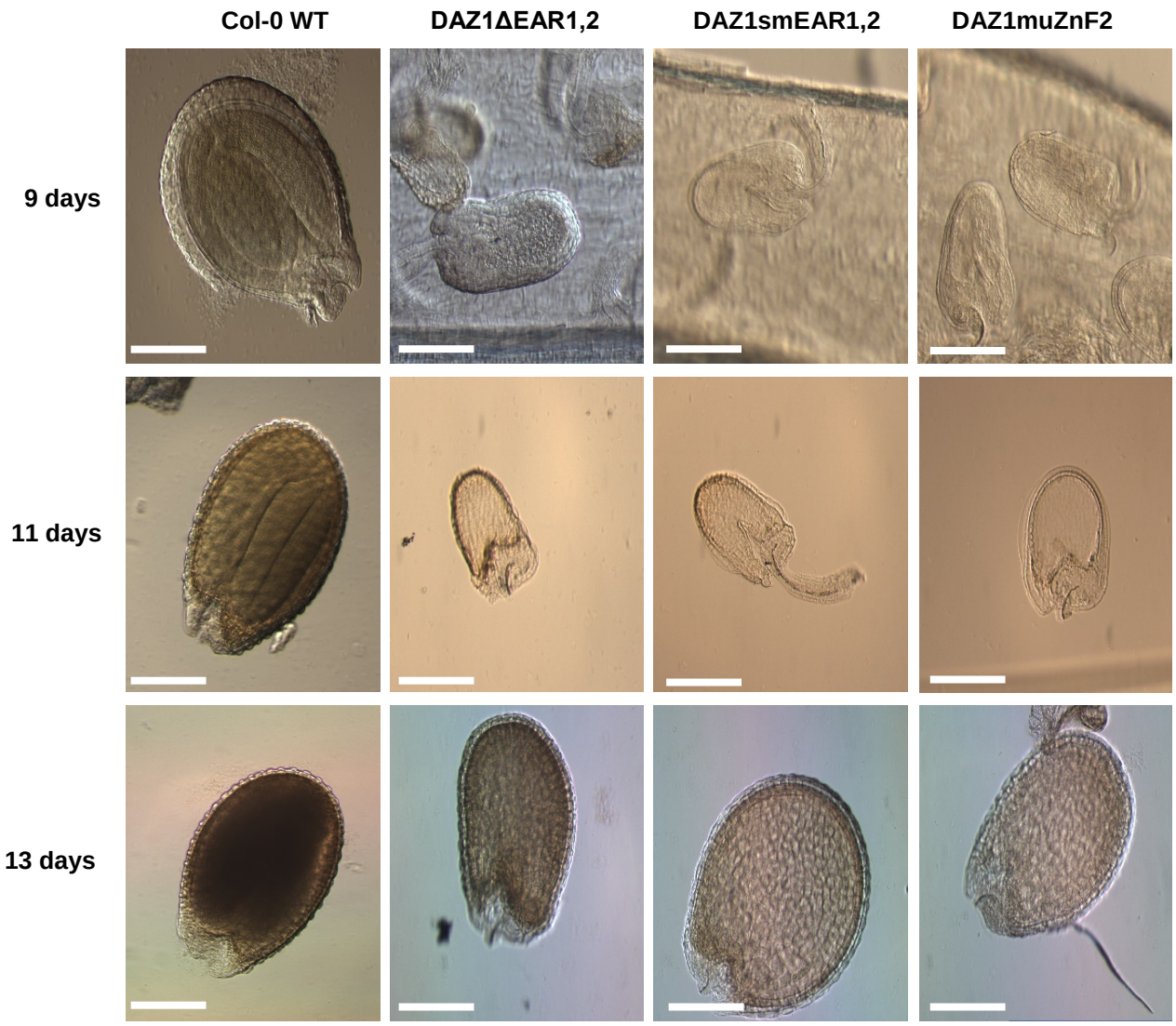


Figure 4.5. Seed development after pollination of *ms1-1* with Col-0 WT and DAZ1 variant plants. Images clearly shows proliferating endosperm and embryo at transition from zygotic stage to mature embryo stage in ovules after crossing with WT pollen (arrows). Some enlarged ovules were observed after crosses with DAZ1 variants 5-13 DAP (arrows indicate single nuclei in ovules) (n = 5 siliques; white scale bar = 100 µm).

Developmental analysis following crosses with Col-0 WT pollen revealed that embryos develop gradually from 3 to 13 DAP from pre-globular to mature embryo

stages (Figure 4.5). In ovules from crosses with pollen from DAZ1 variants plants, some enlarged ovules were observed at 5-13 DAP without any egg cell or central cell development (Figure 4.5). Some of ovules showed single larger nuclei in the ovule indicating failure of fertilisation or early post-fertilisation failure of central cell development.

4.4 Discussion

Double fertilisation, a key feature of the success of angiosperms, happens deep inside the maternal tissues and involves two different fusion events. In successful fertilisation, one sperm cell fuses with the egg cell to form the zygote to initiate the growth of the embryo, while the second sperm cell fuses with the central cell to trigger endosperm development. The final product of normal fertilisation is the seed in which the embryo develops from zygotic to mature embryo stage and involves three distinct genetically characterised components (1) embryo, (2) endosperm, and (3) seed coat (Digonnet *et al.* 1997).

The characterisation of ovule-defective mutants in *Arabidopsis thaliana* has shown that female gametophyte development is dependent on the sporophyte, specifically integument development. This has led to the identification of a variety of factors that contribute to this relationship including transcription factors, kinases and components of plant hormones signal transduction pathways (Bencivenga *et al.* 2011, Shi and Yang 2011, Vielle-Calzada *et al.* 2012).

Male gametophyte development is also dependent on the maternal sporophyte because pollen grains interact with the stigma, style and transmitting tract during pollen tube (Leydon *et al.* 2014). Furthermore, the initial stimulation of growth of ovules following pollination with fertilisation defective pollen tubes suggest that there is a paternal effect on ovule development. This phenomenon has been termed POEM (pollen tube-dependent ovule enlargement morphology) and occurs only

when the ovules accepts the pollen tube content (PTC) facilitating the initial development of the ovules without fertilisation (Kasahara *et al.* 2017).

4.4.1 Seed abortion

The large-scale abortion of flowers and immature fruits is a common phenomenon in plants. The proportion of flowers that develop into mature fruit varies considerably among species, ranging from less than one percent to one hundred percent (Lloyd 1980, Lloyd *et al.* 1980, Stephenson 1984). In addition, molecular studies with Arabidopsis and other plants have identified the cis-control regions of several genes active during seed development, particularly those encoding storage proteins, and the transcription factors (TEs) that play a role in their regulation (Vasil *et al.* 1995, Braybrook *et al.* 2006). Nevertheless, the identities of most regulators of seed development and their direct targets are largely unknown (Brandon H. Le *et al.* 2010). Lack of fertilisation of all ovules despite a large numbers of pollen grains deposited on the stigma is known for various plant species (Mogensen 1975, Guth and Weller 1986, Cruzan 1989). Some theoretical studies also suggest that female reproductive success not limited by pollen availability but by the resources required for fruit and seed species like *Catalpa speciose* (Stephenson 1979) in which artificial pollination failed to increase fecundity, while in a few species of *Asclepias* like *A. verticillata* (Willson and Price 1980), addition of resources increased fecundity (Arathi *et al.* 1999).

In 2010 Ron lab demonstrated that two different phenotypes accounts for the reduced seed set in siliques of self-pollinated *kpl-1/+* and *kpl-1/kpl-1* plants: undeveloped ovules and aborted seeds. Furthermore, those class characterise the undeveloped ovules, from 1 to 2 DAP pistils of self-pollinated *kpl-1/kpl-1* plants. Additionally, observed that ovules that had not been fertilised and ovules that had been fertilised both were developing correctly, with a developing embryo and proliferation endosperm (Ron *et al.* 2010). However, also observed ovules which only one fertilisation event had occurred: either the egg or embryo developing but,

the central cell was unfertilised, or the central cell had been fertilisation and endosperm nuclear divisions were evident, but the egg cell was unfertilised therefore, ovules form the *kpl-2/kpl-2* mutant displayed similar phenotypes (Ron *et al.* 2010). Recently, the identification in *Arabidopsis thaliana* of the receptor-like kinase FERONIA has shed some light on optional signalling mechanisms responsible for loses of ability of the embryo sac to attract pollen tubes after-fertilisation initiation process. Furthermore, in Feronia and sirene mutants, the pollen tube can reach the embryo sac, but, instead of arresting and delivering the two sperm cells, the pollen tube does not arrest and continues to grow and invades the mutants (Boisson-Dernier *et al.* 2008) displays an absence of release of the sperm cells from the pollen tube similar to what was observed in feronia/sirene (Capron *et al.* 2008).

Another results demonstrated by takeuchi and higashiyama group in 2012, suggested that genes directly involved in male/female and host/parasite interactions are believed to be under positive selection therefore, in flowering plant *Arabidopsis thaliana* has more than 300 defendin-like (DEFL) genes, which are likely to be involved in both natural immunity and cell to cell communication including pollen-pistil interaction. Thus, in mutant DAZ1 variants (DAZ1smEAR1,2, DAZ1ΔEAR1,2 and DAZ1muZnF2) loss the interaction between the pollen and pistil to deliver the function sperm cells. However, little information of theses relationship between the molecular evolution of DEFL genes and their functions (Takeuchi and Higashiyama 2012). Result identified a cluster of DEFL genes cluster in *A. thaliana* and DEFL (Cysteine-rich peptide [CRP810_1]) peptides, which named as a AtLURE1 peptides, are pollen tube attractants guiding pollen tube to the ovular micropyle. In mutant DAZ1 variants (DAZ1smEAR1,2, DAZ1ΔEAR1,2 and DAZ1muZnF2) might be AtLURE1 peptides influence interrupted by mutation in DAZ1 and not able to guild pollen tube towards the micropyle (Takeuchi and Higashiyama 2012).

4.4.2 Is incomplete gamete differentiation the cause of seed abortion?

The investigations described show that *daz1-1 daz2-1* mutants which express different DAZ1 variants with reduced function, develop a proportion of pollen with two sperm cells. It is likely that these twin sperm cells are delivered to the embryo sac, since the single undifferentiated germ cell present in *daz1-1 daz2-1* pollen was shown to be released into the embryo sac by labelling germ cells with fluorescent protein markers (Borg *et al.* 2014).

Because, differentiation of the embryo sac occurs contemporarily and in coordination with the development of the diploid sporophytic tissue of the ovule. Megasporogenesis takes place in the nucellus when integument primordia elongate from the chalazal region (Matias-Hernandez *et al.* 2010). However, due to truncation of the EAR domain, or as a result of mutations in the EAR motifs or in ZnF2, sperm cell differentiation and fitness may be compromised. It is therefore suggested that twin sperm cells which express DAZ1 variants are delivered, but due to reduced DAZ1 function are incompletely differentiated. It is suggested that some genes required for fertilisation may not be optimally expressed such as *GEX2* and *CGS1/HAP2* that are known to be required for gamete adhesion and fusion during fertilisation and are known to be regulated by the DUO1-DAZ1 network (Borg *et al.* 2014).This could lead to single fertilisation events or to a lack of embryo or endosperm development due to incomplete male gamete differentiation. More detailed observations, for example using marked *daz1 daz2* mutant gametes, are needed to determine whether single or defective double fertilisation events occur and how these impact on initiation of zygote and central cell development.

The last phase of the pollen tube (PT) growth and guidance is controlled by the female gametophyte in angiosperms and its requires species-preferential chemotropic guidance molecules (Shimizu and Okada 2000, Higashiyama *et al.* 2001, Lausser *et al.* 2009, Márton and Dresselhaus 2010). In maize ZmEA1 (Zea mays Egg Apparatus 1) was the first identified signalling molecules accomplishing

the properties of a female-gametophyte-derived pollen tube attractant. ZmEA1 is an intronless single gene specifically expressed in the egg apparatus (egg and synergid cells) of maize and encodes a hydrophobic precursor protein of 94 amino acids (of which 47 are hydrophobic) with a predicted N-terminal transmembrane domain. Data shows that ZmEA1 is a member of a novel class of polymorphic small proteins and it is secreted from the egg apparatus towards the cell walls of micropylar nuclear cells, and it is knockdown impairs micropylar pollen tube guidance in maize (Márton *et al.* 2012).

Chapter 5

Genetic screening for suppressors of reduced DAZ1 function in *A. thaliana*

5.1 Summary

Background: A C-terminal truncation of DAZ1 which removes the EAR motif domain has a severe, but non-lethal effect on DAZ1 function in *Arabidopsis* (Borg *et al,* 2014). The limited germ cell division and male transmission of *daz1-1 daz2-1* pollen expressing this protein (DAZ1ΔEAR1,2-mCherry) enabled the isolation of homozygous *daz1-1 daz2-1* plants which showed highly reduced fertility (Rutley 2015). These plants were used to screen for genetic suppressors of reduced fertility in the M2 generation following EMS-mutagenesis.

Results: An M2 population of 5000 *daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$, DAZ1ΔEAR1,2-mCherry plants was screened for restored fertility. M2 plants showing partially restored fertility were identified in six independent M2 seed pools and were designated *dazRF* lines. Silique measurements and seed set analysis confirmed fertility restoration in M3, M4 and M5 *dazRF* lines. Five M3, M4 and M5 *dazRF* lines which showed stable restoration of fertility were analysed to confirm their parental genotypes. Stable M5 generation *dazRF* lines were backcrossed to the parental line to remove unlinked mutations. The pollen phenotypes of stable fertile F1 and F2 back-crossed lines restored germ cell division and fertility above that of parental lines.

Conclusion: The isolation of EMS-induced *dazRF* mutants which restored the fertility of daz1$^{-/-}$ *daz2*$^{-/-}$ DAZ1ΔEAR1,2-mCherry plants suggest that the putative mutations can compensate for partial DAZ1 function. Further genetic characterisation and identification of the mutations responsible will help to decipher the loci responsible.

5.2 Introduction

A great challenge for plant breeding is always the collection or development of novel germplasm resources. Established methods of generating novel diversity in germplasm includes chemically-induced mutations which can result in point missense or nonsense substitutions in hundreds of genes in individual lines (Wang *et al.* 2014). In typical mutagenesis experiments EMS-treated seeds are grown to form an M1 population, which is self-fertilised to generate M2 progeny which are then screened for novel phenotypes (Serrat *et al.* 2014). EMS mutagenesis generates randomly distributed point mutations throughout the genome via chemical modification of nucleotides, leading to C to T changes resulting in C/G to T/A substitutions (Kim *et al.* 2006). EMS-mutagenesis has also been applied to suspension-cultured cells for screen of somatic cell phenotypes (Chen *et al.* 2013). In crop plant studies alternative mutagens such as N-methyl-N-nitrosourea have also been used, for example to treat fertilised rice spikelets resulting in high mutagenesis rates (Suzuki *et al.* 2008).

The goal of the work presented was to perform a forward genetic screen of EMS-mutagenised Arabidopsis plants to identify novel suppressors which restore the fertility of plants with compromised DAZ1 function. The rationale is that mutants which restore DAZ1 function may uncover novel mutations that might directly or indirectly influence the DAZ1 protein or the DUO1-DAZ1 network controlling germ cell development.

5.3 Results

5.3.1 Analysis of fertility restored EMS-derived M2 generations

Triple homozygous *daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$ *T*$^{+/+}$ plants expressing the proDAZ1:DAZ1ΔEAR1,2-mCherry transgene (T^{+}) showed highly reduced fertility with few viable seeds, aborted seeds and a larger number of failed or unfertilised ovules. M0 seeds were bulked up and mutagenised by EMS treatment as described

(Rutley 2015). The parental triple homozygous *daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$ *T*$^{+/+}$ M1 plants were grown in 13 separate batches and allowed to self to yield 13 M2 seed pools (A to M). From each M2 seed pool, around 150 M2 plants were phenotypically screened for restored fertility by observation for long and fertile siliques (Figure 5.1), followed by examination of pollen phenotypes by DAPI staining of pollen by fluorescence microscopy. The number and phenotypes of seeds present in siliques from plants showing restored fertility allowed seed number to be scored and classified into three categories: (1) normal, viable seeds, (2) aborted seeds, and (3) white stubs (failed or unfertilised ovules).

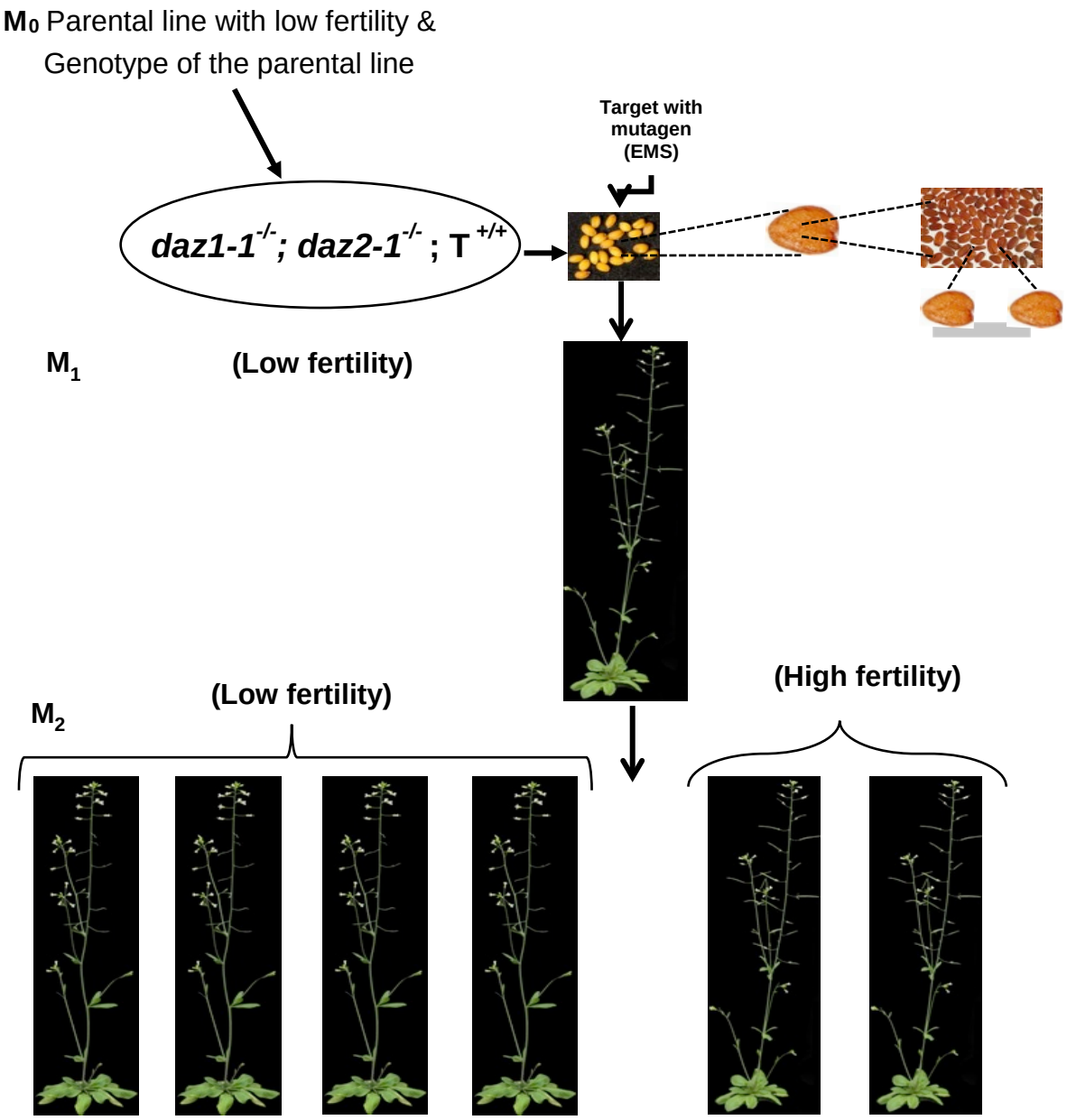


Figure 5.1. Diagram of EMS mutagenesis screen. Mo seed of triple homozygous *daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$ *T*$^{+/+}$ plants were treated with EMS, the M1 grown and M2 seed harvested from 13 M1 pools (M2 seed pools). M2 pools were screened for restoration of fertility.

5.3.2 Phenotypic and genotypic analysis of the *dazRF*- M2 generation

Phenotypic analysis of pollen carried out by using fluorescence microscopy was used to examine the rescue of germ cell division (DAPI-staining) and transgene expression (RFP signal) in pollen in lines showing restored fertility. Out of 13 *dazRF* M2 pools, six gave rise to individual plants with increased fertility relative to the parental line grown alongside as a control. Fertility restoration was observed in *dazRF* M2 pools C, E, F, I, J and K. The average fertility of individuals from each pool is shown in Figure, 5.2.

