Nkrumah-Buansi, Martha (2013) Analysis of tapetally expressed genes during Arabidopsis thaliana pollen development. PhD thesis, University of Nottingham.

ANALYSIS OF TAPETALLY EXPRESSED GENES DURING ARABIDOPSIS THALIANA POLLEN DEVELOPMENT

Martha Nkrumah-Buansi, MSc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

JULY 2013

ABSTRACT

The formation of viable pollen relies upon a complex interaction of genes in time and space within the anther. One of the most important maternal tissues involved in the production of functional pollen is the tapetum, which is a highly active tissue that plays a major secretory role during pollen development. This project involved the molecular analysis of genes that are expressed in the anther tapetum and are critical for functional pollen development. A number of these are thought to be regulated by, or interact with MALESTERILITY1 (MS1), a transcriptional regulator of male gametogenesis (Yang *et al*, 2007) or ABORTED MICROSPORE (Xu *et al,* 2010). Work involved analysis of an ABC transporter (At3g13220), which has been shown to be critical for viable pollen formation and confirmed as directly regulated by AMS (Xu *et al*., 2010). Another protein, POB2, which appears to be involved in ubiquitin-based proteolytic breakdown, is thought to interact with the MS1 protein. POB2 was identified from a previous screen of a stamen specific yeast-2-hybrid library using the MS1 protein. This interaction has been subsequently confirmed in this work by further yeast two hybrid analyses and bifunctional fluorescent complementation. Further work involved verification of this interaction *in vitro* and *in planta* by pull-downs and transient expression of proteins in *E. coli* and *Nicotiana benthamiana* respectively. Other work focused on identifying factors that regulate MS1 expression; this identified novel male sterile mutants derived from screening fast neutron mutagenised seed carrying the MS1Prom:MS1-GFP functional fusion protein. Microscopic observation of the fluorescent reporter showed changes in the stage specific expression of MS1 in some of these mutants. Backcrossing of the male sterile mutants with the parental plants (carrying the MS1Prom:MS1-

GFP fusion construct) and the *ms1* mutant confirmed one as a new mutant and the other three as being allelic to the *ms1* mutation. Gene mapping of this mutant was subsequently conducted and suggest that it may be located on chromosome 3. These results are providing insight into the regulatory network of MS1 and AMS during anther development.

LIST OF PUBLISHED PAPERS

Xu, J., Yang, C., Yuan, Z., Zhang, D., Gondwe, M.Y., Ding, Z., Liang, W., Zhang, D. and Wilson, Z. A. (2010) The ABORTED MICROSPORES regulatory network is required for postmeiotic male reproductive development in Arabidopsis thaliana. Plant Cell **22**: 91–107.

ACKNOWLEDGEMENTS

I would like to express my gratitude and thanks to my supervisor Prof. Zoe Wilson, for all of her help, support and patience. I have learnt so many numerous skills from her over the course of my PhD, which have laid down the foundations for my future career.

I would also like to thank Dr. Caiyun Yang for her help and advice during my studies. I must thank all the members of Zoe Wilson's lab group, past and present, for their sincere advice and help.

Finally big thanks to my family and friends whose support was invaluable and has seen me through to the end. With a special thanks to the International Office University of Nottingham, for the Alumni Scholarship which contributed to the financial support and completion of my thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2-DE: 2-dimensional gel electrophoresis

AAT: amino acid transport and metabolism.

ABC: ATP Binding Cassette

ABRC: *Arabidopsis* Biological Resource Centre

AD: Activation Domain

ADF: Activation Domain Forward

ADR: Activation Domain Reverse

AG: Affymetrix Gene chip

AIMS: Arabidopsis Information Management System

AtDB: Arabidopsis thaliana Database

ATP: Adenosine Tri-Phosphate

bHLH-ZIP: basic-Helix–Loop–Helix leucine -Zipper domain

BiFC: Bimolecular Fluorescence Complementation

BLAST: Basic Local Alignment Search

BTB: BR-C, ttk and bab

bZIP: basic Zipper

CaCl$_2$: Calcium Chloride

CAGE: cap analysis of gene expression

CAT: Carbohydrate Transport and Metabolism;

cDNA: complementary Deoxyribonucleic Acid

ChIP: Chromatin immunoprecipitation

CSHL: Cold Spring Harbor Laboratory

DB: DNA-Binding

dNTP: Deoxynucleotide triphosphate

DTT: Dithiothreitol

dUTP: Deoxyuridine triphosphate

EDTA: Ethylenediaminetetraacetic acid

EMS: Ethane Methyl Sulphonate

ERF: Ethylene-Responsive element binding Factors

ESI-MS/MS: Electrospray Ionization Tandem Mass Spectrometry/Magnetic Spectrometry