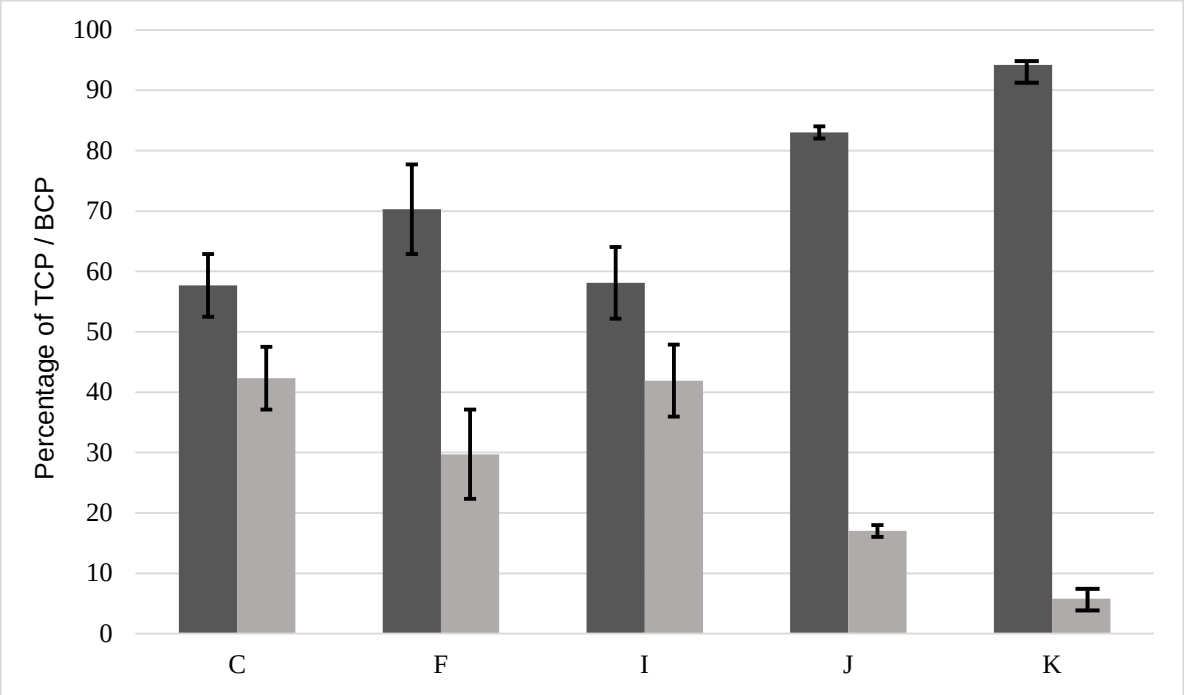


Figure 5.2. Phenotypic analysis of *dazRF* M2 lines showing fertility restoration. The chart shows percentage of TCP (dark bar) and BCP (light bar) for *dazRF* individuals from M2 seed pools with restored fertility. Additionally, parental lines shows 9.3% TCP and 90.87% BCP respectively. Error bars represent SE (n = 4 counts).

5.3.3 Silique measurements and seed set in *dazRF* M2 lines

Measurements of silique length were carried by dissecting microscopy for *dazRF* M2 lines. Mature siliques were opened to examine seed development and some had very low numbers of seeds. Measurements of silique lengths were carried out by comparing *dazRF* lines with M2 parental line background plants as negative controls and Col-0 WT plants as positive controls. *dazRF* plants showed comparatively long siliques (10.9 to 14 mm) which were shorter than Col-0 WT (17.5mm) , but longer than parental line negative controls (8.2 to 9.8 mm). The average silique length is shown for *dazRF* lines from different M2 pools (Figure 5.3).

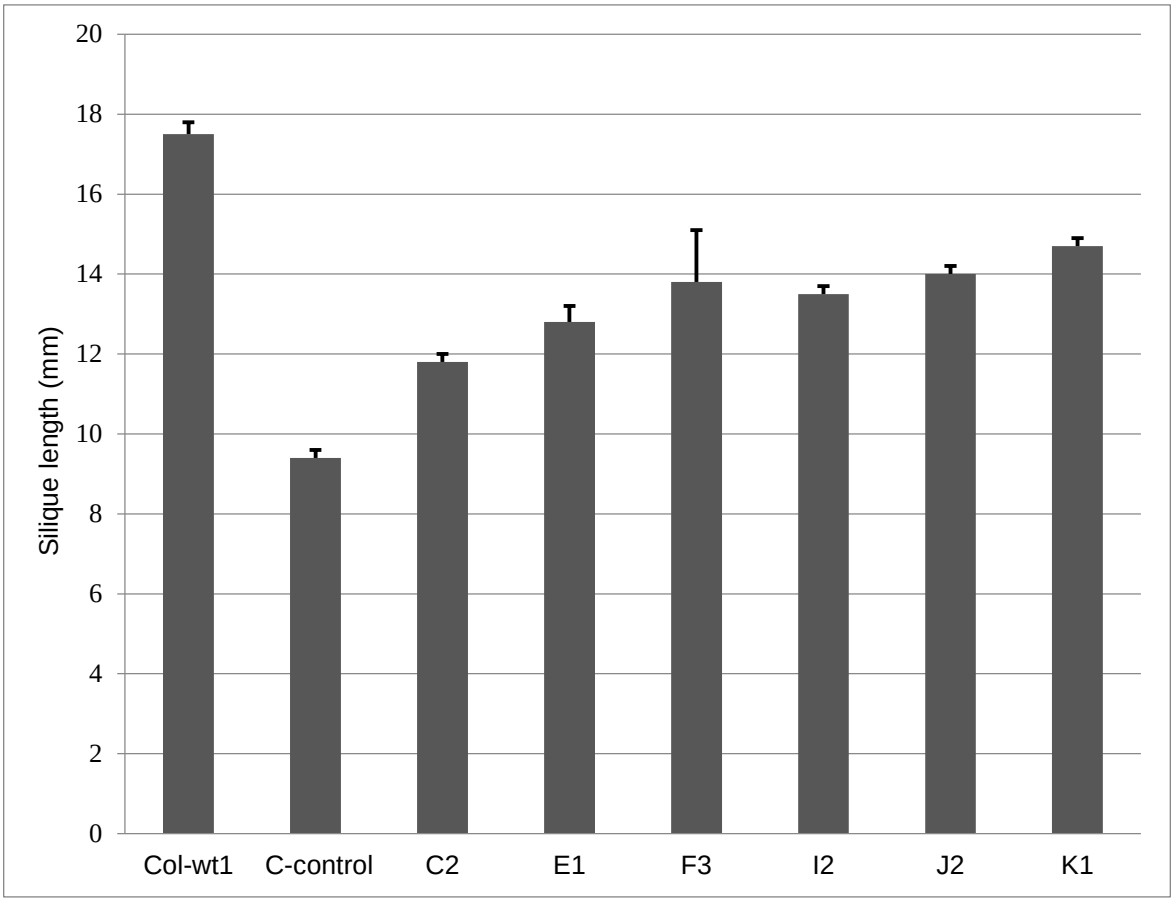


Figure 5.3. Measurements of silique length in *dazRF* M2 lines. Silique length of *dazRF* M2 lines compared with parental negative C-control plants and positive control wild type Col-0 plant. Error bars represent SE (n = 4 siliques).

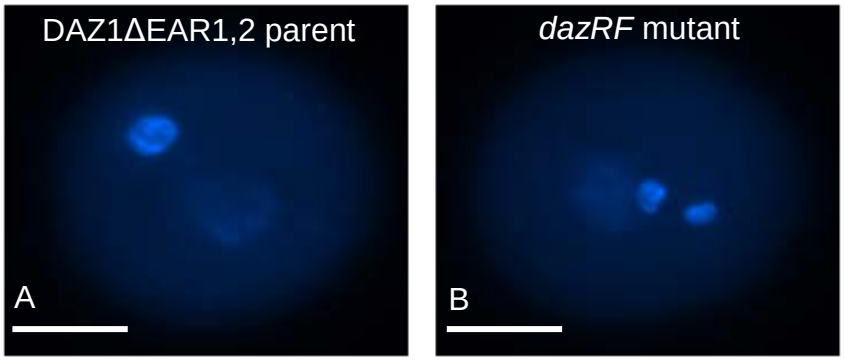


Figure 5.4. Typical pollen phenotype of the parental line and *dazRF* mutant. Images of DAPI-stained pollen show (A) bicellular pollen from DAZ1ΔEAR1,2 parent and (B) restored tricellular pollen phenotype in *dazRF* line C2.

For the analysis of seed set seed numbers were counted in individual siliques of *dazRF* M2 plants to further validate fertility restoration. Three different classes were counted: (1) normal seeds, (2) aborted seeds, and (3) white stubs (failed or undeveloped ovules but, data not included for the analysis). Col-0 control plant showed an average seed number of 64 per silique, while parental control plants from the different pools show highly reduced fertility with only low numbers (<3) of normal seeds and a greater number of aborted seeds compared with Col-0 (Figure 5.5). *dazRF* lines showed increased seed numbers (8 to 35) compared to parental control lines and showed variability between *dazRF* lines. For *dazRF* lines E, F, I, J and K the number of normal seeds was greater than the number of aborted seeds, while this was reversed in *dazRF*-C plants (Figure 5.5). This suggests that unknown factors triggered silique elongation in *dazRF-C* in which relatively low numbers (<10) of normal seeds developed.

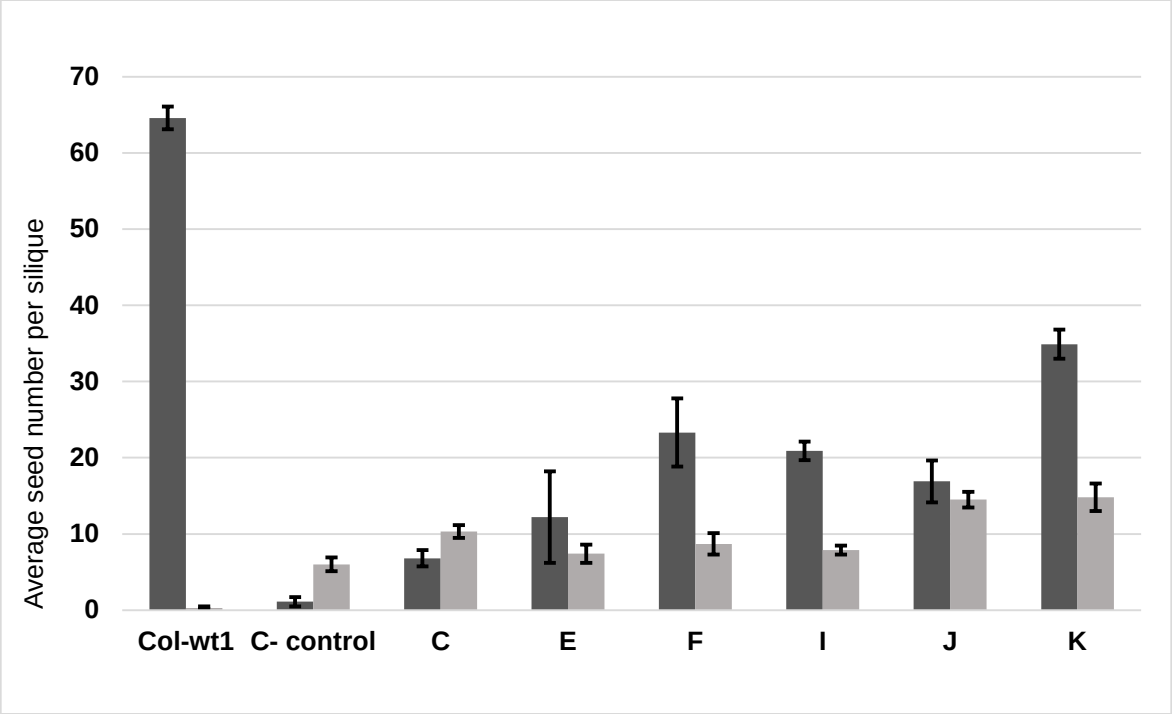


Figure 5.5. Seed set analysis in *dazRF* M2 plants. The average number of normal (dark bars) and aborted seeds (light bars) per silique in *dazRF* M2 plants are shown, compared with Col-0 positive and background (M2 pool C) negative control plants. White stub data not included for the analysis. Error bars represent SE (n = 10 siliques).

Figure 5.6. Seed set in siliques in *dazRF* M2 plants. (A) Col-0 WT shows full seed set, (B) negative control pool C silique with single normal seed and aborted seeds and white stubs, (C) *dazRF*-I2 and (D) *dazRF*-K1 both show high seed set with a few aborted seeds.

Collectively, the analysis of silique length, generative cell division and seed set suggest that putative mutations in *dazRF* lines restore or compensate for the reduced function of DAZ1ΔEAR1,2, leading to rescue of germline development and fertility. Figure 5.6 shows high numbers of normal and apparently viable seed develop in *dazRF* M2 individuals.

5.3.4 Genotypic analysis of *dazRF* M2 generation plants

Genotyping of *dazRF* lines was carried out to confirm their expected genotypes. Genomic DNA was extracted and multiplex PCR performed using gene-specific and T-DNA left border (LB) primer mixtures. Primers were designed to amplify the T-DNA left border flanking region for each allele and a region spanning the T-DNA insertion site *DAZ2* (Figure 5.7). All *dazRF* M2 lines gave a single 0.7 kb T-DNA LB product for *daz1-1* and no *DAZ1* gene-specific product and a single 0.6 kb T-DNA LB product for *daz2-1* and no *DAZ2* gene-specific product. This analysis revealed that all *dazRF* M2 lines are homozygous for *daz1-1 daz2-1* T-DNA insertion alleles (Figure 5.7).

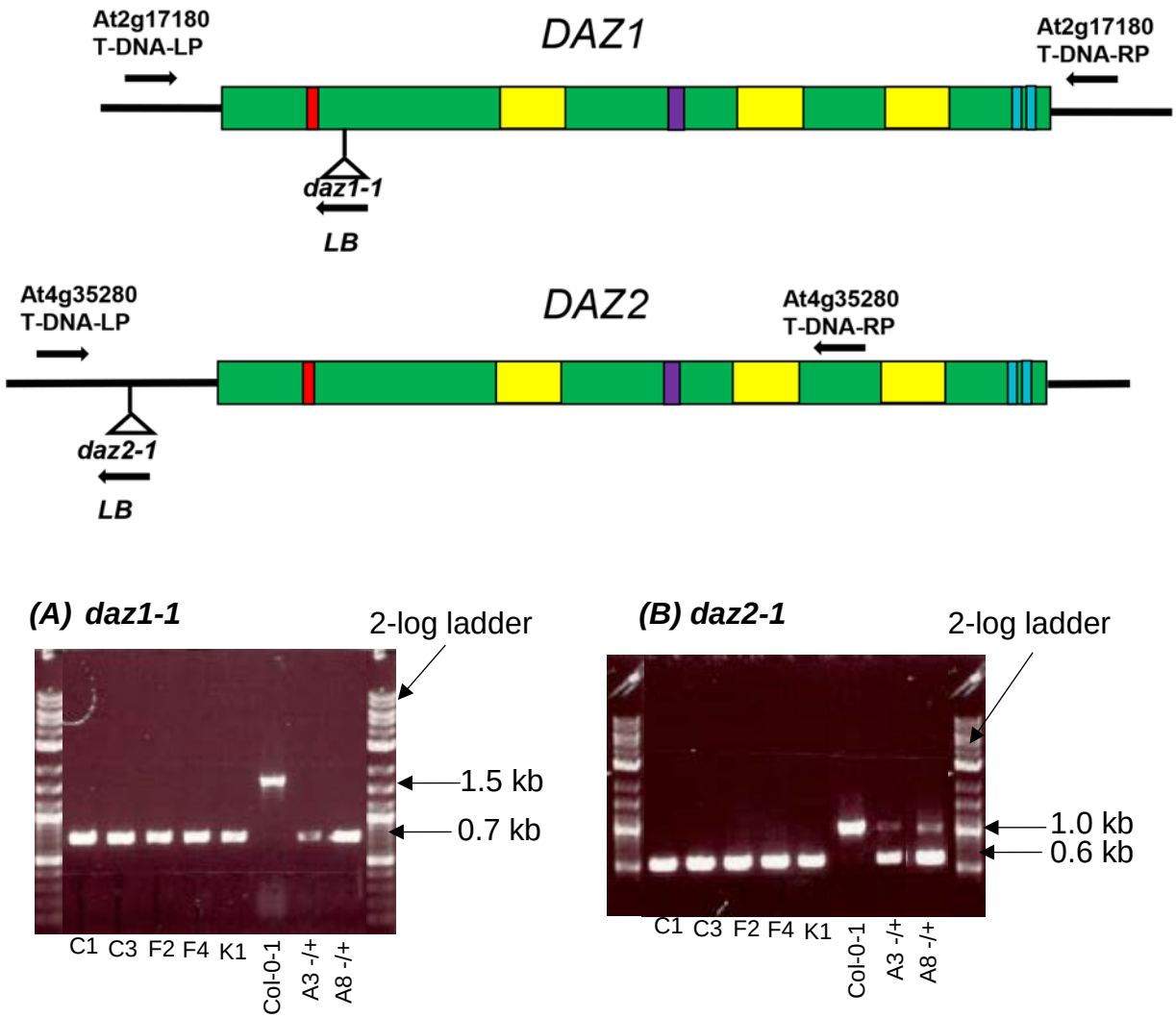


Figure 5.7. Genotypic analysis of *dazRF* M2 plants. Illustration of the DAZ1/DAZ2 proteins T-DNA left border primer binding site.Gel electrophoresis images showing PCR products of left T-DNA border (LB) primer with At2g117180 / At4g35280 (T-DNA-LP and T-DNA RP) which amplify the LB-flanking region of (A) *daz1-1* and (B) *daz2-1* alleles. Control gDNAs: Col-0 WT and *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants (A3, A8).

5.4 Analysis of *dazRF* M3 generation plants

To further characterise *dazRF* M2 families, phenotypic screening (Figure 5.9) and genotyping (Figure 5.10) was carried out to examine whether increased fertility

phenotypes were transmitted to progeny and to re-confirm the parental genotype. The pollen phenotypes of *dazRF* M2 lines isolated from all six original M1 pools (C, E, F, I, J, K) were examined in the M3 generation. The mean ratio of TCP to BCP was determined in families of M3 generation plants. Parental control plants consistently showed a low (10-15 %) percentage of TCP, whereas most *dazRF* M3 lines maintained a high (75-88 %) percentage of TCP and the I2 line, a slightly reduced (63%) percentage (Figure 5.8). Overall, the results show that *dazRF* M2 lines derived from six M2 seed pools retained fertility restoration in the M3 generation.

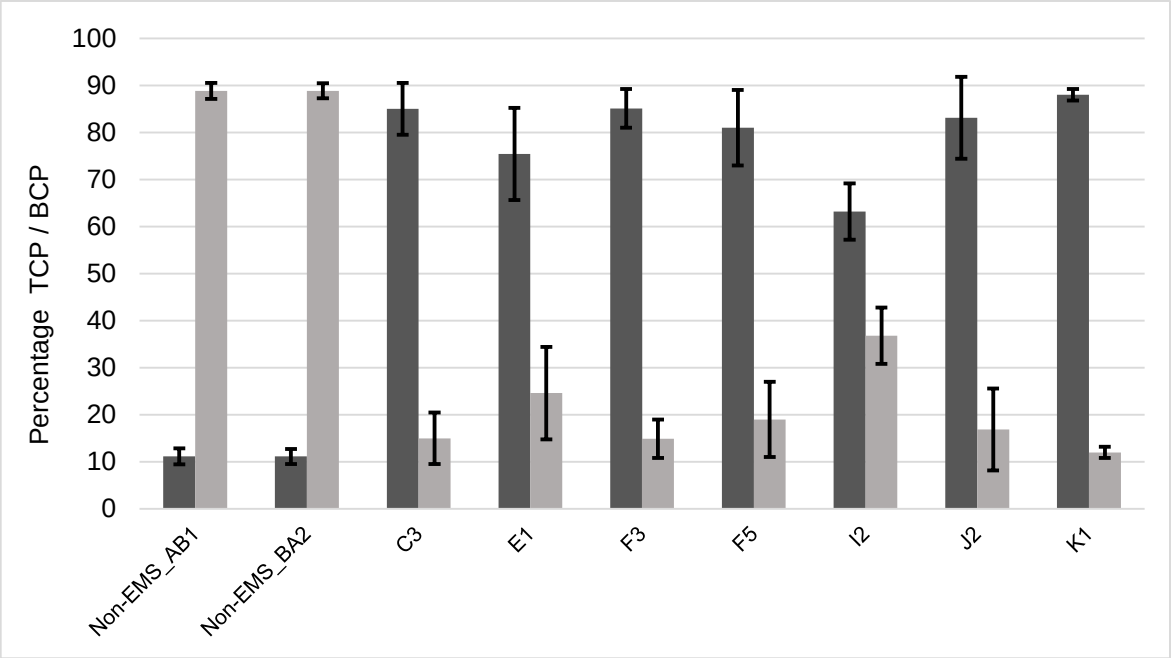


Figure 5.8. Frequency of TCP and BCP in *dazRF* M3 lines. The data show mean percentages of TCP and BCP for M3 families. Error bars represent SE (n = 4 plants).