ESSA: European Scientists Sequencing *Arabidopsis*

EST: Expressed Sequence Tags

GEBD: Genome Express Browser Server

GFP: Green Fluorescent Protein

GM: Genetically Modified

GMO: Genetically Modified Organisms

GN: Genomic DNA

GST Glutathione-S-transferase

HCl: Hydrochloric acid

HM: Homozygous

HMW: High Molecular Weight

HT: Heterozygous

ID: Identity

JA: Jasmonic Acid

KCL: Potassium Chloride

KO: Knock-Out

LB: Luria Broth

LCS: Leica Control Software

MIPS: Munich Information Center for Protein Sequences

MIRNA: Micro RNA

M-MLV RT: Moloney-Murine Leukemia Virus Reverse Transcriptase

mRNA: messenger Ribonucleic Acid

MS: Mass Spectrometry

MS: Murashige and Skoog Basal Medium

NASC: Nottingham *Arabidopsis* Stock Centre

P120: Proliferating cell nuclear protein P120

P450: Proliferating cell nuclear protein 450

PBS: Phosphate buffered saline

PCD: Programmed Cell Death

PCR: Polymerase Chain Reaction

PCR: Polymerase Chain Reaction

PDA: Potato Dextrose Agar

PDS: Particle Delivery System

PHD: Plant Homeo Domain

POZ: Poxvirus and Zinc finger

PPI: Protein-Protein Interaction

PSV: Protein Sorting Vacuole

PTGS: Post-Transcriptional Gene Silencing

RD: Responsive Dehydration

RFLP: Restriction Fragment Length Polymorphism

RLT: Rapid Lysis Tissue

RNA: Ribonucleic Acid

RNAi: Ribonucleic Acid Interference

ROS: Reactive Oxygen species

RPE: RNA Pure Extraction

RT: Reverse Transcriptase

RT-PCR: Reverse Transcriptase-Polymerase Chain Reaction

SAGE: Serial Analysis of Gene Expression

SAIL: Syngenta Arabidopsis Insertion Library

SD: Synthetic Defined

SDS: Sodium Dodecyl Sulfate

SDS-PAGE: Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

SDW: Sterile Distilled Water

SIGNAL: Salk Institute Genomic Analysis Laboratory

SJTU: Shanghai Jiao Tong University

SNP: Single Nucleotide Polymorphism

SSR: Simple Sequence Repeats

TAE: Trisacetate

TAIR: The Arabidopsis Information Resource

TBSV: Tomato Bushy Stunt Virus

T-DNA: Transferred-Deoxyribonucleic Acid

TF: Transcription Factors

TIGR: The Institute for Genomic Research

TILLING: Targeted Induced Local Lesions in Genomes

TM: Trade Mark

UAS: Upstream Activating Sequence

UK: United Kingdom

UoN: University of Nottingham

USA: United States of America

USDA: United States Department of Agriculture

UTR: Untraslated Regions

UV: Ultraviolet

VLCFA: Very Long Chain Fatty Acids

Wt: Wild type

Y2H: Yeast-2-Hybrid

YFP: Yellow Fluorescent Protein

YPD: Yeast extract Peptone Dextrose

CHAPTER 1 INTRODUCTION

1. INTRODUCTION

1.1 *ARABIDOPSIS THALIANA*: A MODEL PLANT FOR RESEARCH IN PLANT BIOLOGY

Arabidopsis thaliana is a small plant in the mustard family that has become the model system of choice for research in plant biology (Meinke *et al*., 1998). By focusing on the molecular genetics of this simple angiosperm, significant advances in understanding plant growth and development have been made for a wide range of studies in plant biology (Ariizumi and Toriyama, 2011). The 120-megabase genome of *Arabidopsis* is organized into five chromosomes (*www.arabidopsis.org*). Having its entire genome sequenced has enhanced the value of *Arabidopsis* as a model for plant biology and the analysis of complex organisms in general.

A growing realization among biologists working on *Arabidopsis* research is the fact that this simple angiosperm can serve as a convenient model not only for plant biology but also for addressing fundamental questions of biological structure and function common to all eukaryotes (Meinke *et al*., 1998). More recently the ability to translate information from Arabidopsis to crops has been enhanced by the release of different genome sequences, for example, *Brachypodium,* tomato, barley and maize (Nussbaumer *et al*., 2012).

Research with *Arabidopsis* has also clarified the important role that analysis of plant genomes can play in understanding basic principles of biology relevant to a

variety of species, including humans (Meinke *et al.*, 1998). The value of *Arabidopsis* for research in plant physiology, biochemistry, and development has been enhanced by significant advances involving the establishment of transformation protocols for detailed molecular analysis (Karin and Jürgen 2010). Since *Arabidopsis* was first adopted as a model plant its genome has been saturated with mutations identifying theoretically every gene (Bouchez and Höfte 1998).

1.1.1   THE BIOLOGY OF ARABIDOPSIS

Arabidopsis thaliana has a broad natural distribution throughout Europe, Asia, and North America (Meyerowitz and Somerville, 1994). Many different ecotypes have been collected from natural populations for experimental analysis. However, the Columbia and Landsberg ecotypes have widely been accepted as standards in genetic and molecular studies. The entire life cycle which includes seed germination, formation of a rosette plant, bolting of the main stem, flowering, and maturation of the first seeds, is completed in 6 weeks. When it comes to size, almost everything about *Arabidopsis* is small. Flowers are 2 mm long, self-pollinate as the bud opens, and can be crossed by applying pollen to the stigma surface. Seeds are 0.5 mm in length at maturity and are produced in slender fruits known as siliques. Seedlings develop into rosette plants that range from 2 to 10 cm in diameter, depending on growth conditions. Leaves are covered with small unicellular hairs known as trichomes that are convenient models for studying morphogenesis and cellular differentiation. *Arabidopsis* plants can also be grown in petri plates or maintained in pots located either in a greenhouse or under