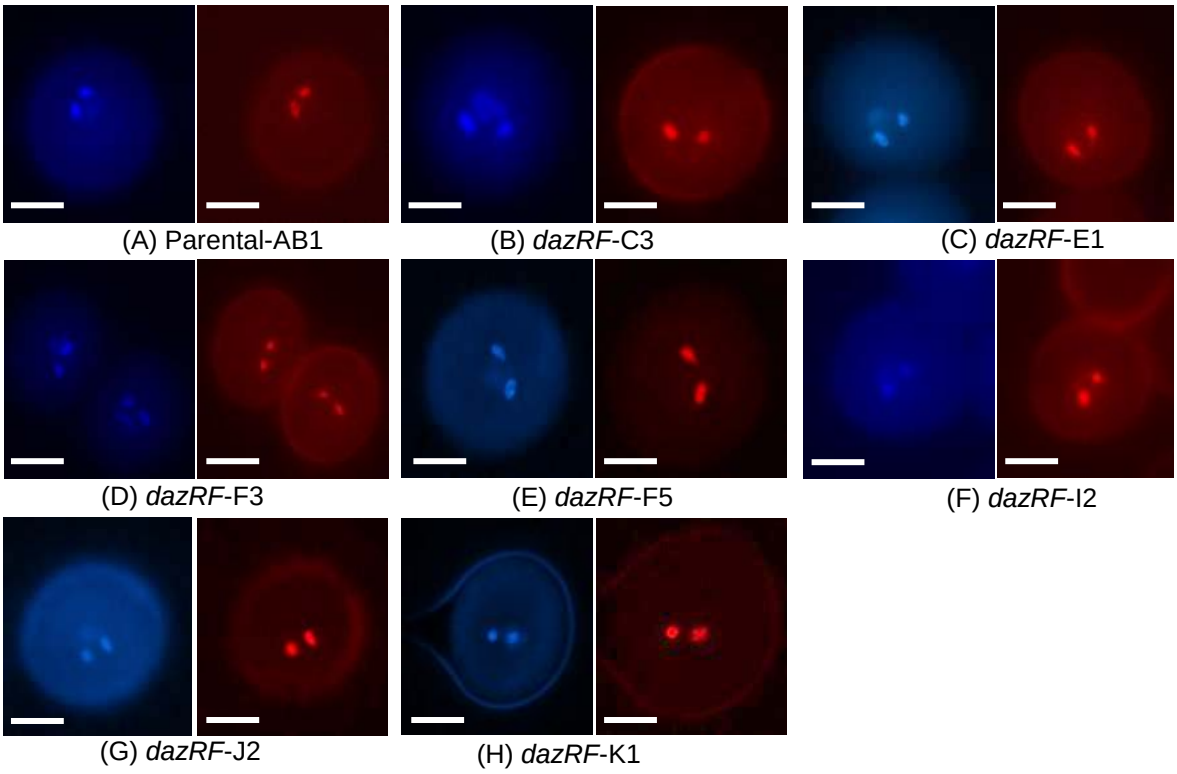


Figure 5.9. Tricellular pollen phenotypes and DAZ1ΔEAR1,2-mCherry expression of *dazRF* M3 plants. Images show DAPI and RFP signals. (A) parental-AB1 line, (B) *dazRF*-C3, (C) *dazRF*-E1, (D) *dazRF*-F3, (E) *dazRF*-F5, (F) *dazRF*-I2, (G) *dazRF*-J2, and (H) *dazRF*-K1. Scale bars = 10 µm.

5.4.1 Genotypic analysis of *dazRF* M3 generation plants

Following the isolation of *dazRF* M3 lines which showed restoration of fertility, genomic DNA was extracted from individuals originating from five M2 pools and multiplex PCR was performed to confirm their genotypes. The results confirmed the daz1-1 daz2-1 homozygous genotype for two M3 individuals for each line (Figure 5.10).

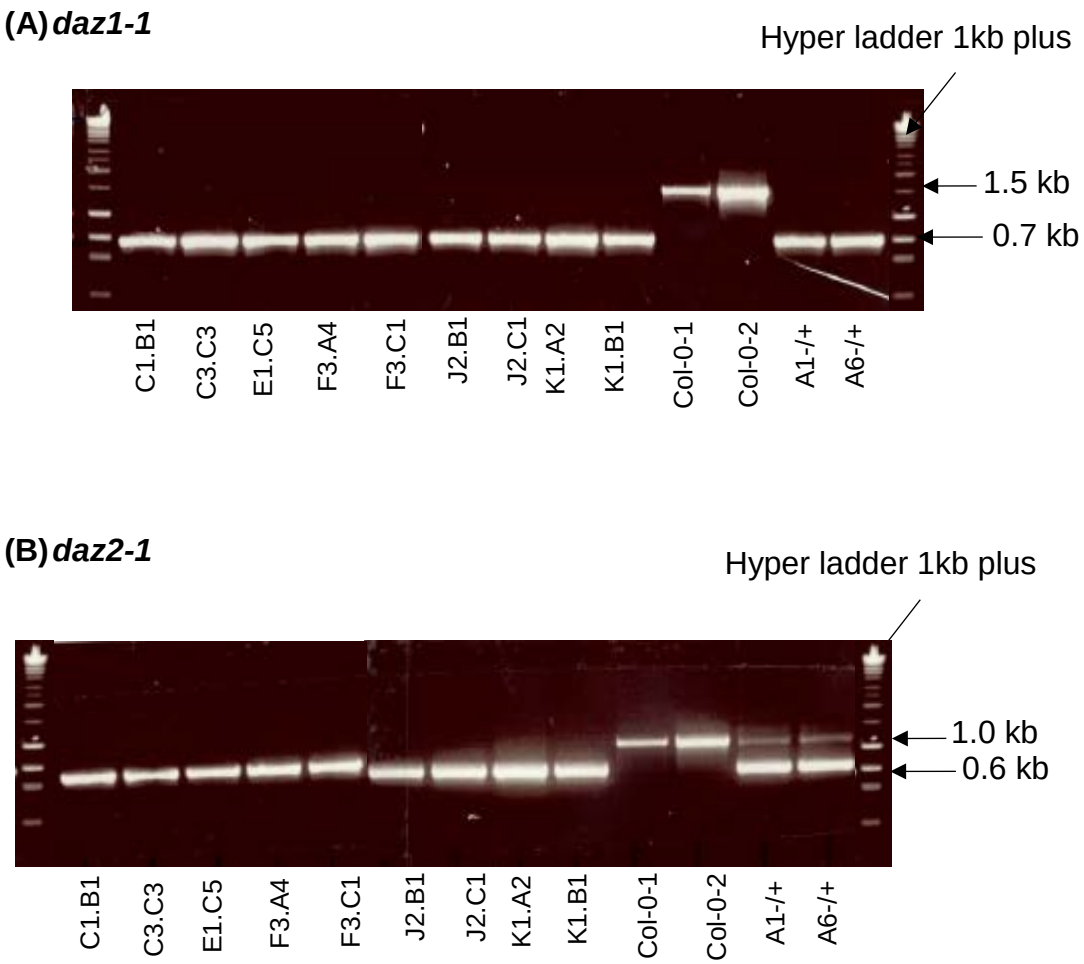


Figure 5.10. Genotypic analysis of *dazRF* M3 plants. Gel electrophoresis images showing multiplex PCR products obtained with T-DNA left border (LB) primer with gene-specific primers, which amplify (A) *daz1-1* and (B) *daz2-1* alleles. Control gDNAs: Col-0 WT and *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants (A1, A6). See Figure 5.7 for primer positions.

5.5 Stable fertility restoration in *dazRF* M4 lines

Analysis of *dazRF* M3 generation plants identified five lines originating from independent M2 pools with high fertility. Further examination of M4 families derived from M3 individuals confirmed that five lines (C3, F3, F5, I2, and K1) showed stable restoration of fertility and a high proportion of TCP (Figure 5.11). The sister lines

F3.C1 and F5.B2 showed the highest (93-95%) percentage TCP, with the F3.C3 line at 78 % TCP (Figure 5.11).

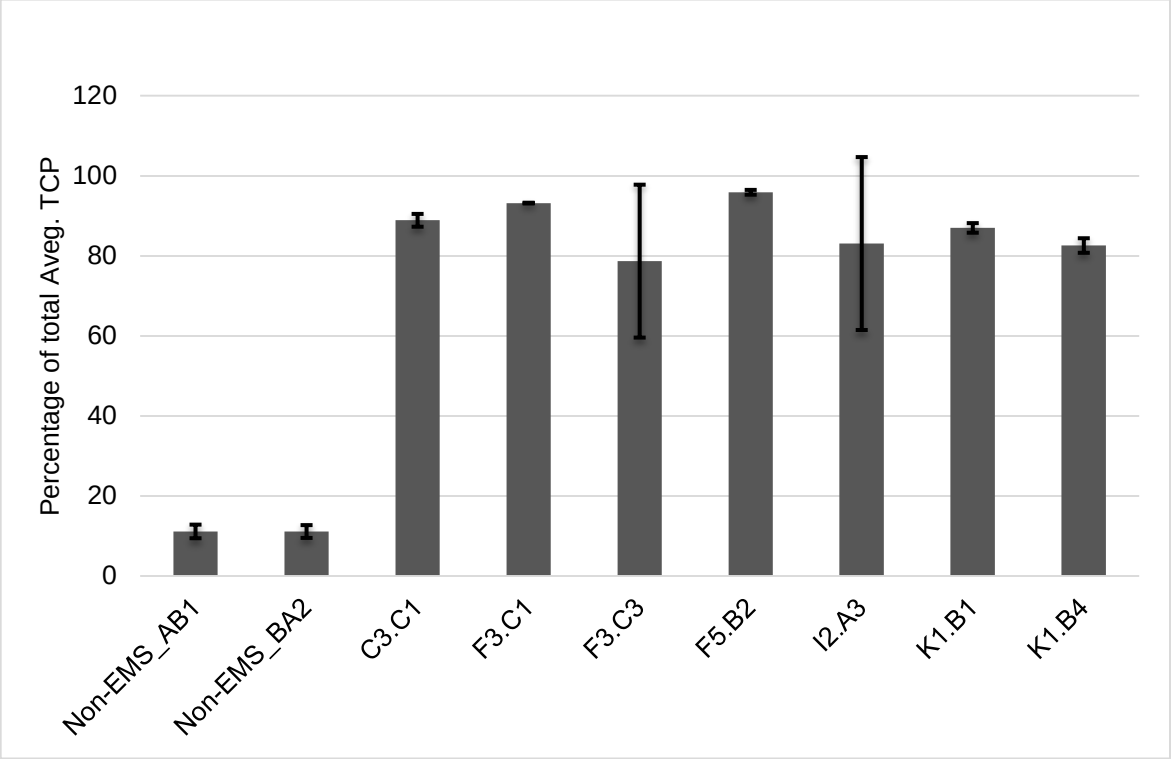


Figure 5.11. Percentage of tricellular pollen in *dazRF* M4 plants. The data show mean percentages of TCP for M4 families. Error bars represent SE (n = 4 plants).

5.5.1 Plant height analyses of *dazRF* M4 lines

The height of the primary inflorescences of a population of *dazRF* M4 lines and parental lines were measured. The average height of the parental plants was greater than that of *dazRF* M4 triple homozygous plants (Figure 5.11). The average height of the parental lines was 43 to 46 cm whereas, *dazRF* M4 plants ranged from 29 to 38 cm, with *dazRF*-C3.C1 showing the most marked reduction in height (Figure 5.12; Figure 5.13).

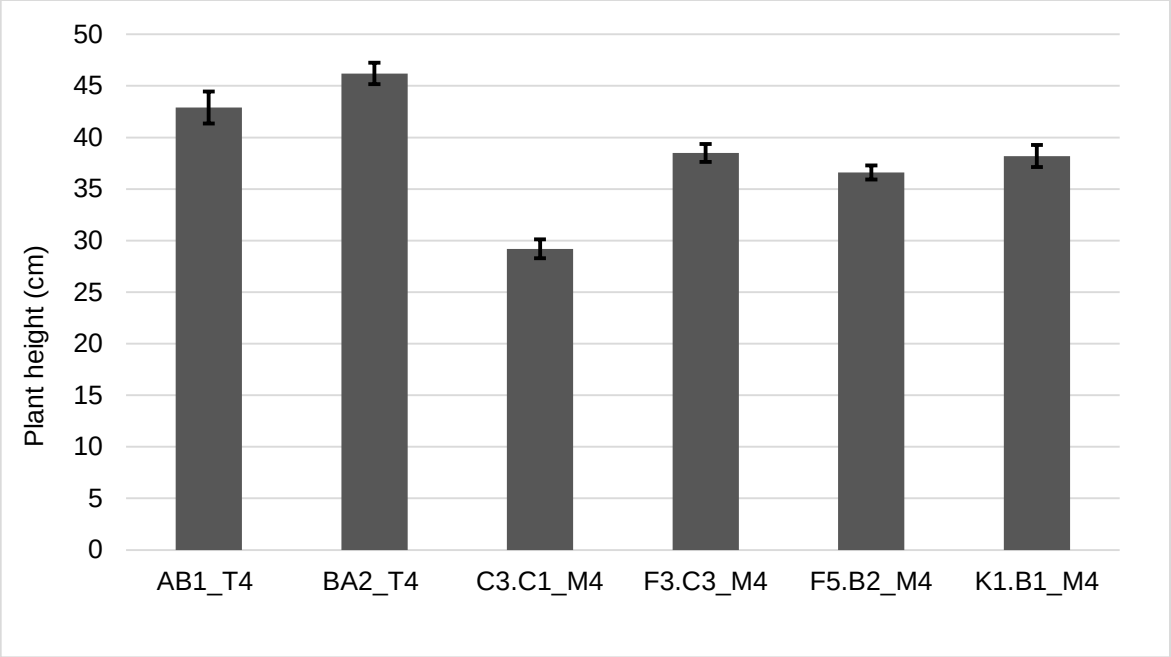


Figure 5.12. Mean plant height of *dazRF* EMS M4 plants. Parental non-mutagenised lines (AB1_T4 and BA2_T4) and *dazRF*-M4 lines are shown. Error bars represent SE (n = 4 plants).

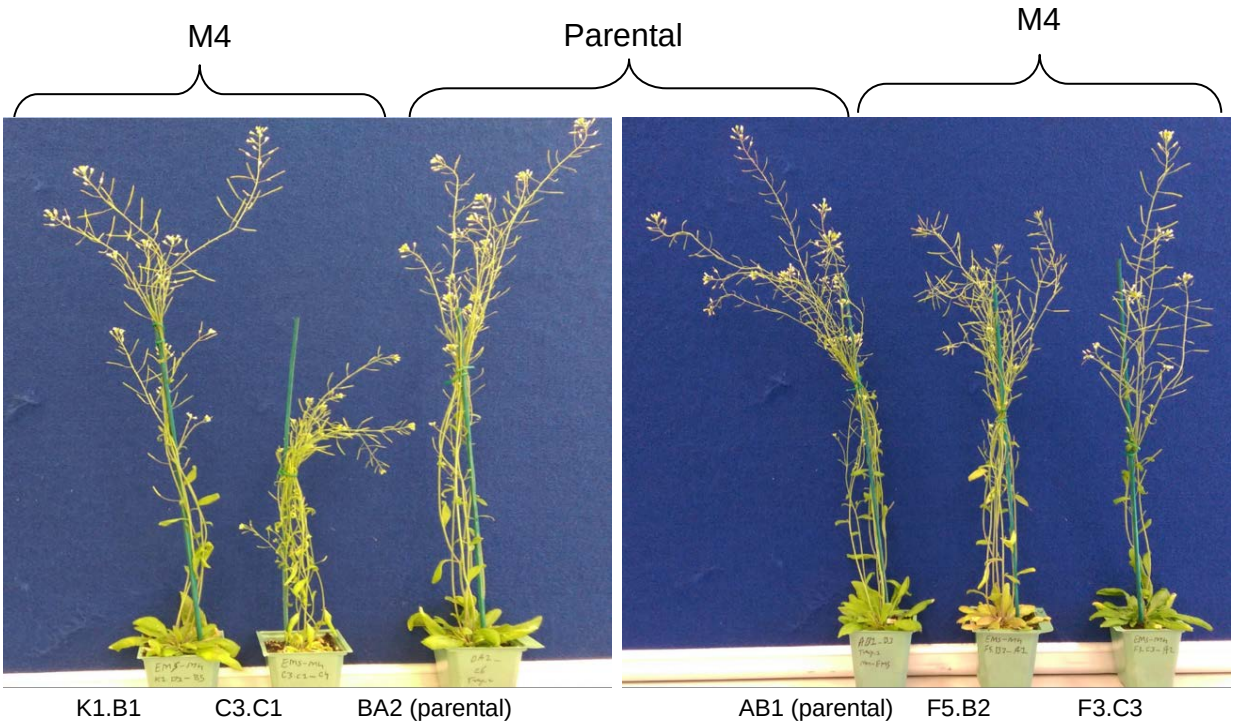


Figure 5.13. Phenotypes of *dazRF* M4 plants and parental lines.

5.5.2 Seed set in *dazRF* M4 plants

Seed number per silique in four *dazRF* M4 lines (C3.C1, F3.C3, F5.B2 and K1.B1) lines was determined. These lines showed stable restoration of fertility based on their germ cell division frequencies in the M3 generation. *dazRF* M4 lines showed variation in seed number. The C3.C1 and K1.B1 lines had the lowest number (21 to 22) of normal seeds, with 4 to 9 aborted seeds per silique, whereas F3.C3 and F5.B2 sister lines had 35 to 51 normal and 0 to 10 aborted seeds (Figure 5.14).

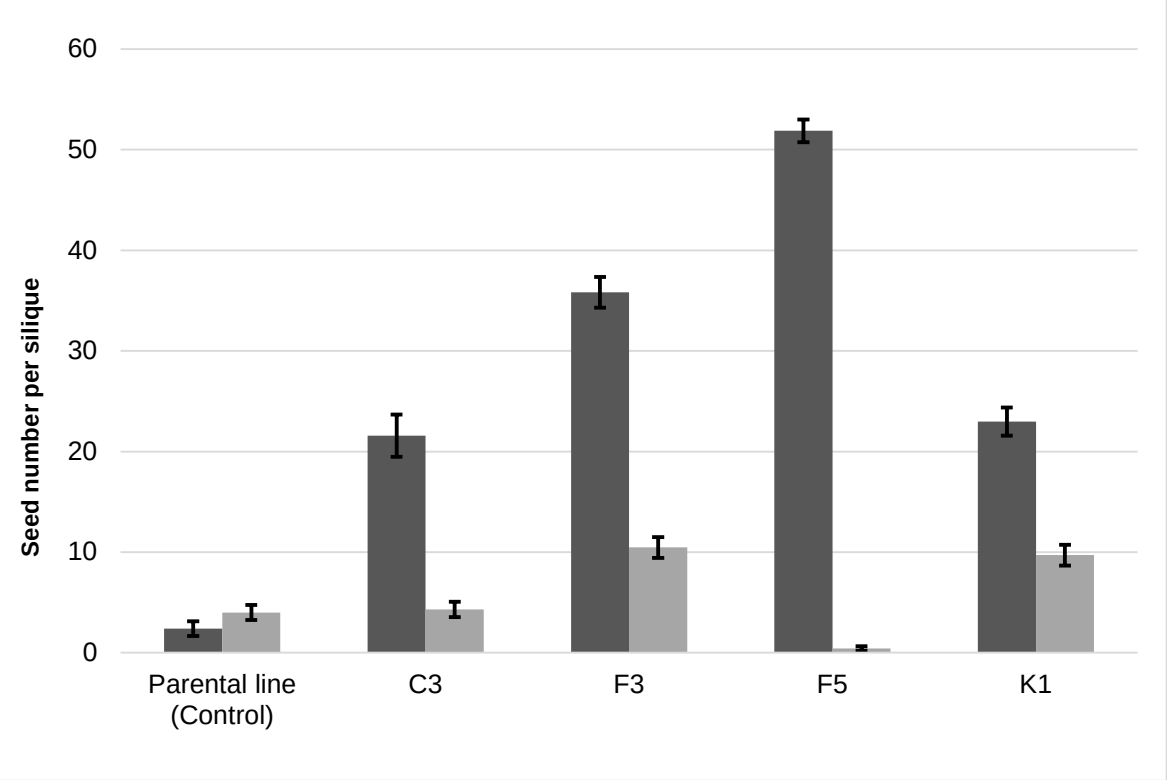
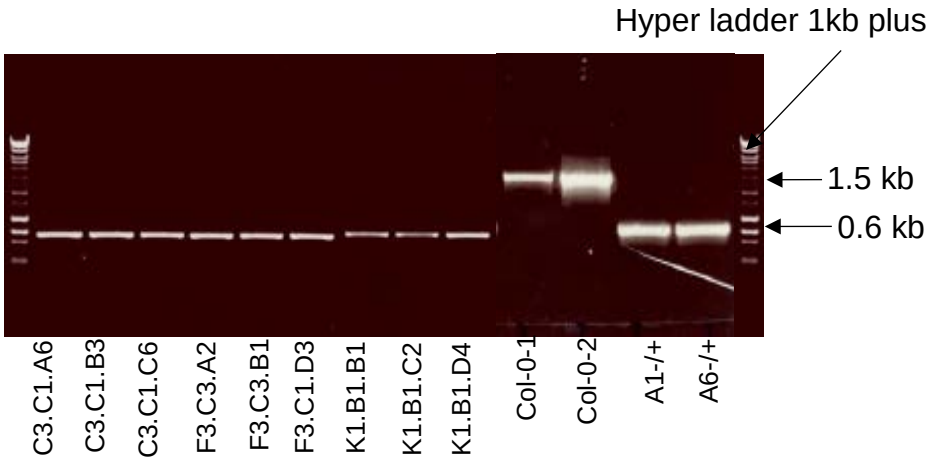


Figure 5.14. Average seed number in *dazRF* M4 plants. The average number of normal (dark bars) and aborted seeds (light bars) per silique is shown. Error bars represent SE (n = 30 siliques).