fluorescent lights in the laboratory. Bolting starts about 3 weeks after planting, and the resulting inflorescence forms a linear progression of flowers and siliques for several weeks before the onset of senescence (Meyerowitz and Somerville, 1994). Flowers are composed of an outer whorl of four green sepals and inner whorls containing four white petals, six stamens bearing pollen, and a central gynoecium that forms the silique (Figure 1.1). The roots are simple in structure, easy to study in culture, and do not establish symbiotic relationships with nitrogen-fixing bacteria. Natural pathogens include a variety of insects, bacteria, fungi, and viruses (Meyerowitz and Somerville, 1994).
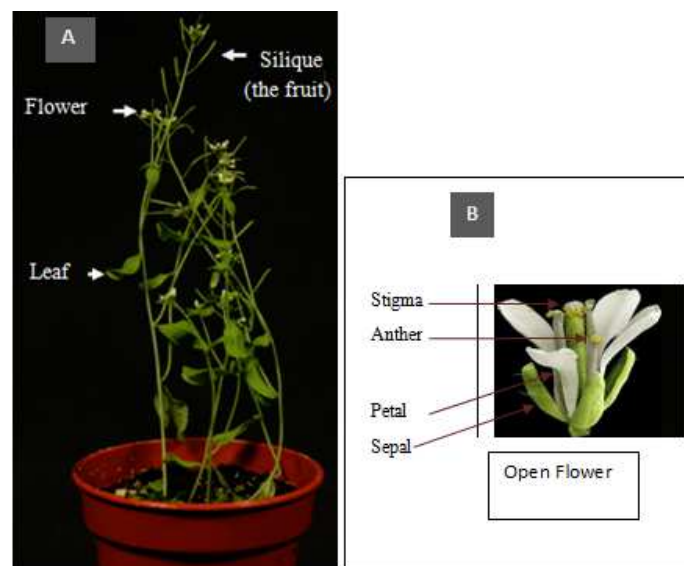
Figure 1.1 (**A**) *A. thaliana* plant at an early stage of flowering has the roots, young buds, stem, mature leaves, closed and open flowers, the inflorescences and siliques. Its developing flower has four basic organs: sepals, petals, stamens (anthers and filament), and carpels (stigma and style) (**B**). These organs are arranged in a series of whorls: four sepals on the outer whorl, followed by four petals inside this, six stamens and a central carpel region. The long fruit called the silique harbours the seeds upon successful fertilisation.

1.2 THE IMPORTANCE OF STUDYING PLANT REPRODUCTION

The study of plant development and the genetic engineering of genes is increasingly become important as more than 90% of the world's land is unsuitable for growing crops. Therefore there is increasing importance in the study of traits that may enable growing areas to be extended, for example nutrient uptake, drought resistance, heat resistance, cold tolerance, salt tolerance, shade tolerance and disease resistance. Plant fertility is of importance since it is frequently vital for crop yield, since fertilisation is a prerequisite for seed formation for most crops. The rise in global population has put a demand on more food production leading to issues of food security (The Future of Food and Farming 2011; Food and Agriculture Organization of the United Nations (FAO-UN) 2008; The Royal Society, Reaping the benefits: Science and the sustainable intensification of global agriculture 2009). Consequently genetic engineering provides a way of potentially solving these problems by using a variety of approaches to identify genes and processes that will help to significantly increase crop yields and food production. Through plant reproduction, developmental traits such as seed size, growth rate, organ size and number (more seeds), plant architecture, flowering time, senescence, maturity and stature can be improved for increasing yields of high quality food in a minimum amount of time, at a reasonable cost. The use of mutant screens to identify genes associated with gamete formation has also provided access to novel information regarding the processes of reproduction in higher plants (Xu *et al*., 2010; Chen *et al*., 2011; Ariizumi and Toriyama, 2011; Ma *et al*, 2012).

1.3 ARABIDOPSIS ANTHER AND POLLEN DEVELOPMENT

The process of male gamete formation in plants involves a series of events that lead to the production and release of mature pollen grains from the anther (Ma, 2005). The anther contains both reproductive and non-reproductive (somatic) cells (Figure 1.2) (Ma, 2005).
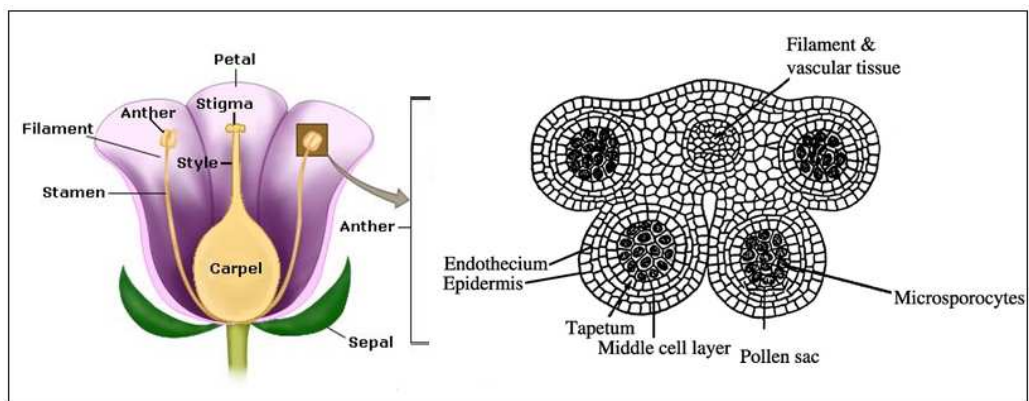
Figure 1.2 The floral structure of Arabidopsis gametophyte (tutorvista.com) and a anther cross section of Arabidopsis anther (Wilson and Yang, 2004)

The stamen is the pollen-producing reproductive organ of a flower. Stamens typically consist of a stalk called the filament and an anther that contains microsporangia. Anthers are most commonly two-lobed and are attached to the filament either at the base or in the middle portion. The sterile tissue between the lobes is called the connective. A typical anther contains four microsporangia. The microsporangia form sacs or pockets (locules) in the anther. The two separate locules on each side of an anther may fuse into a single locule. Each microsporangium contains four separate maternal cell layers, the epidermis, endothecium, middle cell layer, tapetum, which surround the central

6

microsporocytes, or pollen mother cells (PMCs). These PMCs undergo meiosis to form haploid spores. The spores may remain attached to each other in a tetrad or separate after meiosis. Each microspore then divides mitotically to form an immature microgametophyte called a pollen grain. The pollen is eventually released by the opening (dehiscence) of the anther. Anther development has been divided into 14 stages (Sanders *et al.*, 1999). At stage 1, the anther contains three cell layers, L1, L2 and L3 (Figure1.3). The L1 and L2 layers form the epidermis and archesporial cells, respectively, and the L3 layer gives rise to the vascular and connective tissues.

At stage 2, the archesporial cells divide further into primary parietal and primary sporogenous cells to form the stage 3 anthers. The primary parietal cell layer divides again to form two layers of secondary parietal cells. At stage 4, further division and differentiation of secondary parietal cells generate the endothecium, middle layer and tapetum. The primary sporogenous tissue gives rise to microsporocytes (pollen mother cells) at stage 5, and these undergo meiosis to form microspores during stages 6–8 (Figure 1.3).

From stages 9-14 (Figure 1.4) the microspores differentiate into pollen grains, the filament elongates, the anther enlarges and expands, cell degeneration occurs, and the anther enters a dehiscence program that ends with flower opening. Dehiscence results in anther wall breakage at the stomium region located between the two locules of each anther half, or theca, and the release of pollen grains for subsequent pollination and fertilization (Sanders *et al.*, 1999).
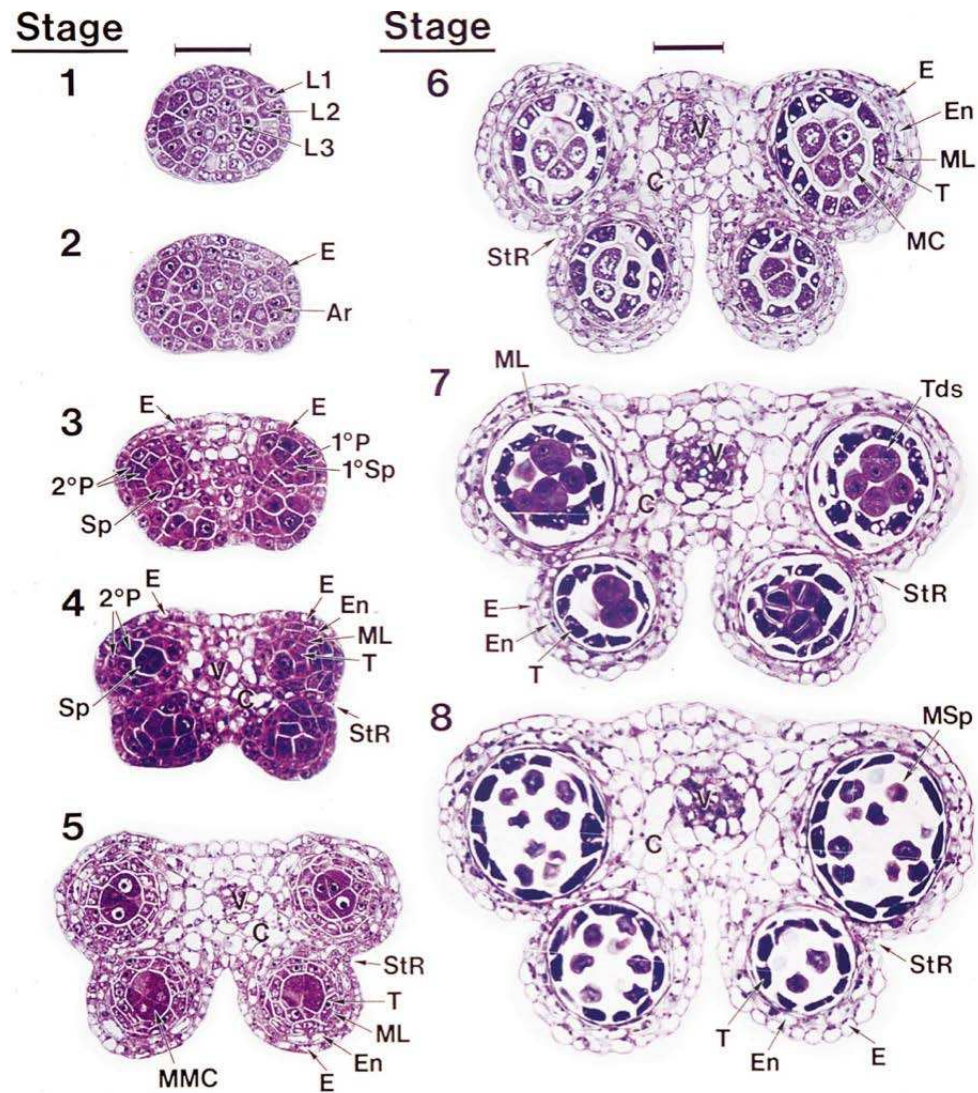
Figure 1.3 Phase one of wild-type *Arabidopsis thaliana* anther development (Sanders *et al.*, 1999). Transverse sections of *Arabidopsis* anther development from stages 1-8. *Ar*, Archesporial cell; *C*, connective; *E*, epidermis; *En*, endothecium; *L1*, *L2*, and *L3*, the three cell-layers in stamen primordia; *MC*, meiotic cell; *ML*, middle layer; *MMC*, microspore mother cells; *MSp*, microspores; *1°P*, primary parietal layer; *2°P*, secondary parietal cell layers; *1°Sp*, primary sporogenous layer; *Sp*, sporogenous cells; *StR*, stomium region; *T*, tapetum; *Tds*, tetrads; *V*,vascular region. *Bar* over stage 1=25 mm and this is the scale for stages 1 to 4. *Bar* over stage 6=25 mm and this is the scale for stages 5 to 8.
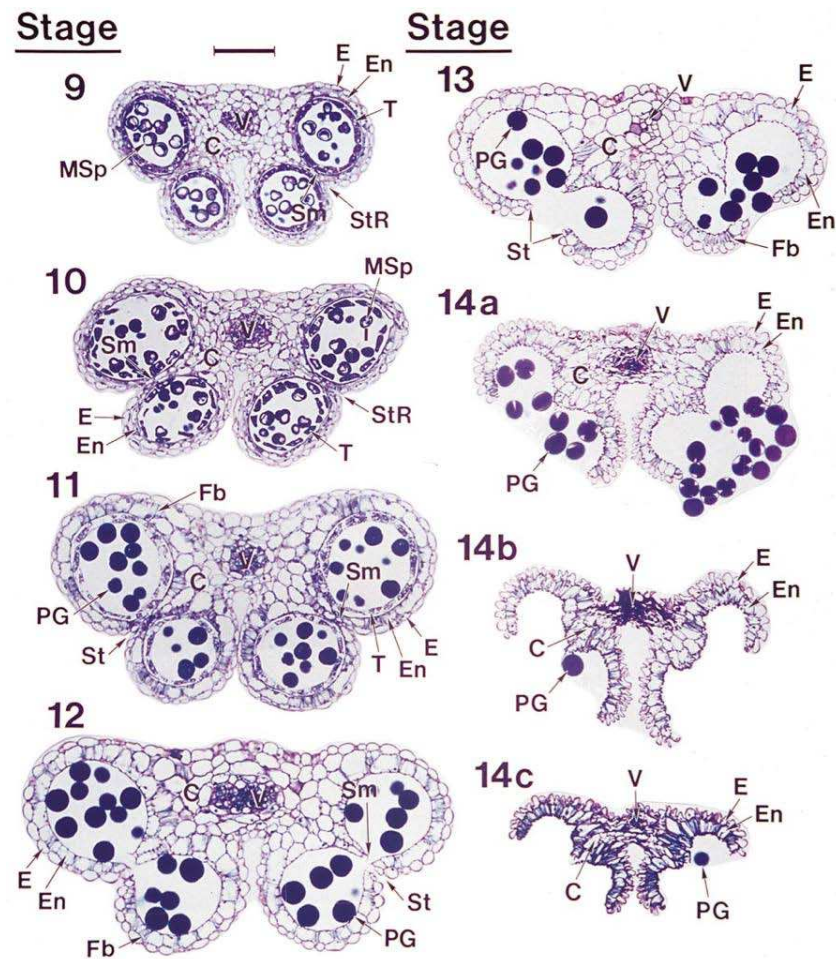
Figure 1.4 Phase two of wild-type *Arabidopsis thaliana* anther development (Sanders *et al*, 1999). From stages 9-14 the microspores differentiate into pollen grains, the filament elongates, the anther enlarges and expands, cell degeneration occurs, and the anther enters a dehiscence program that ends with flower opening. Dehiscence results in anther wall breakage at the stomium region located between the two locules of each anther half, or theca, and the release of pollen grains for subsequent pollination and fertilization. Transverse sections of *Arabidopsis* anther showing stages 9 to 11 (late anther development), 12 to 13 (dehiscence), and 14a to 14c (senescence). *C*, connective; *E*, epidermis; *En*, endothecium; *Fb*, fibrous bands; *MSp*, microspores; *PG*, pollen grains; *Sm*, septum; *St*, stomium; *StR*, stomium region; *T*, tapetum; *V*, vascular region. *Bar*=50mm.