5.5.3 Genotypic analysis of *dazRF* M4 lines

After phenotypic analysis of *dazRF* M4 lines genomic DNA was extracted to confirm their genotypes. Below gel images shows *daz1-1 and daz2-1* fragment size (Figure 5.15).

(A) ***daz1-1***



(B) ***daz2-1***

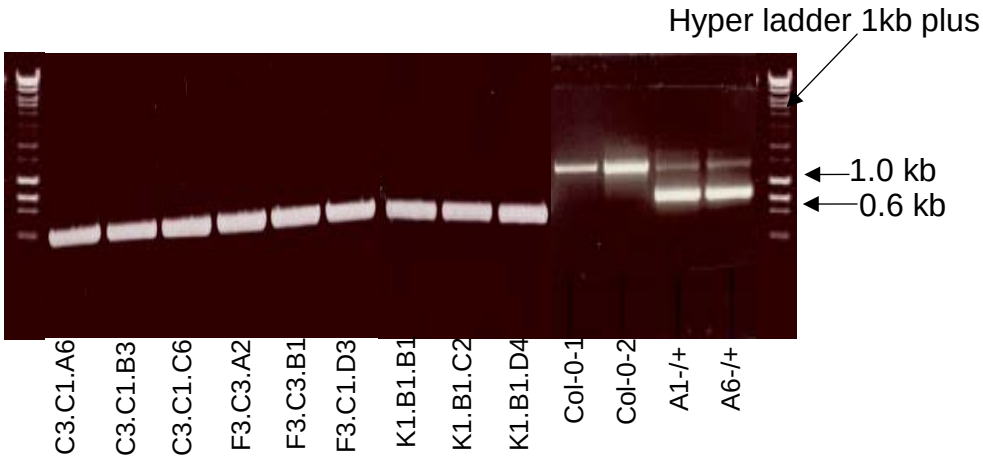


Figure 5.15. Genotypic analysis of *dazRF*-M4 plants. Gel electrophoresis images showing multiplex PCR products obtained with T-DNA left border (LB) primer with gene-specific primers, which amplify (A) *daz1-1* and (B) *daz2-1* alleles. Control

gDNAs: Col-0 WT and *daz1-1*$^{-/-}$ *daz2-1*$^{+/-}$ plants (A1, A6). See Figure 5.7 for primer positions.

5.6 Genetic analysis of *dazRF* lines

To identify the putative mutations responsible for restored fertility in stable *dazRF* lines whole genome sequencing provides an advanced platform. Some lines showed fertility restoration, but also other phenotypes such as leaf senescence and dwarfing (Figure 5.13). In these *dazRF* lines other unlinked mutations are likely to be involved and it is desirable to remove these by multiple backcrosses before sequencing as outlined in Figure 5.16. After 5-6 backcrosses, genomic DNA extracted from separate pools of low and high fertility plants could be used for whole genome sequencing. Identification of homozygous single nucleotide polymorphisms (SNPs) in the high fertility pool that are absent (or heterozygous) from the low fertility pool and are polymorphic with the Col-O reference genome, would represent candidate *dazRF* mutations.

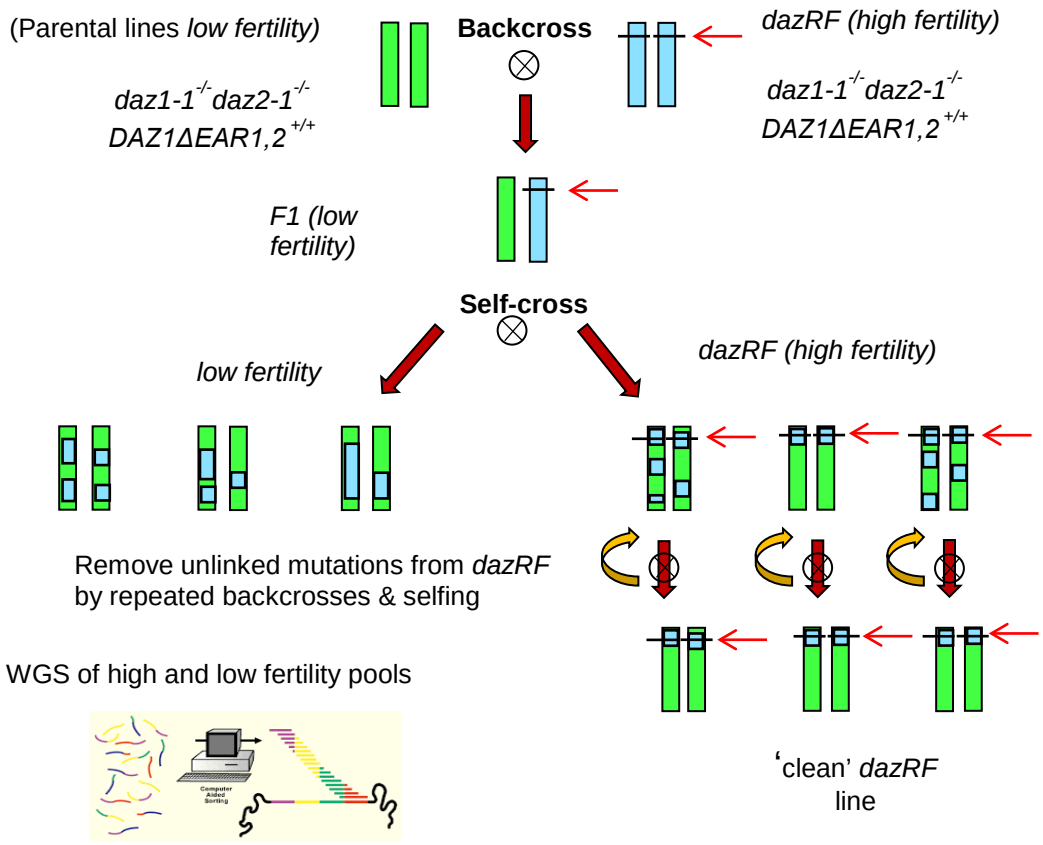


Figure 5.16. Scheme for removal of unlinked mutations from *dazRF* lines prior to whole genome sequencing. Parental lines and *dazRF* lines (*daz1-1*$^{-/-}$ *daz2-1*$^{-/-}$ *DAZ1ΔEAR1,2* $^{+/+}$) are back-crossed for several rounds and selfing. Low and high fertility pools are used for whole genome sequencing and their sequences compared.

5.7 Analysis of *dazRF* backcrossed generations

Based on phenotypic screening of *dazRF* lines through M3, M4 and M5 generations for stable restored fertility, *dazRF*-C3.C1, *dazRF*-F3.C3, and *dazRF*-K1.B1 lines were selected for backcrossing. M4 individuals were grown and used as the female parent in crosses with T4 progeny (AB1 or BA2) of the parental non-mutagenised line.

5.7.1 Phenotypic analysis of *dazRF* BC1 plants

The *dazRF* lines C3.C1, F3.C3, and K1.B1 showed stable restored fertility of 87 to 93 % TCP in the M4 generation. F1 backcrossed individuals from all three lines showed approximately half of the original frequency of TCP observed in M4 parental lines, but higher percentages of TCP than the non-mutagenised parental lines AB1 or BA2 (Figure 5.18). Examination of transgene expression in pollen revealed that it remained highly penetrant in lines C3.C1 and F3.C3 showing 86 to 96 % RFP positive pollen. However, it was reduced to 65 % to 72% in K1.B1 indicating partial suppression of transgene expression after backcrossing (Figure 5.17).

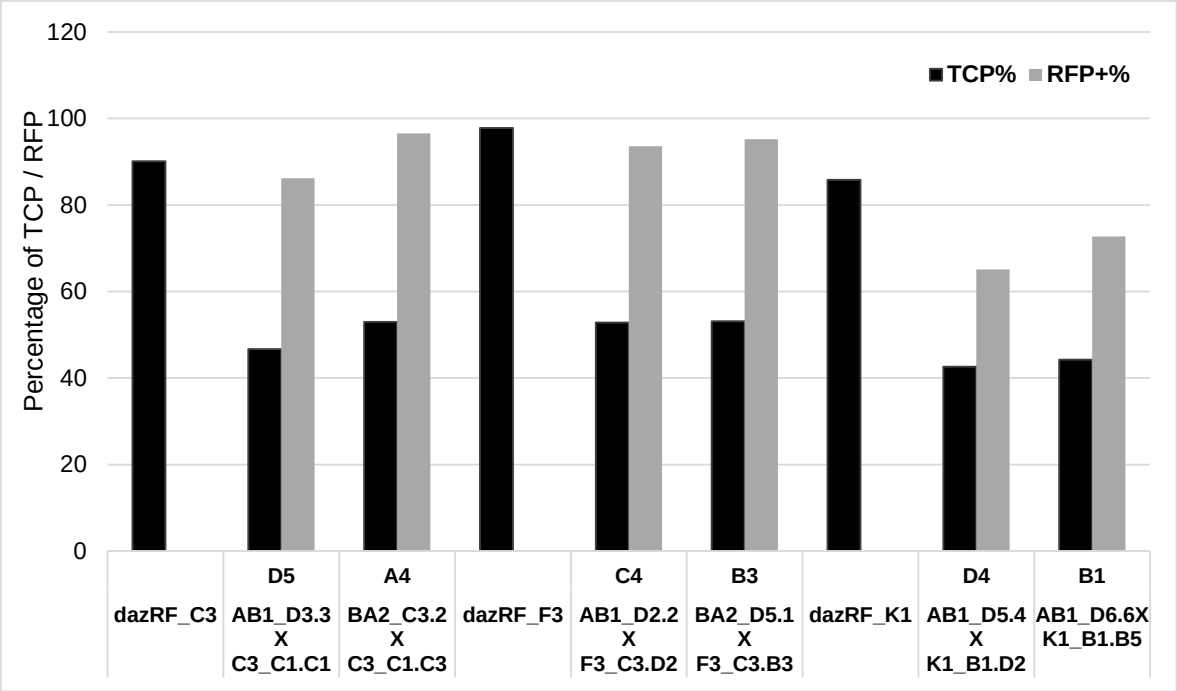


Figure 5.17. Pollen phenotypes of *dazRF* F1 (BC1) plants and parental *dazRF* M4 lines. Percentage TCP and RFP positive pollen are shown for parental *dazRF* M4 progeny and F1 progeny of *dazRF*-C3.C1, *dazRF*-F3.C3 and *dazRF*-K1.B1 M4 lines backcrossed with parental AB1or BA2 lines. Dark bars indicate percentage TCP. Light bars percentage of RFP positive pollen.

5.7.2 Phenotypic analysis of *dazRF* BC2 plants

F1 backcrossed lines were further backcrossed with non-mutagenised parental lines (AB1 and BA2) and rescue of germ cell division was examined by DAPI-staining (Figures 5.18, and 5.19). F1 (BC2) progeny from crosses with two sister lines (C4-D4 and B3-A1) ranged from 36 to 48 % TCP, showing restoration of fertility compared to the non-mutagenised parent lines AB1 and BA2 (Figure 5.18, Figure 5.20). More surprisingly the expression of the RFP in BC2 progeny ranged from 45 to 55 % in progeny of both sister lines, indicating partial suppression of the DAZ1ΔEAR1,2-mCherry transgene compared with BC1 plants (Figure 5.17).

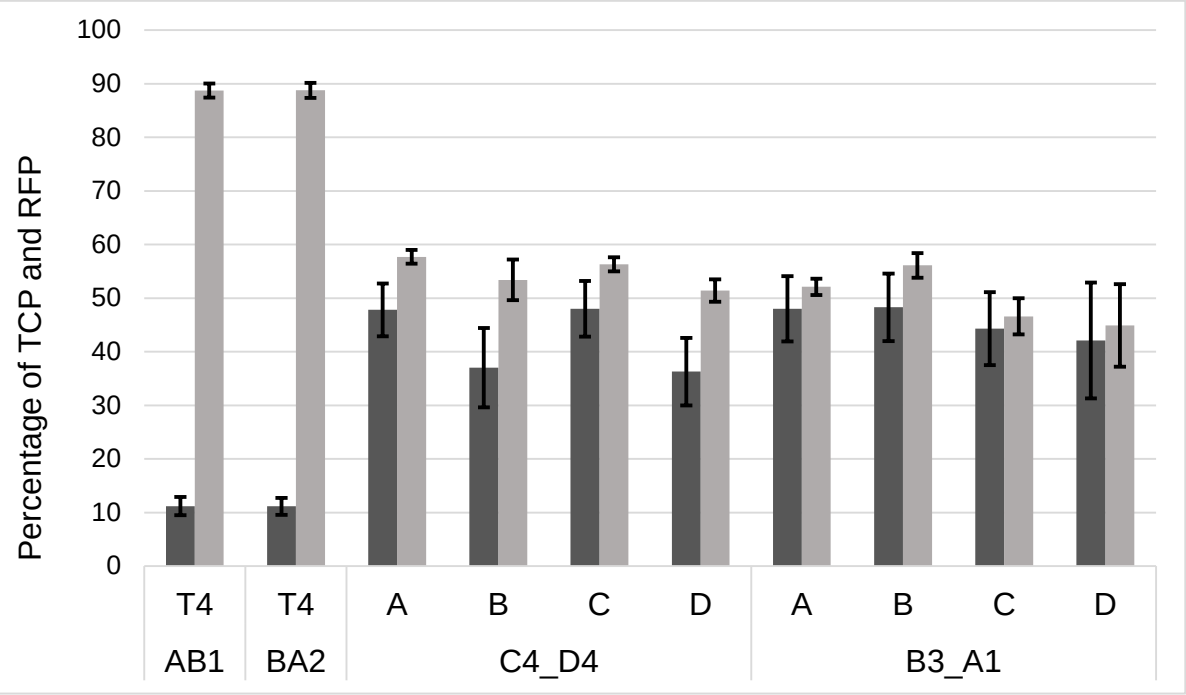


Figure 5.18. Phenotypic screening of the *dazRF* F2-backcrossed line. F1 backcrossed lines were further backcrossed with non-mutagenised parental lines AB1 or BA2. The mean percentages of TCP (dark bars) and RFP positive pollen (light bars) is shown for sister BC2 plants. Error bars represent SE (*n* = 4 counts).

(A)

(B)

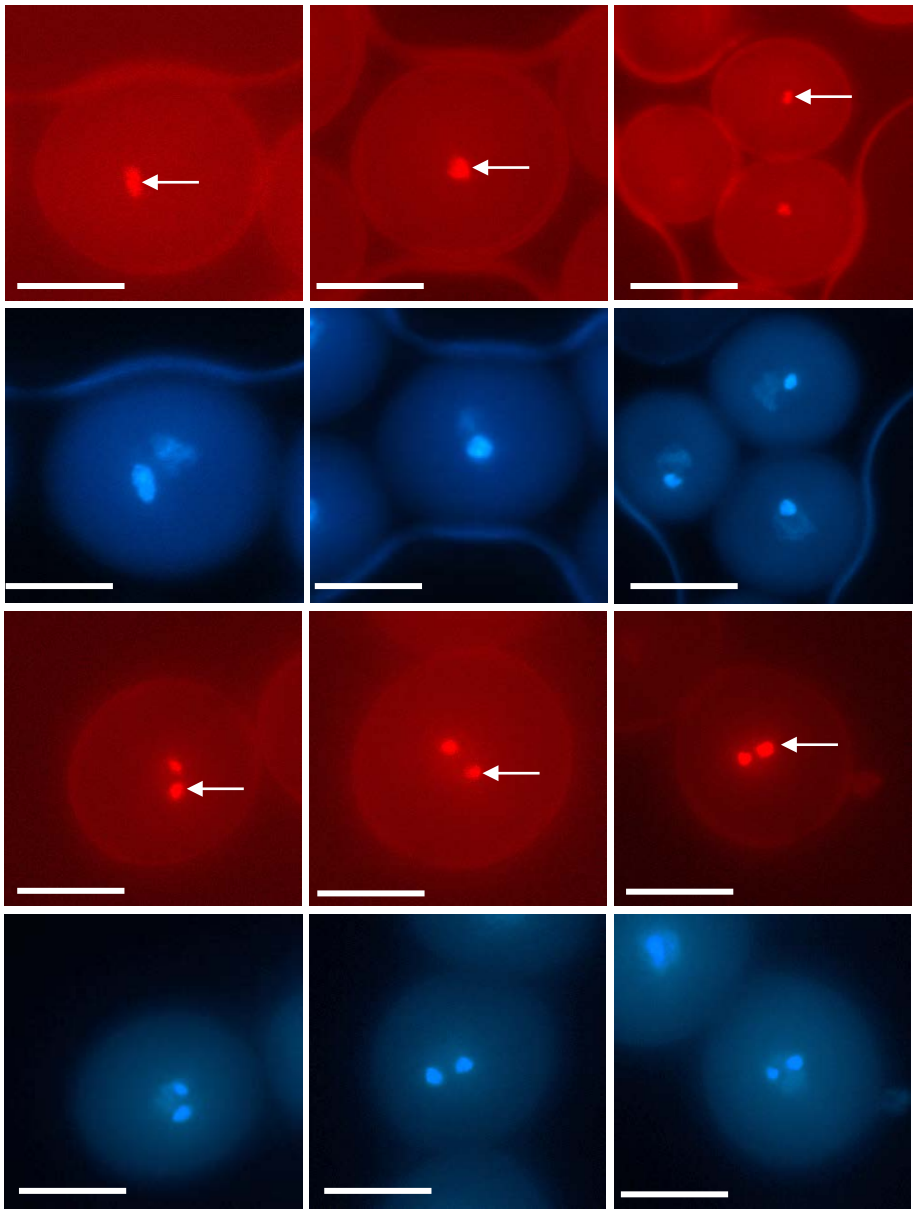
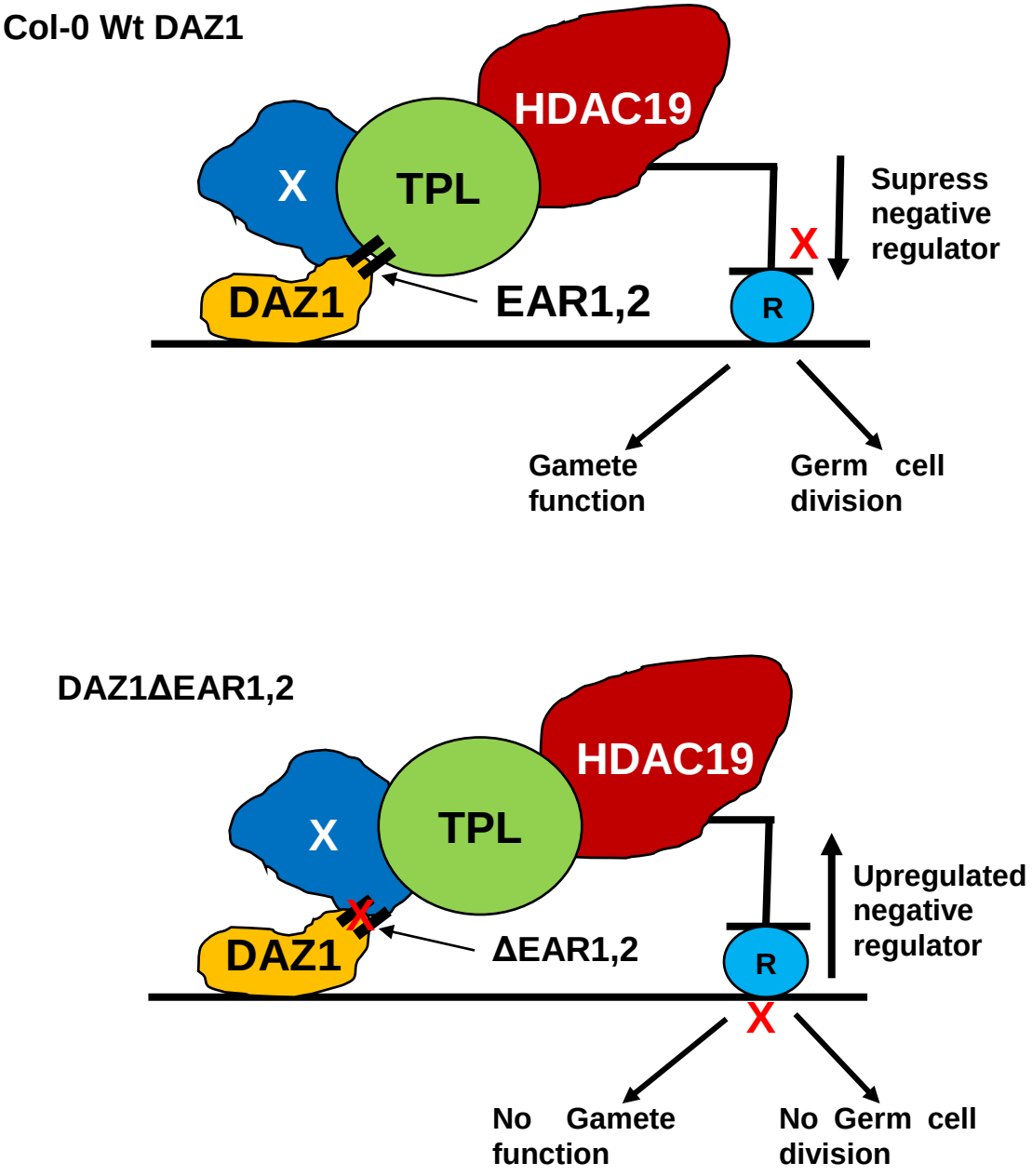


Figure 5.19. Pollen phenotypes of parental and *dazRF* BC2 plants. (A) Non-mutagenised parental AB1-A2 T4 generation, (B) Backcrossed *dazRF*-C4-D4A.

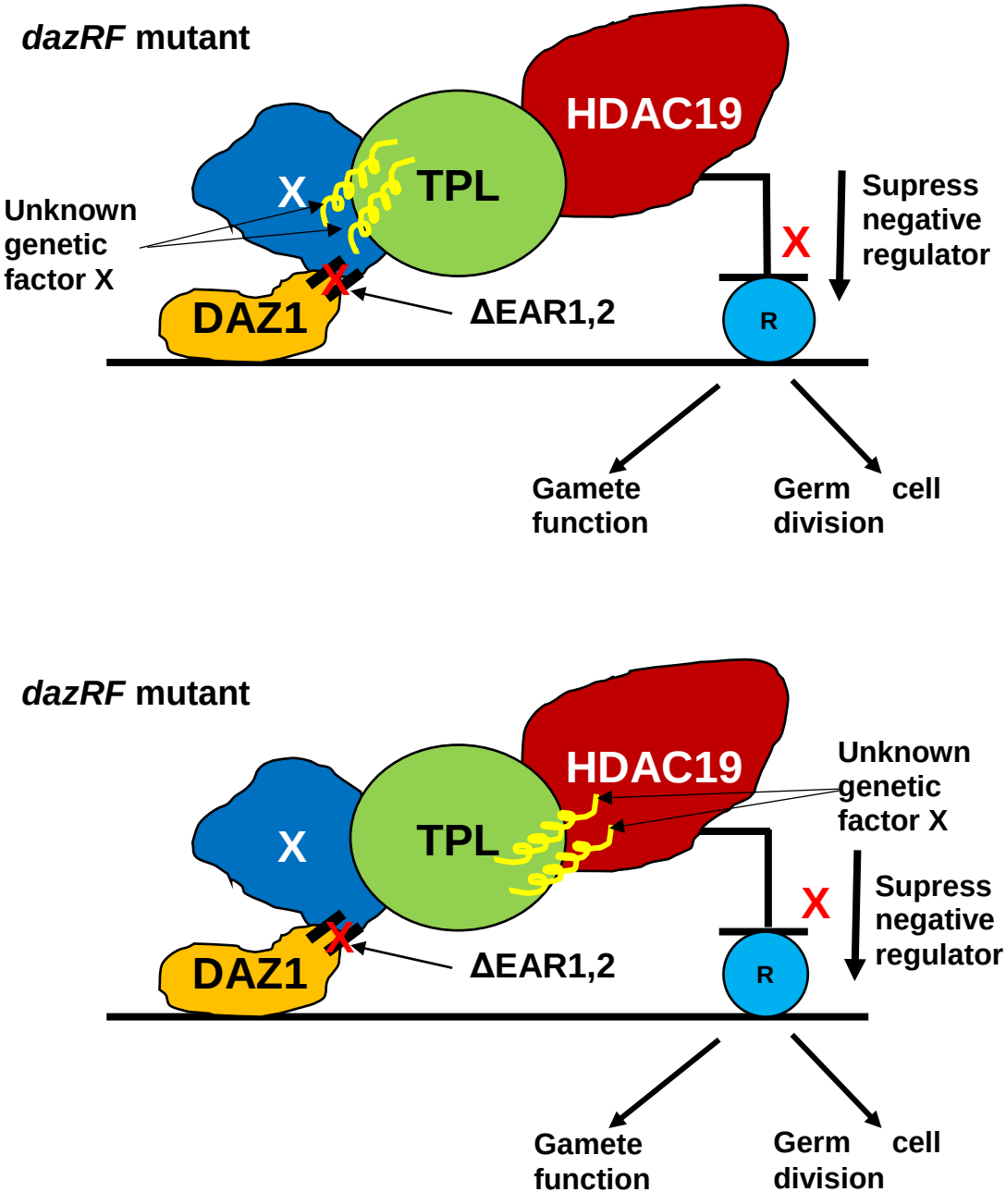
5.8 Discussion

Screening mutagenised populations generated by EMS and by T-DNA insertion has led to the isolation of fertility-related Arabidopsis mutants including numerous male-sterile plants (Ross *et al.* 1997, Bhatt *et al.* 1999, Siddiqi *et al.* 2000, Mercier *et al.* 2001). Similarly, the gametophytic mutants *duo pollen1* (*duo1*), *duo2* and *duo3* which affect male germ cell division and development (Durbarry *et al.* 2005, Brownfield *et al.* 2009) or *gemini pollen1* (*gem1*) and *gemini pollen2* (*gem2*) which affect division asymmetry of the microspore (Park *et al.* 1998, Twell *et al.* 2002, Park *et al.* 2004) were all isolated from EMS mutagenised populations.

Suppressor screens have been commonly used to identify novel alleles and factors that impact on specific developmental pathways. For example, in Yi and Jack (1998) isolated temperature-sensitive suppressors of *ap3-1*. APETALA3 (AP3) is an *Arabidopsis* floral organ identity gene essential for petal and stamen development and *ap3* mutants show organ transformation in which sepals develop in place of petals and carpels in place of stamens (Bowman *et al.* 1989, Jack *et al.* 1992). In mutagenised homozygous *ap3-1* M1 plants grown at 16 ºC, second whorl organs develop as sepaloid petals and the third whorl organs as stamens with carpelloid characters, but flowers are mostly self-fertile. At an elevated temperature of 23 ºC homozygous *ap3-1* M2 plants show a stronger typical *ap3*-1 phenotype and are male sterile (Yi and Jack 1998). Several independent suppressors alleles were isolated which exhibited suppressed phenotypes and restored fertility (Yi and Jack 1998).



(A) Schematic diagram shows potential DAZ1 interaction with TOPLESS protein and supress negative regulator (R) to control DAZ1 function. However, deletion of EAR1,2 motif in DAZ1 fail to proceed germ cell division (Borg *et al.* 2014).



(B) Schematic diagram shows potential unknown genetic factor X either interact with TOPLESS protein upstream of DAZ1 or downstream of DAZ1 where mutation in HDAC19 to communicate with TOPLESS and restore the DAZ1 function and proceed germ cell division.

Figure 5.20. Potential models for the restoration of DAZ1 function in triple homozygous (*daz1-1*$^{-/-}$; *daz2-1*$^{-/-}$; T$^{+/+}$) *A. thaliana* plants by mutations in unknown genetic factors.

In the suppressor screen performed for *daz1 daz2* DAZ1ΔEAR1,2 lines, phenotypic data for M2 to M5 *dazRF* lines were collected. Some *dazRF* lines showed long siliques with highly restored generative cell division (81-94% TCP). The isolation of EMS-induced mutants that restore the fertility to *daz1 daz2* DAZ1ΔEAR1,2 plants suggest that genetic factors can compensate for partial DAZ1 function, however, the nature of the mutations responsible remain unknown. A schematic diagram shows possible interactions in the DUO1/DAZ1 pathway that could allow mutations in unknown factors to restore DAZ1 function (Figure 5.20. B). The further genetic characterisation and identification of the mutation(s) responsible for *dazRF* phenotypes will help to pinpoint genetic factors which may interact with DAZ1 or which support DAZ1 function in male germline development.

Chapter 6

Conclusions and perspectives

6.1 Summary of thesis findings

The discovery of the *DUO1* gene in Arabidopsis was a breakthrough in understanding male germline development in angiosperms as DUO1 was the first germline-specific transcription factor to be identified in plants (Durbarry *et al.* 2005, Rotman *et al.* 2005). Subsequent research extended knowledge of the role of the DUO1 and its network to the discovery of genes encoding the C_2H_2 zinc finger proteins DAZ1 and DAZ2 which have been identified as key regulators of male germline development which act downstream of DUO1 (Borg *et al.* 2014). DUO1 plays a critical role in the male germline by co-ordinating germ cell division with sperm cell specification to control the production of twin sperm cells (Brownfield *et al.* 2009). Moreover, DAZ1 and DAZ2 expression is directly controlled by DUO1, and DAZ1 and DAZ2 are also required for mitotic division of the generative cell and for determination of sperm cell fate (Rotman *et al.* 2005, Borg *et al.* 2011, Borg *et al.* 2014). DAZ1 and DAZ2 show functional redundancy, as double mutant *daz1-1 daz2-1* pollen grains are bicellular at maturity and undivided germ cells fail to fuse with either central cell or egg cell (Borg *et al.* 2014).

In Borg *et al* (2014) experimental results have revealed that modification of the two EAR motifs in DAZ1 protein results in failure of germ cell division, such that EAR1 and EAR2 motifs are required for efficient germ cell division and male transmission. Although studies are incomplete, current data indicate that mutations in ZnF1 results in failure of germ cell division supporting an essential role in DAZ1 function (Rutley 2015). The work presented in this thesis showed that the DAZ1 basic region (BR) and conserved region (CR) have a functional role in DAZ1 (Chapter 3). Based on alignment of sequences of DAZ1 and DAZ2 orthologues, both BR and CR are conserved in dicot species, such as in the crop species tomato, potato, rice and maize, but the BR is not present in monocot species such as rice, barley and maize. Nevertheless, the BR is conserved in the basal angiosperm species *Amborella trichopoda* indicating an earlier origin and loss from the monocot clade (Figure 3.5.B). Understanding the functional importance of these regions in DAZ1 was a

major objective of this thesis. In Chapter 3, analysis of BR (muBR, ΔBR), and CR (LLAA, S119A) variants uncovered their functional importance *in planta*. DAZ-BR and DAZ-CR variants were shown to be important for the ability of DAZ1 to trans-activate the male germ cell-specific GEX1 promoter. Further, the localisation of DAZ-BR and DAZ-CR variants was unaffected in sperm cells.

A previous study involving mutagenesis of ZnF domains of DAZ1 concluded that mutant ZnF1 is essential for the DAZ1 function *in planta* (Rutley 2015). Here, this domain analysis was extended to study the role and independence of ZnF2 and ZnF3. The number of Q-type zinc finger domains varies between DAZ1 and DAZ2 homologues therefore modification or loss of zinc finger domains may be adaptive and could affect the function of DAZ1. Work described in Chapter 3, validated the functional role of ZnF2 *in planta* and this was extended to demonstrate the contribution of ZnF3 for optimal DAZ1 function. Thus, both Q-type zinc fingers (ZnF2 and ZnF3) play an important, but non-essential role in supporting DAZ1 control of germ cell division (Rutley 2015).

In 2014, it was shown that homozygous *daz1-1 daz2-1* plants expressing DAZ1ΔEAR1,2 showed highly reduced fertility but unexpectedly showed seed abortion (Rutley 2015). The studies in Chapter 4 determined whether this seed abortion phenotype was specific to the removal of the EAR motifs, by examining the fertility of homozygous *daz1-1 daz2-1* plants expressing other mutant DAZ1 proteins. Observation of seed abortion in *daz1-1 daz2-1* DAZ1muZnF2 plants (Figure 4.5) showed that seed abortion phenotype is not specific to the EAR domain, but likely arises due to any reduction in DAZ1 function.

Chapter 5 was dedicated to the isolation of suppressors of the reduced fertility of *daz1 daz2* mutants expressing DAZ1 with impaired function. Stable *dazRF* suppressor lines were maintained through to the M5 generation and the results suggest that unknown mutations can restore germ cell division and fertility. The

further analysis of stable *dazRF* lines showed that these lines and putative mutations still restore plant fertility and rescue impaired DAZ1 function after backcrossing.

6.2 DAZ1 BR and CR domains have functional roles

Phenotypic analysis showed that full-length AtDAZ1-mCherry shows complete rescue of germ cell division in *daz1 daz2* pollen, but DAZ1-BR mutants (DAZ1muBR-mCherry and DAZ1ΔBR-mCherry) show only limited rescue (Figure 3.19). DAZ-CR mutants (DAZ1LLAA-CLLM-mCherry and DAZ1S119A-mCherry) also showed partial rescue of *daz1 daz2* germ cell division (Figure 3.20), indicating that both regions are important and/or affect the function of DAZ1 in germ cell division. Whether monocot DAZ1 orthologous are able to rescue germ cell division despite their lack of a homologous BR remains to be determined. Further, the activity of the germline-specific *GEX1* promoter was analysed with DAZ1-BR and DAZ1-CR variants and the results indicate that the DAZ1 BR and DAZ1 CR are also important for promoter *trans*-activation.

6.3 Variability of protein localisation in DAZ1/DAZ2 and orthologues

The nucleus is the significant feature of the eukaryotic cell and segregates the genetic material and the transcriptional machinery in the nucleus from the translational and metabolic machinery in the cytoplasm. One of the key features of such regulation is the selective bi-directional transport between the cytoplasm and nucleus. Whereas, transport happens through large membrane embedded structures termed nuclear pore complexes (NPCs), comprised of approximately 30 proteins collectively termed nucleoporins (NUPs) (Aitchison and Wozniak 2007, Alber *et al.* 2007, Marfori *et al.* 2011). DAZ1 and DAZ2 protein both contain a putative nuclear localisation signal (NLS). AtDAZ1-mCherry was shown to be enriched in the nucleus, but is also present (with lower signals) in the cytoplasm of sperm cell, DAZ1-BR variants showed a similar distribution, thus the DAZ1 BR is not required for the nuclear enrichment in DAZ1 germ cells (Figure 3.26C, D).

Surprisingly, At-DAZ2-mCherry shows tight nuclear localisation (Figure 3.26 A, B). The basis of this difference is not clear, but regions outside the BR are likely to be important for protein interactions with components which may allow retention in the nucleus (DAZ2) or with nuclear export factors or cytoplasmic retention factors (for DAZ1).

Localisation of DAZ1 in tobacco leaf cells tell a different story for DAZ1muBR-GFP and DAZ1ΔBR-GFP. Mutation and deletion of DAZ BR disturbs the normally tight localisation compared to full-length DAZ1-GFP (Figure 3.18A). The different outcomes in leaves and germ cells indicate that details of the mechanisms of nuclear localisation may operate differently in sperm cells compared with leaf cells. It is speculated that localisation of DAZ1 may be influenced by the presence of a nuclear export signal, protein interactions, the size of the protein or other factors (Marfori *et al.* 2011).

Analysis of the localisation of modified DAZ1 proteins and DAZ1 orthologues indicated that they do not behave constantly. Whereas, DAZ1muZnF2,3-mCherry and DAZ1smEAR1,2-mCherry variants showed nuclear-enriched signals similar to AtDAZ1-mCherry (Figure 3.27), DAZ1muZnF1-mCherry and DAZ1muZnF1,2-mCherry showed tight nuclear localisation similar to AtDAZ2-mCherry (Figure 3.27, Figure 3.28). This raised the question of how DAZ1muZnF1 and DAZ1muZnF2,3 behave differently? The mutation in ZnF1 may alter DAZ1 interaction with its normal substrates which might include chromatin or other protein complexes, thereby affecting nuclear export pathways. Alternatively changes in ZnF1 may induce or cause structural changes to DAZ1 that expose the BR to the nuclear import machinery. Previous research shows that classical NLSs sequences are usually found at N and C terminus of proteins, between domains or in flexible loop regions. This is reliable with the mode of interaction, which discovered by structural studies, therefore suggesting that cNLSs adopt an ordered structure only upon binding Importin-α, in line with the expected mechanism of action of the majority of linear motifs (Neduva and Russell 2005).

DAZ1 orthologs fused to mCherry showed variation in localisation with ZmDAZ1A/B, OsDAZ1-mCherry, and SlDAZ1-mCherry showing nuclear-enrichment and 21% to 32% nuclear-specific signal compared with AtDAZ1-mCherry showing 36 % nuclear specific signal. More surprisingly in BrapaDAZ1-mCherry was more tightly localised to the nucleus, with 76 % of pollen with nuclear-specific signal. Since the BR is not required for nuclear localisation of AtDAZ1 and monocot DAZ1 orthologs do not possess the BR they must be guided to the nucleus by other cryptic signals in their sequences or through interaction with unknown partners. Given the presence of BR sequences in *A. trichopoda* DAZ1 the BR may have been substituted after the divergence of monocots. In conclusion the DAZ1 nuclear import signal remains elusive as even after mutation and deletion of the BR as a putative NLS, DAZ1 still shows nuclear-enrichment.

6.4 ZnF domains are required for DAZ1 function in *A. thaliana*

C_2H_2 zinc finger proteins constitute an abundant family of nucleic acid binding proteins in the genomes of higher and lower eukaryotes. C_2H_2 zinc fingers (ZnFs) display a wide range of functions from DNA or RNA binding to protein-protein interactions (Englbrecht *et al,* 2004). Zinc finger proteins not only act in transcriptional regulation, either directly or through site-specific modification and regulation of chromatin, but participate in RNA metabolism and in additional cellular functions that require specific protein association involving the ZnF domains (Englbrecht *et al.* 2004). Recent research demonstrated that ZnF3 has been lost from AtDAZ1 homologs from *S. lycopersicum* and *Vitis vinifera*) (Rutley 2015). Additionally, no proteins were recognised that lack both ZnF2 and ZnF3 signifying that at a minimum one Q-type ZnF is required for DAZ1 function (Rutley, 2015).

Previous research demonstrated that ZnF1 is essential for DAZ1 function, but DAZ1muZnF2-mCherry and DAZ1muZnF3-mCherry show ~60 % rescue of germ cell division in *daz1 daz2* pollen relative to full-length DAZ1-mCherry (Rutley, 2015).

In this thesis, complementation analysis in T2 generation lines confirmed that DAZ1 ZnF1 is important for DAZ1 activity as DAZ1muZnF1 and DAZ1muZnF1,2 failed to rescue germ cell division. Similarly T2 lines of DAZ1muZnF2 and muZnF3 showed partial rescue with 63 % and 68 % efficiency compared with AtDAZ1-mCherry, similar to previous observations (Figure 3.32, Figure 3.33) (Rutley 2015).

There have been wide ranging studies investigating the functional role of other C_2H_2 zinc finger proteins in plants. For example, it was demonstrated that SUPERMAN is a zinc finger protein and possesses additional putative motifs typical of a transcriptional repressor (Payne *et al.* 2004). Recessive mutations at the *SUPERMAN (SUP)* locus lead to an extension of B-class gene expression and results in supernumerary stamen production at the expense of fourth whorl carpel development in *Arabidopsis thaliana* (Bowman *et al.* 1992, Sakai *et al.* 1995). In addition the *KNUCKLES* locus, like *SUPERMAN*, encodes a small protein containing a single C_2H_2 zinc finger and functions as a transcription repressor (Payne *et al.* 2004).

C_2H_2 zinc finger proteins play important roles in metabolic and developmental pathways as well as in stress responses and defence activation in *A. thaliana* (Ciftci-Yilmaz and Mittler 2008). Recent experimental results shows that the importance of C_2H_2-type ZFPs with repression activity in defence and stress responses. For example, *ZINC FINGER PROTEIN 12* (ZAT12) was originally isolated as a light stress response cDNA and was identified by transcriptome analysis of plants subjected to different biotic and abiotic stress conditions (Mittler *et al.* 2004). Recent data suggested a role for ZAT12 in cold acclimation and in the response of plants to oxidative stress (Rizhsky *et al.* 2004, Davletova *et al.* 2005, Vogel *et al.* 2005).

6.5 Aborted seeds arise from enlarged ovules in partial function DAZ1 mutants

In aangiosperms seed development is major event where plants put their all resources to complete the developmental staged for the future generation. Fusion

of the male female gametes leading to the zygotic stage is critical for seed development (Aw *et al.* 2010), whereas, seeds must accumulate enough reserves to complete the initial stages of germination and seedling establishment (Mena-AlÍ and Rocha 2005). This process requires signals from a double fertilisation event that occurs in the embryo sac (Rodrigues *et al.* 2008). Recent work has shown that ovule enlargement and seed development also depends on the pollen tube contents (PTC) including the vegetative cell nucleus and cytoplasm as well as the two sperm cells (Kasahara *et al.* 2016).