9

1.4 KEY GENES INVOLVED DURING ARABIDOPSIS ANTHER AND POLLEN DEVELOPMENT

The production and release of function pollen from anthers is a complex developmental process involving a diverse range of gene interactions. During early anther development many genes that impact upon male meiosis have now been characterized, but only a few genes that affect the later stages of development during tapetal maturation have been identified (Wilson and Zhang, 2009). A number of genes defective at specific stages during male gametogenesis have previously been characterized and their mutants are being used as tools to investigate anther and pollen development (Sanders *et al*., 1999; Scott *et al*., 2004; Ma, 2005; Wilson and Zhang, 2009; Feng and Dickinson, 2010). Wilson *et al*., (2011) has shown some of the key genes regulating the subsequent development of the different cell layers in anther and pollen formation (Figure 1.5).
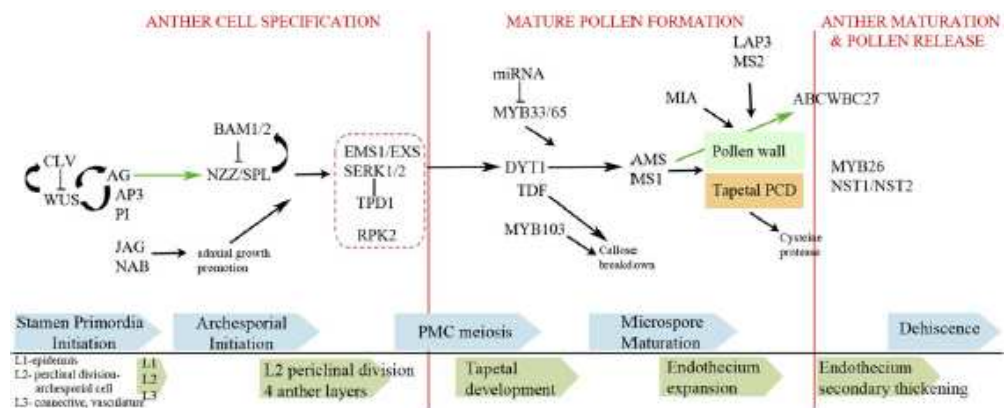
Figure 1.5 Anther and pollen development pathway (from Wilson *et al*., 2011) detailing a number of key genes involved in the regulation of male gametogenesis. Regulation is indicated by arrows. Green arrows indicate proven direct regulation; inhibition is shown by a T-bar.