Seed abortion in *daz1-1 daz2-1* homozygotes expressing mutant DAZ1 proteins is described in Chapter 4. Based on their ability to produce low levels of tricellular pollen and viable seed, double fertilisation events can occur. Previous studies also revealed that *daz1-1 daz2-1* pollen tubes are able to deliver two sperm cell, but are unable to initiate fertilisation (Borg *et al.* 2014). However, in *daz1 daz2* mutants expressing DAZ1 variants some ovules enlarge but abort and do not contain embryo or endosperm. This seed or ovule abortion phenomenon was shown to be not specific to particular DAZ1 protein domains and rather reflects impaired DAZ1 protein function and possibly sperm cell viability. Several possibilities could account for this phenomenon. For example, incomplete sperm cell differentiation in pollen rescued with impaired DAZ1 proteins, such that fertilisation (single or double) results in early failure of the zygote and endosperm prior to division, but stimulates ovule enlargement. Further, DAZ1 without its EAR motifs may not be able to bind with TOPLESS and not able to communicate with or regulate genes in the DUO1-DAZ1 network. This network includes male germ cell specific genes like *GCS1/HAP2*, *GEX2* which are required for gamete adhesion and fusion (Mori *et al.* 2005).

Recent investigations of the failed fertilisation by defective mutant male gametes showed that ovule fertility may be rescued by a second functional wild type pollen tube. This is termed polysiphonogamy because, an ovule accepts multiple pollen tubes (Kasahara *et al.* 2012). Recently experimental data reported that after pollen tube discharge and failure of male gametes to fertilise, the seed coat stretches and

ovules increase in size, a phenomenon termed pollen tube dependent ovule enlargement morphology (POEM), however this event only happened when the ovules accepts the pollen tube contents (PTC). This is suggested as a paternal function of PTC in to promote the development of the ovule (Kasahara *et al.* 2017). It is also reported that PTC leads to induction of gene expression in ovules as a result ovule enlargement and seed coat formation without fertilisation and might increase endosperm nuclei numbers without fertilisation in mutants (Kasahara *et al.* 2016).

6.6 Putative suppressor screening restore fertility in DAZ1 mutant background

Suppressor screening was used to identify novel genes that may interact with or regulate DAZ1 to restore plant fertility. Screening identified *dazRF* restored fertility lines in mutagenized *daz1 daz2* null DAZ1ΔEAR1,2-mCherry plants. *dazRF* M2, M3, M4 and M5 generations were characterised and some lines (C3, F3 and K1) were found to show stable fertility restoration. These results suggest that unknown mutations in these lines influence or overcome partial DAZ1 function. Possible candidate mechanisms might involve ectopic expression of factors which bind to DAZ1 target gene promoters, replacing DAZ1 function. Alternatively expression of modified proteins which can bind to DAZ1ΔEAR1,2-mCherry to enhance its function such as those harbouring EAR motifs, thereby restoring interaction with the corepressor TPL. This latter possibility would be dependent upon the expression of DAZ1ΔEAR1,2-mCherry. This dependence could be tested by CRISPR/Cas9 mediated gene editing of DAZ1ΔEAR1,2-mCherry to prevent expression.

6.7 Conclusion

The major objective of this research was to uncover previously unknown roles of the DAZ1 BR and DAZ1 CR. Both regions were shown to be important for DAZ1 function in germline development in *Arabidopsis thaliana* and mutant DAZ1-BR and DAZ1-CR proteins failed to trans-activate putative DAZ1 target promoter GEX1 in tobacco leaf cells. Although the BR and CR were not required for nuclear-enriched

localisation of DAZ1 in sperm cells, mutant DAZ1-BR proteins were delocalised in tobacco leaf cells indicating differences in the machinery guiding DAZ1 to the nucleus in germ cells and leaves or between species. DAZ1 protein localisation in germ cells was found to be independent of ZnF2, ZnF3 and the EAR motifs, but nuclear enrichment was enhanced by mutations in ZnF1 suggesting different impacts of ZnF domains on the mechanism of protein localisation.

The study of seed abortion phenomena in *daz1 daz2* mutants expressing DAZΔEAR12 was extended and shown to occur in plants expressing DAZ1smEAR1,2, and DAZ1muZnF2. This showed that seed abortion was not particular to specific domains or motifs and likely reflects general DAZ1 loss of function. Although different DAZ1 variants are still able to produce a proportion of tricellular pollen, aborted seed do not develop embryos or endosperm which may indicate stimulation of ovule growth after fertilisation by defective gametes.

In a genetic screen in *Arabidopsis thaliana*, plants that restore fertility to plants with partial DAZ1 function were isolated and characterised. Several of these *dazRF* suppressor lines, which were stable in M2 to M5 generations, could provide access to novel regulators of the DUO1-DAZ1 network that orchestrates germ cell division and sperm differentiation in *Arabidopsis*.

In conclusion, the relative importance of different protein domains in *Arabidopsis thaliana* DAZ1 was demonstrated. This information will help to uncover the molecular mechanisms, protein interactions and function of DAZ1 as a transcription factor in the DUO1-DAZ1 network controlling male germline development. This information also provides valuable background knowledge of the control of fertility that may be of future value to help secure food security for humankind.

Appendices

TableA1: Sequences of primers used (5'-3')

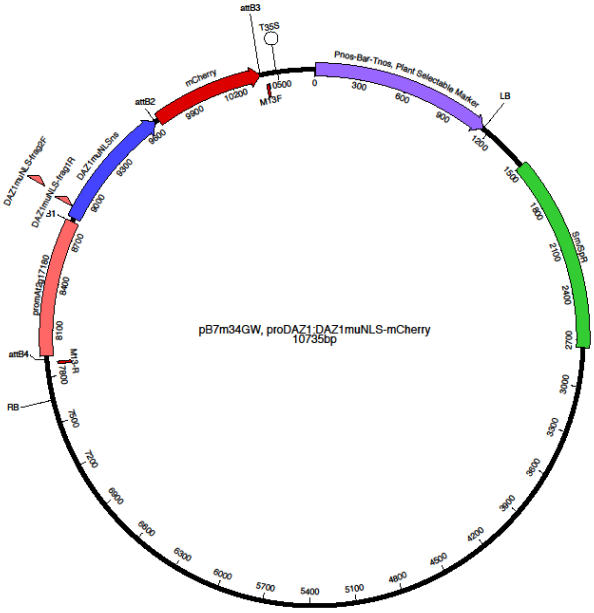
Primer names	Sequences
attB4-F adapter	GGGGACAACTTTGTATAGAAAAGTTG
attB4-R adapter	GGGGACTGCTTTTTTGTACAAACTTG
attB2-F adapter	GGGGACAAGTTTGTACAAAAAAGCAGGCT
attB2-R adapter	GGGGACTGCTTTTTTGTACAAACTTG
M13 F	GTAAAACGACGGCCAG
M13 R	CAGGAAACAGCTATGAC
ProDAZ1 attB4-F	TGTATAGAAAAGTTGAAGTGGCACAAACCAACCC
ProDAZ1 attB1-R	TTTTGTACAAACTTGTATTATTGAGTCTCTTACTAGAG
DAZ1 attB1-F	ACAAAAAAGCAGGCTCTATGTCGAACACTTCAAACTCCG
DAZ1s attB2-R	ACAAGAAAGCTGGGTCTCAAAGTCCTAACCTAAGATCC
DAZ1ns attB2-R	ACAAGAAAGCTGGGTCAAGTCCTAACCTAAGATCC
DAZ2s attB1-F	ACAAAAAAGCAGGCTCTATGAACAACAATCATTCCTATG
DAZ2ns attB2-R	ACAAGAAAGCTGGGTCTTAGAGTCCTAACCTAAGATC
daz1-1 (salk_058012) LP	TGATTTCGAAATGTGGAATGG
daz1-1 (salk_058012) RP	CAACAACTTCCACCCTGAATC
daz1-1 (Salk_101906) LR	CAGATGCTTATGGCATTTTCTG
daz1-1 (Salk_101906) RR	CTCATGTGACCAAAGAGAGCC
SALK LBb 1.3	ATTTTGCCGATTTCGGAAC
DAZ1muZf1 OE-F	CTCTTTGGAAACATGAGATGTC
DAZ1muZf1 OE-R	GACATCTCATGTTTCCAAAGAG

DAZ1muZf2 OE-F	GTTAGGCGGAAATAGAGCAAC
DAZ1muZf2 OE-R	TGTGTTGCTCTATTTCCGC
DAZ1muZf3 OE-F	GTTAGGAGGTAACATGAGATGTCATTG
DAZ1muZf3 OE-R	CATCTCATGTTACCTCCTAACGCTTGAC
DAZ1muNLS_frag 2F	CCAAACAACAACAACACTAACCTCACCAACAACGAAGTC
DAZ1muNLS_frag 1R	GTTGTTGTTGTTTGGGTTTTGATTGTAACTAGGAAG
DAZ1muNLS_frag 2F	ACAATCAAAACCTCACCAACAACGAAGTC
DAZ1muNLS_frag 1R	GTTGTTGGTGAGGTTTTGATTGTAACTAGGAAG
LLAACLLM_F	CCTCCTGTGCAGCAATGAT
LLAACLLM_R	CCATCATTGCTGCACAGGAGG
SACLLM_F	GATCCAGCTGAGGAAGAGC
SACLLM_R	CTCTTCCTCAGCTGGATCCC

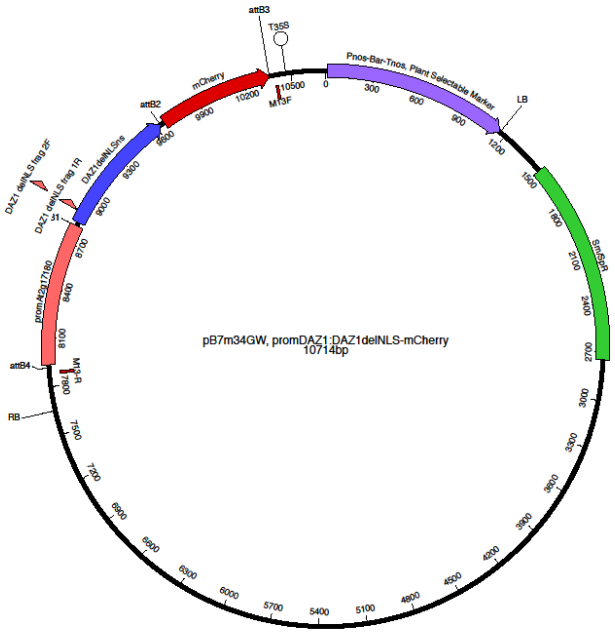
T-DNA binary vector maps

All binary vector constructs were made from individual entry clones in the destination vector pB7m34GW by 3-part recombination (Karimi et al, 2002). A map for each gene construct is shown (alternative name in parentheses).

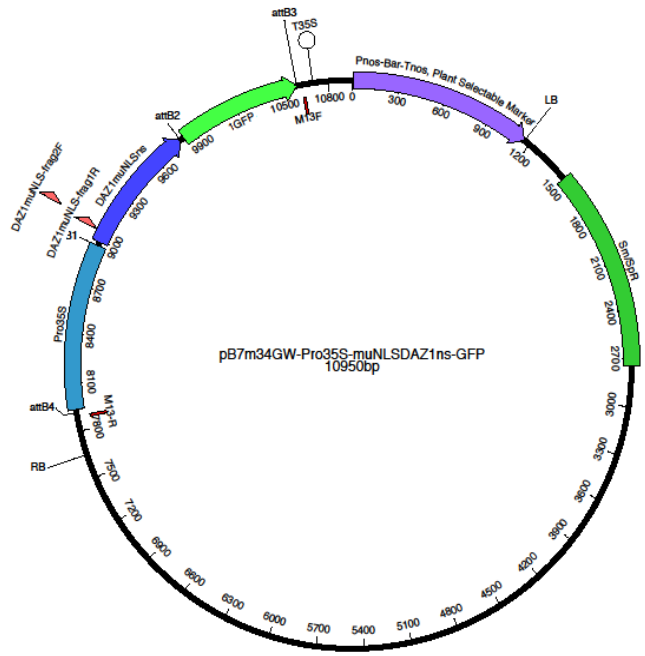
ProDAZ1:DAZ1muBR-mCherry (ProDAZ1:muBRDAZ1-mCherry)



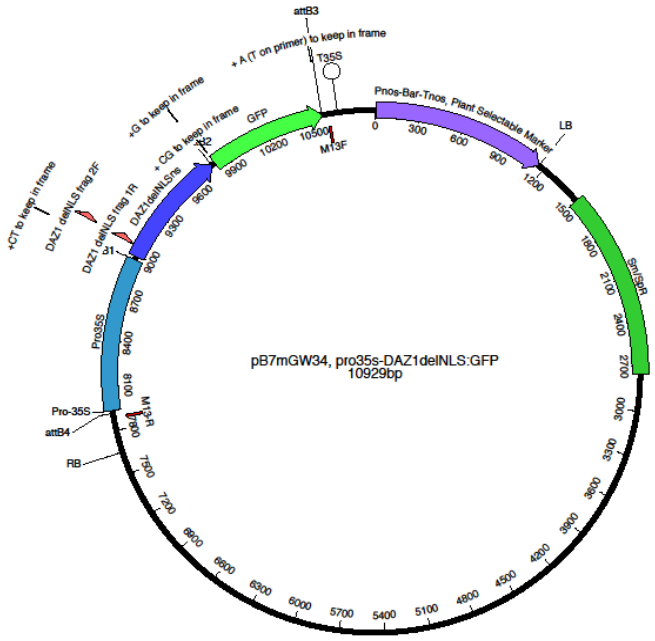
ProDAZ1: DAZ1ΔBR-mCherry (ProDAZ1:delNLSDAZ1ns-mCherry)



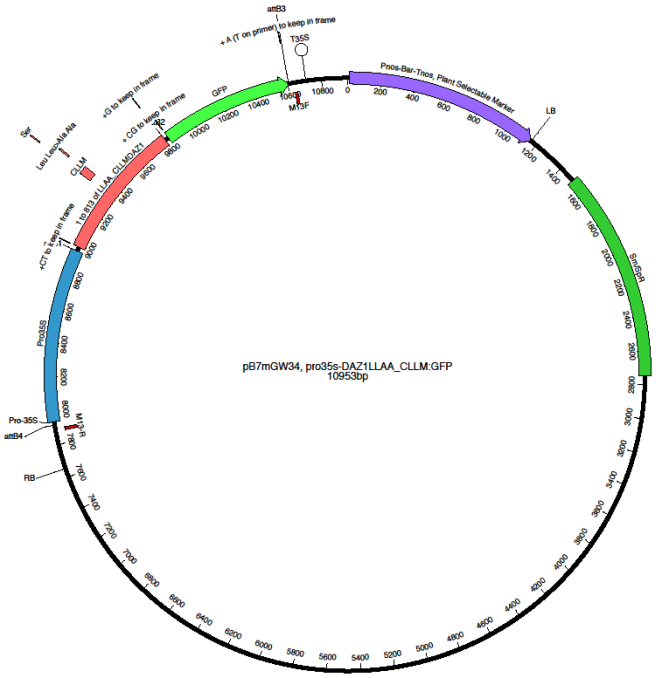
ProDAZ1: DAZ1muBR-GFP (ProDAZ1:muNLSDAZ1ns-GFP)



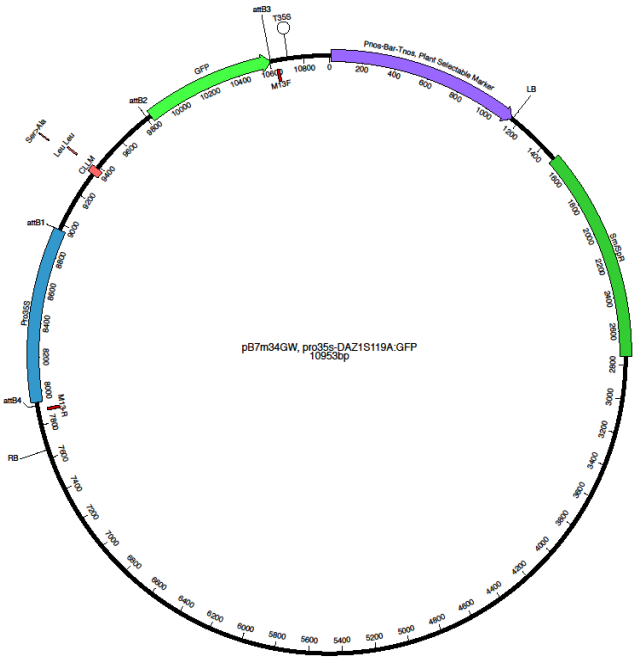
ProDAZ1:DAZ1ΔBR-GFP (ProDAZ1:DAZ1delNLSns-GFP)



ProDAZ1:DAZ1LLAA-CLLM-mCherry



ProDAZ1:DAZ1S119A-GFP



References

Ahmad, M., Jarillo, J. A., Smirnova, O. and Cashmore, A. R. (1998) 'Cryptochrome blue-light photoreceptors of Arabidopsis implicated in phototropism', *Nature*, 392(6677), 720.

Aitchison, J. D. and Wozniak, R. W. (2007) 'Pore puzzle', *Nature*, 450, 621.

Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A. and Rout, M. P. (2007) 'The molecular architecture of the nuclear pore complex', *Nature*, 450, 695.

Albert, V. A., Barbazuk, W. B., Der, J. P., Leebens-Mack, J., Ma, H., Palmer, J. D., Rounsley, S., Sankoff, D., Schuster, S. C. and Soltis, D. E. (2013) 'The Amborella genome and the evolution of flowering plants', *Science*, 342(6165), 1241089.

Alter, O., Brown, P. O. and Botstein, D. (2003) 'Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms', *Proceedings of the National Academy of Sciences*, 100(6), 3351-3356.

Arathi, H., Ganeshaiah, K., Shaanker, R. U. and Hegde, S. (1999) 'Seed abortion in Pongamia pinnata (Fabaceae)', *American Journal of Botany*, 86(5), 659-662.

Aw, S. J., Hamamura, Y., Chen, Z., Schnittger, A. and Berger, F. (2010) 'Sperm entry is sufficient to trigger division of the central cell but the paternal genome

is required for endosperm development in Arabidopsis.', *Development*, 137(16), 2683-2690.

Ba, A. N. N., Pogoutse, A., Provart, N. and Moses, A. M. (2009) 'NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction', *BMC Bioinformatics*, 10(1), 202.

Ballesteros, M. a. L., Bolle, C., Lois, L. M., Moore, J. M., Vielle-Calzada, J.-P., Grossniklaus, U. and Chua, N.-H. (2001) 'LAF1, a MYB transcription activator for phytochrome A signaling', *Genes & development*, 15(19), 2613-2625.

Barnabás, B., Jäger, K. and Fehér, A. (2008) 'The effect of drought and heat stress on reproductive processes in cereals', *Plant, Cell & Environment*, 31(1), 11-38.

Bawa, K. S. and Webb, C. J. (1984) ' Flower, fruit and seed abortion in tropical forest trees: Implications for the evolution of parental and maternal rerproductive patterns', *American Journal of Botany*, 71(5), 736-751.

Becker, J. D., Boavida, L. C., Carneiro, J., Haury, M. and Feijó, J. A. (2003) 'Transcriptional Profiling of Arabidopsis Tissues Reveals the Unique Characteristics of the Pollen Transcriptome', *Plant Physiology*, 133(2), 713-725.

Bencivenga, S., Colombo, L. and Masiero, S. (2011) 'Cross talk between the sporophyte and the megagametophyte during ovule development', *Sexual Plant Reproduction*, 24(2), 113-121.

Berger, F., Grini, P. E. and Schnittger, A. (2006) 'Endosperm: an integrator of seed growth and development', *Current opinion in plant biology*, 9(6), 664-670.

Berger, F. and Twell, D. (2011) 'Germline specification and function in plants', *Annual Review of Plant Biology*, 62, 461-484.

Bhatt, A. M., Lister, C., Page, T., Fransz, P., Findlay, K., Jones, G. H., Dickinson, H. G. and Dean, C. (1999) 'The DIF1 gene of Arabidopsis is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family', *The Plant Journal*, 19(4), 463-472.

Blackmore, S., Wortley, A. H., Skvarla, J. J. and Rowley, J. R. (2007) 'Pollen wall development in flowering plants', *New phytologist*, 174(3), 483-498.

Boisson-Dernier, A., Frietsch, S., Kim, T.-H., Dizon, M. B. and Schroeder, J. I. (2008) 'The peroxin loss-of-function mutation abstinence by mutual consent disrupts male-female gametophyte recognition', *Current Biology*, 18(1), 63-68.

Bonifaci, N., Moroianu, J., Radu, A. and Blobel, G. (1997) 'Karyopherin β2 mediates nuclear import of a mRNA binding protein', *Proceedings of the National Academy of Sciences*, 94(10), 5055-5060.

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.', *The Plant Cell*, 23(2), 534-549.

Borg, M., Brownfield, L. and Twell, D. (2009) 'Male gametophyte development: a molecular perspective', *Journal of Experimental Botany*, 60(5), 1465-1478.

Borg, M., Rutley, N., Kagale, S., Hamamura, Y., Gherghinoiu, M., Kumar, S., Sari, U., Esparza-Franco, M. A., Sakamoto, W. and Rozwadowski, K. (2014) 'An

EAR-dependent regulatory module promotes male germ cell division and sperm fertility in Arabidopsis', *The Plant Cell*, 26(5), 2098-2113.

Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijó, J. A. and Becker, J. D. (2008) 'Comparative Transcriptomics of Arabidopsis Sperm Cells', *Plant Physiology*, 148(2), 1168-1181.

Bowman, J. L., Sakai, H., Jack, T., Weigel, D., Mayer, U. and Meyerowitz, E. M. (1992) 'SUPERMAN, a regulator of floral homeotic genes in Arabidopsis', *Development*, 114(3), 599-615.

Bowman, J. L., Smyth, D. R. and Meyerowitz, E. M. (1989) 'Genes directing flower development in Arabidopsis', *The Plant Cell*, 1(1), 37-52.

Bowman, J. L., Smyth, D. R. and Meyerowitz, E. M. (1991) 'Genetic interactions among floral homeotic genes of Arabidopsis', *Development*, 112(1), 1-20.

Braybrook, S. A., Stone, S. L., Park, S., Bui, A. Q., Le, B. H., Fischer, R. L., Goldberg, R. B. and Harada, J. J. (2006) 'Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis', *Proceedings of the National Academy of Sciences*, 103(9), 3468-3473.

Brockman, H. E., de Serres, F. J., Ong, T.-m., DeMarini, D. M., Katz, A. J., Griffiths, A. J. F. and Stafford, R. S. (1984) 'Mutation tests in Neurospora crassa: A report of the U.S. environmental protection agency gene-tox program', *Mutation Research/Reviews in Genetic Toxicology*, 133(2), 87-134.

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(3), e1000430.

Brownfield, L. and Twell, D. (2009) 'A dynamic DUO of regulatory proteins coordinates gamete specification and germ cell mitosis in the angiosperm male germline', *Plant Signaling & Behavior*, 4(12), 1159-1162.

Calarco, Joseph P., Borges, F., Donoghue, Mark T. A., Van Ex, F., Jullien, Pauline E., Lopes, T., Gardner, R., Berger, F., Feijó, José A., Becker, Jörg D. and Martienssen, Robert A. (2012) 'Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA', *Cell*, 151(1), 194-205.

Caldwell, D. G., McCallum, N., Shaw, P., Muehlbauer, G. J., Marshall, D. F. and Waugh, R. (2004) 'A structured mutant population for forward and reverse genetics in Barley (Hordeum vulgare L.)', *The Plant Journal*, 40(1), 143-150.

Calviño, A. (2014) 'Effects of ovule and seed abortion on brood size and fruit costs in the leguminous shrub Caesalpinia gilliesii (Wall. ex Hook.) D. Dietr', *Acta Botanica Brasilica*, 28, 59-67.

Capron, A., Gourgues, M., Neiva, L. S., Faure, J.-E., Berger, F., Pagnussat, G., Krishnan, A., Alvarez-Mejia, C., Vielle-Calzada, J.-P. and Lee, Y.-R. (2008) 'Maternal control of male-gamete delivery in Arabidopsis involves a putative GPI-anchored protein encoded by the LORELEI gene', *The Plant Cell*, 20(11), 3038-3049.

Cashmore, A. R., Jarillo, J. A., Wu, Y.-J. and Liu, D. (1999) 'Cryptochromes: blue light receptors for plants and animals', *Science*, 284(5415), 760-765.

Causier, B., Ashworth, M., Guo, W. and Davies, B. (2012) 'The TOPLESS Interactome: A Framework for Gene Repression in Arabidopsis', *Plant Physiology*, 158(1), 423-438.

Cebolla, A., Vinardell, J. M., Kiss, E., Olah, B., Roudier, F., Kondorosi, A. and Kondorosi, E. (1999) 'The mitotic inhibitor ccs52 is required for endoreduplication and ploidy-dependent cell enlargement in plants', *The EMBO Journal*, 18(16), 4476-4484.

Chandler, J. W., Cole, M., Flier, A. and Werr, W. (2009) 'BIM1, a bHLH protein involved in brassinosteroid signalling, controls Arabidopsis embryonic patterning via interaction with DORNRÖSCHEN and DORNRÖSCHEN-LIKE', *Plant Molecular Biology*, 69(1-2), 57-68.

Chattopadhyay, S., Ang, L.-H., Puente, P., Deng, X.-W. and Wei, N. (1998) 'Arabidopsis bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression', *The Plant Cell*, 10(5), 673-683.

Chaudhury, A. M., Ming, L., Miller, C., Craig, S., Dennis, E. S. and Peacock, W. J. (1997) 'Fertilization-independent seed development in Arabidopsis thaliana', *Proceedings of the National Academy of Sciences*, 94(8), 4223-4228.

Chen, Y. L., Liang, H. L., Ma, X. L., Lou, S. L., Xie, Y. Y., Liu, Z. L., Chen, L. T. and Liu, Y. G. (2013) 'An Efficient Rice Mutagenesis System Based on Suspension-Cultured Cells', *Journal of Integrative Plant Biology*, 55(2), 122-130.

Ciftci-Yilmaz, S. and Mittler, R. (2008) 'The zinc finger network of plants', *Cellular and Molecular Life Sciences*, 65(7-8), 1150-1160.

Clough, S. J. and Bent, A. F. (1998) 'Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana', *The Plant Journal*, 16(6), 735-743.

Coen, E. S. and Meyerowitz, E. M. (1991) 'The war of the whorls: genetic interactions controlling flower development', *Nature*, 353, 31-37.

Conner, J. and Liu, Z. (2000) 'LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development', *Proceedings of the National Academy of Sciences*, 97(23), 12902-12907.

Conti, E., Müller, C. W. and Stewart, M. (2006) 'Karyopherin flexibility in nucleocytoplasmic transport', *Current opinion in structural biology*, 16(2), 237-244.

Conti, E., Uy, M., Leighton, L., Blobel, G. and Kuriyan, J. (1998) 'Crystallographic Analysis of the Recognition of a Nuclear Localization Signal by the Nuclear Import Factor Karyopherin α', *Cell*, 94(2), 193-204.

Cruzan, M. B. (1989) 'Pollen tube attrition in Erythronium grandiflorum', *American Journal of Botany*, 562-570.

Davletova, S., Schlauch, K., Coutu, J. and Mittler, R. (2005) 'The Zinc-Finger Protein Zat12 Plays a Central Role in Reactive Oxygen and Abiotic Stress Signaling in Arabidopsis', *Plant Physiology*, 139(2), 847-856.

Digonnet, C., Aldon, D., Leduc, N., Dumas, C. and Rougier, M. (1997) 'First evidence of a calcium transient in flowering plants at fertilization', *Development*, 124(15), 2867-2874.

Dingwall, C. and Laskey, R. A. (1991) 'Nuclear targeting sequences—a consensus?', *Trends in Biochemical Sciences*, 16, 478-481.

Dingwall, C., Sharnick, S. V. and Laskey, R. A. (1982) 'A polypeptide domain that specifies migration of nucleoplasmin into the nucleus', *Cell*, 30(2), 449-458.

Doyle, J. J. (1994) 'Evolution of a Plant Homeotic Multigene Family: Toward Connecting Molecular Systematics Andmolecular Developmental Genetics', *Systematic Biology*, 43(3), 307-328.

Duek, P. D. and Fankhauser, C. (2003) 'HFR1, a putative bHLH transcription factor, mediates both phytochrome A and cryptochrome signalling', *The Plant Journal*, 34(6), 827-836.

Dumas, C., Berger, F., Faure, J. and Matthys-Rochon, E. (1998) 'Gametes, Fertilization and Early Embryogenesis in Flowering', *Advances in Botanical Research*, 28, 231-254.

Durbarry, A., Vizir, I. and Twell, D. (2005) 'Male Germ Line Development in Arabidopsis. duo pollen Mutants Reveal Gametophytic Regulators of Generative Cell Cycle Progression', *Plant Physiology*, 137(1), 297-307.

Engel, M. L., Chaboud, A., Dumas, C. and McCormick, S. (2003) 'Sperm cells of Zea mays have a complex complement of mRNAs', *The Plant Journal*, 34(5), 697-707.

Engel, M. L., Holmes-Davis, R. and McCormick, S. (2005) 'Green sperm. Identification of male gamete promoters in Arabidopsis', *Plant Physiology*, 138(4), 2124-2133.

Englbrecht, C. C., Schoof, H. and Böhm, S. (2004) 'Conservation, diversification and expansion of C2H2 zinc finger proteins in the Arabidopsis thaliana genome', *BMC Genomics*, 5(1), 39.

Escobar-Restrepo, J.-M., Huck, N., Kessler, S., Gagliardini, V., Gheyselinck, J., Yang, W.-C. and Grossniklaus, U. (2007) 'The FERONIA receptor-like kinase mediates male-female interactions during pollen tube reception', *Science*, 317(5838), 656-660.

Fairchild, C. D., Schumaker, M. A. and Quail, P. H. (2000) 'HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction', *Genes & Development*, 14(18), 2377-2391.

Feldmann, K. A., Malmberg, R. L. and Dean, C. (1994) 'Mutagenesis in Arabidopsis', *Cold Spring Harbor Monograph Series*, 27, 137-137.

Fischer, A., Baum, N., Saedler, H. and Theiβen, G. (1995) 'Chromosomal mapping of the MADS-box multigene family in Zea mays reveals dispersed distribution of allelic genes as well as transposed copies', *Nucleic Acids Research*, 23(11), 1901-1911.

Franco-Zorrilla, J. M., López-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P. and Solano, R. (2014) 'DNA-binding specificities of plant transcription factors and their potential to define target genes', *Proceedings of the National Academy of Sciences*, 111(6), 2367-2372.

Fu, H. and Dooner, H. K. (2002) 'Intraspecific violation of genetic colinearity and its implications in maize', *Proceedings of the National Academy of Sciences*, 99(14), 9573-9578.

Galstyan, A., Cifuentes-Esquivel, N., Bou-Torrent, J. and Martinez-Garcia, J. F. (2011) 'The shade avoidance syndrome in Arabidopsis: a fundamental role for atypical basic helix–loop–helix proteins as transcriptional cofactors', *The Plant Journal*, 66(2), 258-267.

Garcia-Bustos, J., Heitman, J. and Hall, M. N. (1991) 'Nuclear protein localization', *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 1071(1), 83-101.

Ge, X., Chang, F. and Ma, H. (2010) 'Signaling and transcriptional control of reproductive development in Arabidopsis', *Current Biology*, 20(22), R988-R997.

Gilchrist, E. J., O'Neil, N. J., Rose, A. M., Zetka, M. C. and Haughn, G. W. (2006) 'TILLING is an effective reverse genetics technique for Caenorhabditis elegans', *BMC Genomics*, (7), 262-262.

Görlich, D. and Kutay, U. (1999) 'Transport between the cell nucleus and the cytoplasm', *Annual Review of Cell and Developmental Biology*, 15(1), 607-660.

Görlich, D. and Mattaj, I. W. (1996) 'Nucleocytoplasmic transport', *Science*, 271(5255), 1513-1519.

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

Grossniklaus, U., Spillane, C., Page, D. R. and Köhler, C. (2001) 'Genomic imprinting and seed development: endosperm formation with and without sex', *Current Opinion in Plant Biology*, 4(1), 21-27.

Guth, C. J. and Weller, S. G. (1986) 'Pollination, fertilization and ovule abortion in Oxalis magnifica', *American Journal of Botany*, 73(2), 246-253.

Hall, M. N., Hereford, L. and Herskowitz, I. (1984) 'Targeting of E. coli β-galactosidase to the nucleus in yeast', *Cell*, 36(4), 1057-1065.

Harter, K., Kircher, S., Frohnmeyer, H., Krenz, M., Nagy, F. and Schäfer, E. (1994) 'Light-regulated modification and nuclear translocation of cytosolic G-box binding factors in parsley', *The Plant Cell*, 6(4), 545-559.

Haudry, A., Platts, A. E., Vello, E., Hoen, D. R., Leclercq, M., Williamson, R. J., Forczek, E., Joly-Lopez, Z., Steffen, J. G. and Hazzouri, K. M. (2013) 'An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions', *Nature Genetics*, 45(8), 891.

Higashiyama, T., Yabe, S., Sasaki, N., Nishimura, Y., Miyagishima, S.-y., Kuroiwa, H. and Kuroiwa, T. (2001) 'Pollen tube attraction by the synergid cell', *Science*, 293(5534), 1480-1483.

Honys, D. and Twell, D. (2003) 'Comparative Analysis of the Arabidopsis Pollen Transcriptome', *Plant Physiology*, 132(2), 640-652.

Honys, D. and Twell, D. (2004) 'Transcriptome analysis of haploid male gametophyte development in Arabidopsis', *Genome Biology*, 5(11), R85.

Huang, X.-Y., Chao, D.-Y., Koprivova, A., Danku, J., Wirtz, M., Müller, S., Sandoval, F. J., Bauwe, H., Roje, S. and Dilkes, B. (2016) 'Nuclear localised MORE SULPHUR ACCUMULATION1 epigenetically regulates sulphur homeostasis in Arabidopsis thaliana', *PLOS Genetics*, 12(9), e1006298.

Huck, N., Moore, J. M., Federer, M. and Grossniklaus, U. (2003) 'The Arabidopsis mutant feronia disrupts the female gametophytic control of pollen tube reception', *Development*, 130(10), 2149-2159.

Hudson, M., Ringli, C., Boylan, M. T. and Quail, P. H. (1999) 'The FAR1 locus encodes a novel nuclear protein specific to phytochrome A signaling', *Genes & Development*, 13(15), 2017-2027.

Hudson, M. E., Lisch, D. R. and Quail, P. H. (2003) 'The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway', *The Plant Journal*, 34(4), 453-471.

Jack, T., Brockman, L. L. and Meyerowitz, E. M. (1992) 'The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens', *Cell*, 68(4), 683-697.

Jiao, Y., Lau, O. S. and Deng, X. W. (2007) 'Light-regulated transcriptional networks in higher plants', *Nature Reviews Genetics*, 8(3), 217-230.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and Dean, C. (2000) 'Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time', *Science*, 290(5490), 344-347.

Kagale, S. and Rozwadowski, K. (2011) 'EAR motif-mediated transcriptional repression in plants: An underlying mechanism for epigenetic regulation of gene expression', *Epigenetics*, 6(2), 141-146.

Kalderon, D., Roberts, B. L., Richardson, W. D. and Smith, A. E. (1984) 'A short amino acid sequence able to specify nuclear location', *Cell*, 39(3), 499-509.

Kasahara, Ryushiro D., Maruyama, D., Hamamura, Y., Sakakibara, T., Twell, D. and Higashiyama, T. (2012) 'Fertilization Recovery after Defective Sperm Cell Release in Arabidopsis', *Current Biology*, 22(12), 1084-1089.

Kasahara, R. D., Notaguchi, M. and Honma, Y. (2017) 'Discovery of pollen tube-dependent ovule enlargement morphology phenomenon, a new step in plant reproduction', *Communicative & Integrative Biology*, 10(4), e1338989.

Kasahara, R. D., Notaguchi, M., Nagahara, S., Suzuki, T., Susaki, D., Honma, Y., Maruyama, D. and Higashiyama, T. (2016) 'Pollen tube contents initiate ovule enlargement and enhance seed coat development without fertilization', *Science Advances*, 2(10), e1600554.

Kazan, K. (2006) 'Negative regulation of defence and stress genes by EAR-motif-containing repressors', *Trends in Plant Science*, 11(3), 109-112.

Kellogg, E. A. and Bennetzen, J. L. (2004) 'The evolution of nuclear genome structure in seed plants', *American Journal of Botany*, 91(10), 1709-1725.

Kiang, L., Heichinger, C., Watt, S., Bähler, J. and Nurse, P. (2009) 'Cyclin-dependent kinase inhibits reinitiation of a normal S-phase program during G2 in fission yeast', *Molecular and Cellular Biology*, 29(15), 4025-4032.

Kieffer, M., Stern, Y., Cook, H., Clerici, E., Maulbetsch, C., Laux, T. and Davies, B. (2006) 'Analysis of the Transcription Factor WUSCHEL and Its Functional Homologue in Antirrhinum Reveals a Potential Mechanism for Their Roles in Meristem Maintenance', *The Plant Cell*, 18(3), 560-573.

Kim, Y., Schumaker, K. S. and Zhu, J.-K. (2006) 'EMS mutagenesis of Arabidopsis' in *Arabidopsis Protocols Springer,* 101-103.

Kim, Y. M., Woo, J. C., Song, P. S. and Soh, M. S. (2002) 'HFR1, a phytochrome A-signalling component, acts in a separate pathway from HY5, downstream of COP1 in Arabidopsis thaliana', *The Plant Journal*, 30(6), 711-719.

Kleiner, O., Kircher, S., Harter, K. and Batschauer, A. (1999) 'Nuclear localization of the Arabidopsis blue light receptor cryptochrome 2', *The Plant Journal*, 19(3), 289-296.

Kobayashi, K., Kanno, S.-i., Takao, M., Yasui, A., Smit, B. and van der Horst, G. T. (1998) 'Characterization of photolyase/blue-light receptor homologs in mouse and human cells', *Nucleic Acids Research*, 26(22), 5086-5092.

Köhler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W. and Grossniklaus, U. (2003) 'The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1', *Genes & Development*, 17(12), 1540-1553.

Krieg, D. R. (1963) 'Ethyl methanesulfonate-induced reversion of bacteriophage T4rII mutants', *Genetics*, 48(4), 561-580.

Lalanne, E. and Twell, D. (2002) 'Genetic control of male germ unit organization in Arabidopsis', *Plant Physiology*, 129(2), 865-875.

Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E. and Corbett, A. H. (2007) 'Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin α', *Journal of Biological Chemistry*, 282(8), 5101-5105.

Latta, R. G. (1995) 'The effects of embryo competition with mixed mating on the genetic load in plants', *Heredity*, 75, 637-643.

Lausser, A., Kliwer, I., Srilunchang, K.-o. and Dresselhaus, T. (2009) 'Sporophytic control of pollen tube growth and guidance in maize', *Journal of Experimental Botany*, 61(3), 673-682.

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R. and Horvath, S. (2010) 'Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors', *Proceedings of the National Academy of Sciences*, 107(18), 8063-8070.

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., Horvath, S., Drews, G. N., Fischer, R. L., Okamuro, J. K., Harada, J. J. and Goldberg, R. B. (2010) 'Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors', *Proceedings of the National Academy of Sciences*, 107(18), 8063-8070.

Lee, J.-Y. and Lee, D.-H. (2003) 'Use of Serial Analysis of Gene Expression Technology to Reveal Changes in Gene Expression in Arabidopsis Pollen Undergoing Cold Stress', *Plant Physiology*, 132(2), 517-529.

Leydon, A. R., Chaibang, A. and Johnson, M. A. (2014) 'Interactions between pollen tube and pistil control pollen tube identity and sperm release in the Arabidopsis female gametophyte', *Portland Press Limited,* 42(2), 340-345.

Liljegren, S. J., Ferr, xe, ndiz, C., Alvarez-Buylla, E. R., Pelaz, S. and Yanofsky, M. F. (1998) 'ARABIDOPSIS MADS-BOX GENES INVOLVED IN FRUIT DEHISCENCE', *Flowering Newsletter*, (25), 9-19.

Lin, C., Ahmad, M. and Cashmore, A. R. (1996) 'Arabidopsis cryptochrome 1 is a soluble protein mediating blue light-dependent regulation of plant growth and development', *The Plant Journal*, 10(5), 893-902.

Liu, Z. and Karmarkar, V. (2008) '*Groucho/Tup1 family co-repressors in plant development*', *Trends in Plant Science, 13,*(3), 137-144.

Lloyd, D. G. (1980) 'Sexual strategies in plants', *New Phytologist*, 86(1), 69-79.

Lloyd, D. G., Webb, C. and Primack, R. B. (1980) 'Sexual strategies in plants. II. Data on the temporal regulation of maternal investment', *New phytologist*, 86(1), 81-92.

Long, J. A., Ohno, C., Smith, Z. R. and Meyerowitz, E. M. (2006) 'TOPLESS Regulates Apical Embryonic Fate in Arabidopsis', *Science*, 312(5779), 1520-1523.

Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J. and Chaudhury, A. (2000) 'Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds', *Proceedings of the National Academy of Sciences*, 97(19), 10637-10642.

Maheswari, P. (1950) *An introduction to the embryology of angiosperms*, Tata Mcgraw-Hill Publishing Company Ltd; Bombay; New Delhi.

Marfori, M., Mynott, A., Ellis, J. J., Mehdi, A. M., Saunders, N. F. W., Curmi, P. M., Forwood, J. K., Bodén, M. and Kobe, B. (2011) 'Molecular basis for specificity of nuclear import and prediction of nuclear localization', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1813(9), 1562-1577.

Márton, M.-L. and Dresselhaus, T. (2010) 'Female gametophyte-controlled pollen tube guidance', *Portland Press Limited*, 38(2), 627-630.

Márton, M. L., Fastner, A., Uebler, S. and Dresselhaus, T. (2012) 'Overcoming hybridization barriers by the secretion of the maize pollen tube attractant ZmEA1 from Arabidopsis ovules', *Current Biology*, 22(13), 1194-1198.

Matias-Hernandez, L., Battaglia, R., Galbiati, F., Rubes, M., Eichenberger, C., Grossniklaus, U., Kater, M. M. and Colombo, L. (2010) 'VERDANDI is a direct target of the MADS domain ovule identity complex and affects embryo sac differentiation in Arabidopsis', *The Plant Cell*, 109.068627.