A database that collates mutant and gene data has also recently been developed for plant male reproduction (PMRD, 2012). In Figure 1.5 it is shown that the initiation of floral organ primordia from pleuripotent cells within the floral meristem involves *AGAMOUS* (*AG*), alongside other homeotic genes to form organ initials according to the ABCDE model (Robles and Pelaz, 2005). AGAMOUS (AG) is activated by WUSCHEL (WUS) at the centre of the floral apex leading to the increase in the number of pluripotent cells and then AG represses WUS to promote differentiation. Stamen initiation is controlled by the homeotic genes *APETALA3 (AP3)*, *PISTILLATA* (*PI*), and *AG*, with the primordia forming as a tetrad of archesporial cells. AG induces microsporogenesis via activation of *NOZZLE*/*SPOROCYTELESS* (*NZZ*/*SPL*) (Ito *et al*., 2004). The transcription factors JAGGED (JAG) and NUBBIN (NAB) are also involved in the process of defining stamen structure (Dinneny *et al*., 2006). Two CLAVATA1-related leucine-rich repeat receptor-like protein kinases, BARELY ANY MERISTEM1 (BAM1) and BAM2, act in a regulatory loop with *NZZ*/*SPL* to promote somatic cell types and to restrict *NZZ*/*SPL* expression to the inner region of the locule (Hord *et al*., 2006; Feng and Dickinson, 2010). Archesporial cell number and tapetal cell fate is controlled by a leucine-rich repeat receptor kinase EXTRA SPOROGENOUS CELLS/EXCESS MICROSPOROCYTES1 (EXS/EMS1) (Canales *et al*., 2002; Zhao *et al*., 2002) and its ligand *TAPETAL DETERMINANT1* (*TPD1*) (Jia *et al*., 2008). The SERK1 and SERK2 complex is thought to form a receptor complex with EMS1 in the tapetal plasma membrane [Albrecht *et al.,* 2005; Colcombet *et al*., 2005], which then binds TPD1 (Yang *et al*., 2003, 2005) (indicated by the red dotted line). Tapetal development is initiated by DYSFUNCTIONAL TAPETUM1 (DYT1) (Zhang *et al*., 2006) and

DEFECTIVE IN TAPETAL DEVELOPMENT AND FUNCTION1 (TDF1) (*Zhu et al*., 2008), with tapetal maturation, pollen wall formation, and tapetal programmed cell death (PCD) involving *ABORTED MICROSPORES* (*AMS*), (Sorensen *et al*., 2003; Yang *et al*., 2007; Xu *et al*., 2010 *MALE STERILITY1* (*MS1*) (Wilson *et al*., 2001).

Tapetal programmed cell death (PCD) is essential for the development of functional pollen with many of the male sterile mutants showing either altered timing of tapetal PCD (Li *et al*., 2006), or a switch towards necrotic tapetal breakdown (Vizcay-Barrena and Wilson, 2006). Niu *et al*., (2013) have recently reported that ETERNAL TAPETUM 1, a basic helix-loop-helix transcription factor that is conserved in land plants, positively regulates programmed cell death in tapetal cells in rice anthers. They demonstrated that the *eat1* mutant exhibits delayed tapetal cell death and aborted pollen formation. They further revealed that EAT1 directly regulates the expression of OsAP25 and OsAP37, which encode aspartic proteases that induce programmed cell death in both yeast and plants.

The final stage of dehiscence involves jasmonic acid (JA)-induced gene expression and transcription factors associated with endothecium secondary thickening, MYB26 (Steiner-Lange *et al*., 2003; Yang *et al*., 2007) and the *NAC SECONDARY WALL THICKENING PROMOTING FACTOR1* (*NST1*) and *NST2* double mutant (Mitsuda *et a*l., 2007). Recently Ariizumi and Toriyama, (2011) reviewed more genes which are involved in precise primexine formation including *DEFECTIVE IN EXINE FORMATION 1* (*DEX1), NO EXINE FORMATION 1 (NEF1)* and *RUPTURED POLLEN GRAIN1* (*RPG1*) [Ariizumi *et al*., 2004; Guan *et al.,* 2008; Paxson-Sowders *et al*., 2001] (Figure 1.6). In addition *CALLOSE*