McConnell, J. R., Emery, J., Eshed, Y., Bao, N., Bowman, J. and Barton, M. K. (2001) 'Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots', *Nature*, 411, 709-713.

McCormac, A. C. and Terry, M. J. (2002) 'Light-signalling pathways leading to the co-ordinated expression of HEMA1 and Lhcb during chloroplast development in Arabidopsis thaliana', *The Plant Journal*, 32(4), 549-559.

McCormick, S. (2004) 'Control of male gametophyte development', *The Plant Cell*, 16(1), 142-153.

Mena-AlÍ, J. I. and Rocha, O. J. (2005) 'Selective Seed Abortion Affects the Performance of the Offspring in Bauhinia ungulata', *Annals of Botany*, 95(6), 1017-1023.

Mercier, R., Vezon, D., Bullier, E., Motamayor, J. C., Sellier, A., Lefèvre, F., Pelletier, G. and Horlow, C. (2001) 'SWITCH1 (SWI1): a novel protein required for the establishment of sister chromatid cohesion and for bivalent formation at meiosis', *Genes & Development*, 15(14), 1859-1871.

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

Moede, T., Leibiger, B., Pour, H. G., Berggren, P.-O. and Leibiger, I. B. (1999) 'Identification of a nuclear localization signal, RRMKWKK, in the homeodomain transcription factor PDX-1', *FEBS Letters*, 461(3), 229-234.

Mogensen, H. L. (1975) 'Ovule abortion in Quercus (Fagaceae)', *American Journal of Botany*, 62(2), 160-165.

Mori, T., Kuroiwa, H., Higashiyama, T. and Kuroiwa, T. (2005) 'GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization', *Nature Cell Biology*, 8(1), 64-71.

Neduva, V. and Russell, R. B. (2005) 'Linear motifs: Evolutionary interaction switches', *FEBS Letters*, 579(15), 3342-3345.

Novak, F. and Brunner, H. (1992) 'Plant breeding: Induced mutation technology for crop improvement', *IAEA Bull*, 4, 25-33.

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. and Ohme-Takagi, M. (2001) 'Repression Domains of Class II ERF Transcriptional Repressors Share an Essential Motif for Active Repression', *The Plant Cell*, 13(8), 1959-1968.

Okuda, S., Tsutsui, H., Shiina, K., Sprunck, S., Takeuchi, H., Yui, R., Kasahara, R. D., Hamamura, Y., Mizukami, A. and Susaki, D. (2009) 'Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells', *Nature*, 458(7236), 357-361.

Østergaard, L. and Yanofsky, M. F. (2004) 'Establishing gene function by mutagenesis in Arabidopsis thaliana', *The Plant Journal*, 39(5), 682-696.

Owen, H. A. and Makaroff, C. (1995) 'Ultrastructure of microsporogenesis and microgametogenesis inArabidopsis thaliana (L.) Heynh. ecotype Wassilewskija (Brassicaceae)', *Protoplasma*, 185(1-2), 7-21.

Park, S. K., Howden, R. and Twell, D. (1998) 'The Arabidopsis thaliana gametophytic mutation gemini pollen1 disrupts microspore polarity, division asymmetry and pollen cell fate', *Development*, 125(19), 3789-3799.

Park, S. K., Rahman, D., Oh, S. A. and Twell, D. (2004) 'Gemini pollen 2, a male and female gametophytic cytokinesis defective mutation', *Sexual Plant Reproduction*, 17(2), 63-70.

Payne, T., Johnson, S. D. and Koltunow, A. M. (2004) 'KNUCKLES/(KNU) encodes a C2H2 zinc-finger protein that regulates development of basal pattern elements of the Arabidopsis gynoecium', *Development*, 131(15), 3737-3749.

Philippe, H., Chenuil, A. and Adoutte, A. (1994) 'Can the Cambrian explosion be inferred through molecular phylogeny?', *Development*, 1994, 15-25.

Pina, C., Pinto, F., Feijó, J. A. and Becker, J. D. (2005) 'Gene Family Analysis of the Arabidopsis Pollen Transcriptome Reveals Biological Implications for Cell Growth, Division Control, and Gene Expression Regulation', *Plant Physiology*, 138(2), 744-756.

Poon, I. K. H. and Jans, D. A. (2005) 'Regulation of Nuclear Transport: Central Role in Development and Transformation?', *Traffic*, 6(3), 173-186.

Purugganan, M. D. (1997) 'The MADS-box floral homeotic gene lineages predate the origin of seed plants: Phylogenetic and molecular clock estimates', *Journal of Molecular Evolution*, 45(4), 392-396.

Raff, R. A. (2012) *The Shape of Life: Genes, Development, and the Evolution of Animal Form*, University of Chicago Press.

Richardson, W. D., Mills, A. D., Dilworth, S. M., Laskey, R. A. and Dingwall, C. (1988) 'Nuclear protein migration involves two steps: Rapid binding at the nuclear envelope followed by slower translocation through nuclear pores', *Cell*, 52(5), 655-664.

Rizhsky, L., Davletova, S., Liang, H. and Mittler, R. (2004) 'The Zinc Finger Protein Zat12 Is Required for Cytosolic Ascorbate Peroxidase 1 Expression during Oxidative Stress in Arabidopsis', *Journal of Biological Chemistry*, 279(12), 11736-11743.

Rodrigues, J. C. M., Tucker, M. R., Johnson, S. D., Hrmova, M. and Koltunow, A. M. G. (2008) 'Sexual and Apomictic Seed Formation in Hieracium Requires the Plant Polycomb-Group Gene FERTILIZATION INDEPENDENT ENDOSPERM', *The Plant Cell*, 20(9), 2372-2386.

Ron, M., Saez, M. A., Williams, L. E., Fletcher, J. C. and McCormick, S. (2010) 'Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis', *Genes & Development*, 24(10), 1010-1021.

Ross, K., Fransz, P., Armstrong, S., Vizir, I., Mulligan, B., Franklin, F. and Jones, G. (1997) 'Cytological characterization of four meiotic mutants of Arabidopsis isolated from T-DNA-transformed lines', *Chromosome Research*, 5(8), 551-559.

Roth, S., Stein, D. and Nüsslein-Volhard, C. (1989) 'A gradient of nuclear localization of the dorsal protein determines dorsoventral pattern in the Drosophila embryo', *Cell*, 59(6), 1189-1202.

Rotman, N., Durbarry, A., Wardle, A., Yang, W. C., Chaboud, A., Faure, J.-E., Berger, F. and Twell, D. (2005) 'A Novel Class of MYB Factors Controls Sperm-Cell Formation in Plants', *Current Biology*, 15(3), 244-248.

Rotman, N., Gourgues, M., Guitton, A.-E., Faure, J.-E. and Berger, F. (2008) 'A dialogue between the SIRENE pathway in synergids and the fertilization independent seed pathway in the central cell controls male gamete release during double fertilization in Arabidopsis', *Molecular Plant*, 1(4), 659-666.

Rotman, N., Rozier, F., Boavida, L., Dumas, C., Berger, F. and Faure, J.-E. (2003) 'Female control of male gamete delivery during fertilization in Arabidopsis thaliana', *Current Biology*, 13(5), 432-436.

Rounsley, S. D., Ditta, G. S. and Yanofsky, M. F. (1995) 'Diverse roles for MADS box genes in Arabidopsis development', *The Plant Cell*, 7(8), 1259-1269.

Ruan, Y.-L., Jin, Y., Yang, Y.-J., Li, G.-J. and Boyer, J. S. (2010) 'Sugar Input, Metabolism, and Signaling Mediated by Invertase: Roles in Development, Yield Potential, and Response to Drought and Heat', *Molecular Plant*, 3(6), 942-955.

Ruan, Y.-L., Patrick, J. W., Bouzayen, M., Osorio, S. and Fernie, A. R. (2012) 'Molecular regulation of seed and fruit set', *Trends in Plant Science*, 17(11), 656-665.

Rushlow, C. A., Han, K., Manley, J. L. and Levine, M. (1989) 'The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila', *Cell*, 59(6), 1165-1177.

Rutley, N. (2015) *Molecular control of Arabidopsis male germline development by DAZ1 and DAZ2*, unpublished thesis Department of Genetics.

Rutley, N. and Twell, D. (2015) 'A decade of pollen transcriptomics', *Plant Reproduction*, 28(2), 73-89.

Sakai, H., Medrano, L. J. and Meyerowitz, E. M. (1995) 'Role of SUPERMAN in maintaining Arabidopsis floral whorl boundaries', *Nature*, 378, 199-203.

Sanders, P. M., Bui, A. Q., Weterings, K., McIntire, K., Hsu, Y.-C., Lee, P. Y., Truong, M. T., Beals, T. and Goldberg, R. (1999) 'Anther developmental defects in Arabidopsis thaliana male-sterile mutants', *Sexual Plant Reproduction*, 11(6), 297-322.

Saze, H. (2008) 'Epigenetic memory transmission through mitosis and meiosis in plants', *Seminars in Cell & Developmental Biology*, 19(6), 527-536.

Saze, H., Scheid, O. M. and Paszkowski, J. (2003) 'Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis', *Nature Genetics*, 34, 65-69.

Schmidt, R. J., Ketudat, M., Aukerman, M. J. and Hoschek, G. (1992) 'Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes', *The Plant Cell*, 4(6), 689-700.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H. (1990) 'Genetic Control of Flower Development by Homeotic Genes in Antirrhinum majus', *Science*, 250(4983), 931-936.

Serrat, X., Esteban, R., Guibourt, N., Moysset, L., Nogués, S. and Lalanne, E. (2014) 'EMS mutagenesis in mature seed-derived rice calli as a new method for rapidly obtaining TILLING mutant populations', *Plant Methods*, 10(1), 5-17.

Sherf, B. A., Navarro, S. L., Hannah, R. R. and Wood, K. V. (1996) 'Dual-luciferase reporter assay: an advanced co-reporter technology integrating firefly and Renilla luciferase assays', *Promega Notes*, 57(2), 2-8.

Shi, D.-Q. and Yang, W.-C. (2011) 'Ovule development in Arabidopsis: progress and challenge', *Current Opinion in Plant Biology*, 14(1), 74-80.

Shimizu, K. K. and Okada, K. (2000) 'Attractive and repulsive interactions between female and male gametophytes in Arabidopsis pollen tube guidance', *Development*, 127(20), 4511-4518.

Siddiqi, I., Ganesh, G., Grossniklaus, U. and Subbiah, V. (2000) 'The dyad gene is required for progression through female meiosis in Arabidopsis', *Development*, 127(1), 197-207.

Sikora, P., Chawade, A., Larsson, M., Olsson, J. and Olsson, O. (2011) 'Mutagenesis as a tool in plant genetics, functional genomics, and breeding', *International Journal of Plant Genomics*, 2011, 13.

Slack, J. M. W., Holland, P. W. H. and Graham, C. F. (1993) 'The zootype and the phylotypic stage', *Nature*, 361, 490-492.

Sommer, H., Beltrán, J. P., Huijser, P., Pape, H., Lönnig, W. E., Saedler, H. and Schwarz-Sommer, Z. (1990) 'Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors', *The EMBO Journal*, 9(3), 605-613.

Sparkes, I. A., Runions, J., Kearns, A. and Hawes, C. (2006) 'Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants', *Nature Protocols*, 1, 2019-2025.

Spillane, C., MacDougall, C., Stock, C., Köhler, C., Vielle-Calzada, J., Nunes, S. M., Grossniklaus, U. and Goodrich, J. (2000) 'Interaction of the Arabidopsis polycomb group proteins FIE and MEA mediates their common phenotypes', *Current Biology*, 10(23), 1535-1538.

Stephenson, A. G. (1979) 'An evolutionary examination of the floral display of Catalpa speciosa (Bignoniaceae)', *Evolution*, 33(4), 1200-1209.

Stephenson, A. G. (1984) 'The regulation of maternal investment in an indeterminate flowering plant (Lotus corniculatus)', *Ecology*, 65(1), 113-121.

Steward, R. (1989) 'Relocalization of the dorsal protein from the cytoplasm to the nucleus correlates with its function', *Cell*, 59(6), 1179-1188.

Stewart, M. (2006) 'Structural basis for the nuclear protein import cycle', *Portland Press Limited*, 34(5), 701-704.

Stockwell, C. A., Hendry, A. P. and Kinnison, M. T. (2003) 'Contemporary evolution meets conservation biology', *Trends in Ecology & Evolution*, 18(2), 94-101.

Stuart, J. M., Segal, E., Koller, D. and Kim, S. K. (2003) 'A gene-coexpression network for global discovery of conserved genetic modules', *Science*, 302(5643), 249-255.

Sun, K., Hunt, K. and Hauser, B. A. (2004) 'Ovule Abortion in Arabidopsis Triggered by Stress', *Plant Physiology*, 135(4), 2358-2367.

Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y. and Kurata, N. (2008) 'MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice', *Molecular Genetics and Genomics*, 279(3), 213-223.

Takeuchi, H. and Higashiyama, T. (2012) 'A species-specific cluster of defensin-like genes encodes diffusible pollen tube attractants in Arabidopsis', *PLoS Biology*, 10(12), e1001449.

Tanaka, I. (1988) 'Isolation of generative cells and their protoplasts from pollen ofLilium longiflorum', *Protoplasma*, 142(1), 68-73.

Teichmann, S. A. and Babu, M. M. (2002) 'Conservation of gene co-regulation in prokaryotes and eukaryotes', *TRENDS in Biotechnology*, 20(10), 407-410.

Terzaghi, W. B., Bertekap Jr, R. L. and Cashmore, A. R. (1997) 'Intracellular localization of GBF proteins and blue light-induced import of GBF2 fusion proteins into the nucleus of cultured Arabidopsis and soybean cells', *The Plant Journal*, 11(5), 967-982.

Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Münster, T., Winter, K.-U. and Saedler, H. (2000) 'A short history of MADS-box genes in plants' in Doyle, J. J. and Gaut, B. S., eds., *Plant Molecular Evolution*115-149.

Theißen, G., Kim, J. T. and Saedler, H. (1996) 'Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes', *Journal of Molecular Evolution*, 43(5), 484-516.

Thresher, R. J., Vitaterna, M. H., Miyamoto, Y., Kazantsev, A., Hsu, D. S., Petit, C., Selby, C. P., Dawut, L., Smithies, O. and Takahashi, J. S. (1998) 'Role of mouse cryptochrome blue-light photoreceptor in circadian photoresponses', *Science*, 282(5393), 1490-1494.

Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S. and Comai, L. (2007) 'Discovery of chemically induced mutations in rice by TILLING', *BMC Plant Biology*, 7, 19-19.

Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C. A., Enns, L. C., Odden, A. R., Greene, E. A., Comai, L. and Henikoff, S. (2004) 'Discovery of induced point mutations in maize genes by TILLING', *BMC Plant Biology*, 4, 12-12.

Tinland, B., Koukolíková-Nicola, Z., Hall, M. N. and Hohn, B. (1992) 'The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization

signals', *Proceedings of the National Academy of Sciences*, 89(16), 7442-7446.

Tsukamoto, T., Qin, Y., Huang, Y., Dunatunga, D. and Palanivelu, R. (2010) 'A role for LORELEI, a putative glycosylphosphatidylinositol-anchored protein, in Arabidopsis thaliana double fertilization and early seed development', *The Plant Journal*, 62(4), 571-588.

Twell, D. (2011) 'Male gametogenesis and germline specification in flowering plants', *Sexual Plant Reproduction*, 24(2), 149-160.

Twell, D., Oh, S.-A. and Honys, D. (2006) 'Pollen Development, a Genetic and Transcriptomic View' in *Springer, Berlin, Heidelberg,* 15-45.

Twell, D., Park, S. K., Hawkins, T. J., Schubert, D., Schmidt, R., Smertenko, A. and Hussey, P. J. (2002) 'MOR1/GEM1 has an essential role in the plant-specific cytokinetic phragmoplast', *Nature Cell Biology*, 4(9), 711.

Twell, D., Park, S. K. and Lalanne, E. (1998) 'Asymmetric division and cell-fate determination in developing pollen', *Trends in Plant Science*, 3(8), 305-310.

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T. C., McElver, J., Aux, G., Patton, D. and Meinke, D. (2004) 'Identification of Genes Required for Embryo Development in Arabidopsis', *Plant Physiology*, 135(3), 1206-1220.

van der Krol, A. R. and Chua, N.-H. (1991) 'The basic domain of plant B-ZIP proteins facilitates import of a reporter protein into plant nuclei', *The Plant Cell*, 3(7), 667-675.

van Noort, V., Snel, B. and Huynen, M. A. (2003) 'Predicting gene function by conserved co-expression', *TRENDS in Genetics*, 19(5), 238-242.

Varagona, M. J., Schmidt, R. J. and Raikhel, N. V. (1992) 'Nuclear localization signal (s) required for nuclear targeting of the maize regulatory protein Opaque-2', *The Plant Cell*, 4(10), 1213-1227.

Vasil, V., Marcotte, W. R., Rosenkrans, L., Cocciolone, S. M., Vasil, I. K., Quatrano, R. S. and McCarty, D. R. (1995) 'Overlap of Viviparous1 (VP1) and abscisic acid response elements in the Em promoter: G-box elements are sufficient but not necessary for VP1 transactivation', *The Plant Cell*, 7(9), 1511-1518.

Vielle-Calzada, J.-P., Hernández-Lagana, E., Rodríguez-Leal, D., Rodríguez-Arévalo, I., León-Martínez, G., Abad-Vivero, U., Demesa-Arévalo, E., Armenta-Medina, A. and Alvarez-Mejía, C. (2012) 'Reproductive versatility and the epigenetic control of female gametogenesis', in *Cold Spring Harbor Laboratory Press,* 17-21.

Vielle-Calzada, J.-P., Thomas, J., Spillane, C., Coluccio, A., Hoeppner, M. A. and Grossniklaus, U. (1999) 'Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity', *Genes & Development*, 13(22), 2971-2982.

Vogel, J. T., Zarka, D. G., Van Buskirk, H. A., Fowler, S. G. and Thomashow, M. F. (2005) 'Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of Arabidopsis', *The Plant Journal*, 41(2), 195-211.

von Arnim, A. G. and Deng, X.-W. (1994) 'Light inactivation of arabidopsis photomorphogenic repressor COP1 involves a cell-specific regulation of its nucleocytoplasmic partitioning', *Cell*, 79(6), 1035-1045.

Wang, H. and Deng, X. W. (2002) 'Arabidopsis FHY3 defines a key phytochrome A signaling component directly interacting with its homologous partner FAR1', *The EMBO Journal*, 21(6), 1339-1349.

Wang, H., Ma, L., Habashi, J., Li, J., Zhao, H. and Deng, X. W. (2002) 'Analysis of far-red light-regulated genome expression profiles of phytochrome A pathway mutants in Arabidopsis', *The Plant Journal*, 32(5), 723-733.

Wang, L., Zhang, B., Li, J., Yang, X. and Ren, Z. (2014) 'Ethyl methanesulfonate (EMS)-mediated mutagenesis of cucumber (Cucumis sativus L.)', *Agricultural Sciences*, 5(8), 716-721.

Weber, H., Borisjuk, L. and Wobus, U. (2005) 'MOLECULAR PHYSIOLOGY OF LEGUME SEED DEVELOPMENT', *Annual Review of Plant Biology*, 56(1), 253-279.

Weigel, D. and Meyerowitz, E. M. (1994) 'The ABCs of floral homeotic genes', *Cell*, 78(2), 203-209.

Weis, K. (2003) 'Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle', *Cell*, 112(4), 441-451.

Willson, M. F. and Price, P. W. (1980) 'Resource limitation of fruit and seed production in some Asclepias species', *Canadian Journal of Botany*, 58(20), 2229-2233.

Wilson, Z. A. and Yang, C. (2004) 'Plant gametogenesis: conservation and contrasts in development', *Reproduction*, 128(5), 483-92.

Xin, Z., Li Wang, M., Barkley, N. A., Burow, G., Franks, C., Pederson, G. and Burke, J. (2008) 'Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population', *BMC Plant Biology*, 8, 103-103.

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(11), 1192-1201.

Yang, K.-Y., Kim, Y.-M., Lee, S., Song, P.-S. and Soh, M.-S. (2003) 'Overexpression of a mutant basic helix-loop-helix protein HFR1, HFR1-ΔN105, activates a branch pathway of light signaling in Arabidopsis', *Plant Physiology*, 133(4), 1630-1642.

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A. and Meyerowitz, E. M. (1990) 'The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors', *Nature*, 346, 35-39.

Yi, Y. and Jack, T. (1998) 'An intragenic suppressor of the Arabidopsis floral organ identity mutant apetala3-1 functions by suppressing defects in splicing', *The Plant Cell*, 10(9), 1465-1477.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W. (2004) 'GENEVESTIGATOR. Arabidopsis Microarray Database and Analysis Toolbox', *Plant Physiology*, 136(1), 2621-2632.

Zuber, H., Noguero, M., Le Signor, C., Thompson, R. and Gallardo, K. (2012) 'Metabolic Specialization of Maternal and Filial Tissues' in Agrawal, G. K. and Rakwal, R., eds., *Springer, Dordrecht, Netherlands,* 407-432.