SYNTHETASE 5 (*CALS5*) regulates callose formation which is likely to be important for enhanced physical efficiency of primexine assembly, by keeping the primexine subunits around microspores (Dong *et al*., 2005). In combination to *CAL5*, *KAONASHI 2* (*KNS2*) is required for synthesis of primexine or callose wall, which are both important for probacula positioning (Suzuki *et al*., 2008). Other key genes mentioned in Figure 1.6 which are essential enzymes for sporopollenin synthesis include *ACYL-CoA SYNTHETASE 5* (*ACOS5*), *CYTOCHROME P450 703A2* (*CYP703A2*), *CYTOCHROME P450 704B2* (*CYP704B2, ARABIDOPSIS DEHYDROFLAVONOL 4-REDUCTASE-LIKE1* (*DRL1*), *LESS ADHESIVE POLLEN 5* (*LAP5*) *and LESS ADHESIVE POLLEN 6* (*LAP6*) all are among the well-conserved families in land plant (de Azevedo Souza *et al*., 2009; Dobritsa *et al*., 2009; Li *et al*., 2010; Tang *et al.*, 2009; Chen *et al*., 2011). Other genes included in Figure 1.6 are *β-KETOACYL REDUCTASE* (*KAR*), a key enzyme essential for the first reductive step in de novo fatty acid synthesis in rice (Aya *et al*., 2009), *TAPETUM DEGENERATION RETARDATION* (*TDR*) and GAMYB shown to be required for lipid biosynthesis and metabolism during early pollen wall development in rice and TDR is the ortholog of AMS (Li *et al*., 2006). In addition, rice *WAX-DEFICIENT ANTHER1 WDA1* is a close homolog of Arabidopsis *CER1*, which is putatively associated with anther cuticle and pollen exine development) which is putatively associated with anther cuticle and pollen exine development (Li and Zhang 2010; Jung *et al*., 2006) (Figure 1.6). Furthermore Chang *et al.,* (2012) have recently identified *NO PRIMEXINE AND PLASMA MEMBRANE UNDULATION* as essential for primexine deposition and plasma membrane undulation during microsporogenesis in Arabidopsis**.**
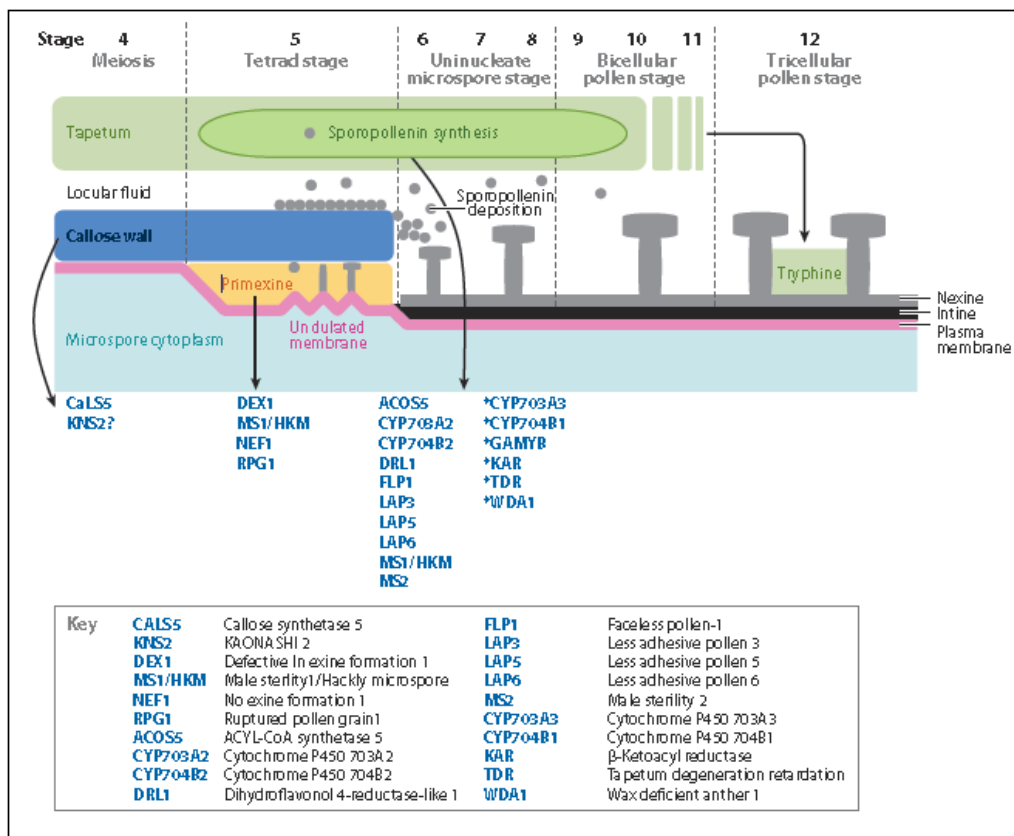
Figure 1.6 Current model of pollen wall formation in *Arabidopsis* (from Ariizumi and Toriyama, 2011). Asterisks indicate genes identified from rice.

1.5 ARABIDOPSIS MS1 REGULATORY NETWORK

ARABIDOPSIS MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis homologous to the Plant Homeo Domain (PHD)-finger family of transcription factors (Wilson *et al*., 2001). It has been reported to play an important role in synthesis of pollen wall materials and in polar secretion (Vizcay-Barrena and Wilson, 2006). Based on Vizcay-Barrena and Wilson, (2006) *MS1* is proposed to play a role in tapetal Programmed Cell Death (PCD) and pollen wall development; they demonstrated that the TUNEL (terminal

deoxynucleotidyl transferase–mediated dUTP nick-end labelling) staining did not occur in the mutant tapetum indicating that it failed to go through PCD, but changed to a necrotic breakdown pathway. Observations by Yang *et al*. (2007) indicated that MS1 is localized in the nucleus within the tapetum and is expressed in a developmentally regulated manner between late tetraspore and microspore release, then rapidly breaks down prossibly by ubiquitin-dependent proteolysis. Previous microarray data indicates that MS1 coordinates the expression of late genes associated with pollen wall formation, which are involved in the biosynthesis of components of the phenyl propanoid pathway, long-chain fatty acids, and phenolics, which are required for sporopollenin biosynthesis (Wijeratne *et al*., 2007; Yang *et al.,* 2007).

By transcriptome analysis, Yang *et al*. (2007) identified three putative protease genes as probable downstream targets of *MS1*. These three groups of putative MS1 downstream genes belonged to putative lipid biosynthetic functions; genes for fatty acid hydroxylases (CYP86-A1 and A8) and putative lipid transfer proteins (Figure 1.7). The putative metabolic function of these downstream genes suggests that MS1 directly or indirectly controls part of phenylpropanoid and fatty acid biosynthetic pathways leading to sporopollenin biosynthesis. The induction of these proteases may also be critical to the progression of PCD and that in the absence, possibly in association with a lack of secretion from the tapetal cells, PCD does not occur (Vizcay-Barrena and Wilson, 2006).

Transcriptomic data and microscopic observations suggest that MS1 also directly, or indirectly, controls the synthesis of the pollen coat. Possible targets include the tapetally expressed oleosin gene family, which mediate pollen coat behaviour by altering the size of lipid bodies, and the production of lipid transfer proteins. It is speculated that the altered expression of pollen wall material, either directly or indirectly, affects the secretion and deposition of such materials during the early stages of microspore maturation (Yang *et al.,* 2007) (Figure 1.7).

However, it is envisaged that MS1 does not directly regulate genes associated with pollen wall biosynthesis, due to the transient nature of MS1 expression and the subsequent phenotypic changes seen in pollen wall development, but acts via one or a number of additional transcription factors, including MYB99 and two NAM genes (At1g61110 and At5g61430) that contain a conserved NAC domain. This was further supported by Ito *et al*., (2007) using a dexamethasone-inducible MS1 construct to explore the possible role of MS1 leucine zipper and PHD-finger domains. Consequently, a model for the genetic control of pollen maturation and tapetum development was proposed (Figure 1.7) (Yang *et al*., 2007).

This model also suggests that since *MS1* is not expressed in the *dysfunctional tapetum1* (*dyt1*) mutant (Zhang *et al*., 2006), these genes are located downstream of *DYT1*, which is strongly expressed in tapetum at late stage 5 (Figure 1.7). In this regulatory pathway the *AMS* gene appears to function at similar stage as *MS1*, judging from the *ams* mutant phenotype (Sorensen *et al*., 2003).
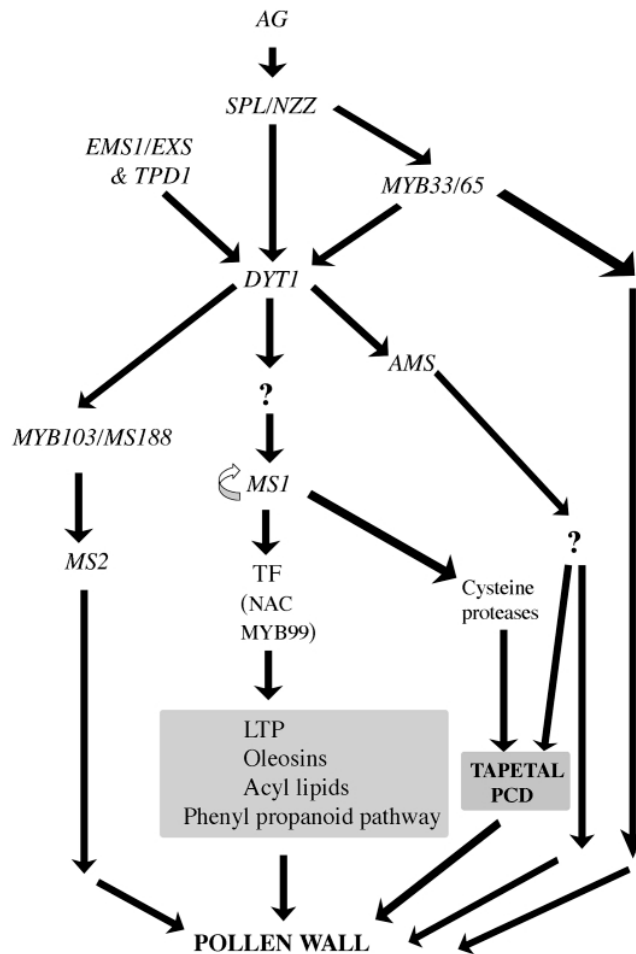
Figure 1.7 Model for MS1 regulatory network during late anther development by (from Yang *et al* 2007).

Recently Xu *et al.,* 2010 showed that *AMS* directly regulates other exine related genes. These findings suggest that *MS1* and *AMS* genes encode transcription factors that control different developmental/cellular processes but act at a similar stage. Furthermore, two transcription factor genes encoding a NAM family member (At1g61110) and MYB99 (At5g62320) were found to be immediately downstream of *MS1* (Figure 1.7). Analysis using the protein synthesis inhibitor cycloheximide combined with inducible MS1 expression, suggested that *MYB99* gene is a direct target of *MS1* (Ito *et al*., 2007).

Previous phylogenetic studies suggest that *MS1* is conserved across plant species (Ito *et al*., 2007) and this has been confirmed by recent work that identified the *MS1* orthologue (*PERSISTANT TAPETAL CELL)* in rice (Li *et al*., 2010). While attempts to integrate microarray data, bioinformatics analysis of *cis*-regulatory sequences and results from previous studies has resulted in a set of hypotheses about genetic interactions in a complex regulatory network, further experiments to expand our understanding of anther development are needed. Understanding how transcription factors (TFs) interact with the basal transcription apparatus at target gene promoters to activate or repress the target gene function also remains vital as PHD-finger transcription factors are essential for the regulation of gene expression (Bienz, 2006). Their response to development and intercellular signals seem to be found universally in the nucleus, and their functions tend to lie in the control of chromatin or transcription (Bienz, 2006

1.6 ARABIDOPSIS AMS REGULATORY NETWORK

The *AMS* gene encodes a MYC class basic helix-loop-helix transcription factor which plays a crucial role in tapetal cell development and post-meiotic transcriptional regulation of microspore development within the developing anther (Sorensen *et al*., 2003). The *AMS* gene is specifically expressed in the tapetum, with high expression during pollen mitosis I and the bicellular microspore stages (Sorensen et al., 2003). Analyses conducted by Xu *et al*. (2010) showed that mutation of *AMS* results in a down regulation of gene expression commencing from meiosis and that many genes with altered expression in the *ams* mutant are associated with metabolic changes occurring in the tapetum and in the

biosynthesis of pollen wall materials (Table 1.1). It is thought that many of these expression changes may be indirect effects of abnormal tapetal development, however qChIP-PCR analysis indicated that AMS selectively binds the promoters of 13 of these genes (Xu *et al*, 2010) (Table 1.2), suggesting direct regulation of genes associated with tapetal metabolic processes.

Table 1.1 Functional classification by eukaryotic orthologous group analysis of genes with altered expression in the *ams* mutant as identified by sam analysis (<5% false discovery rate) and showing a >Log2-Fold change in expression (from Xu *et al*., 2010).

Function	Down	Mei	PMI	Bi	PMII	UP	Mei	PMI	Bi	PMII	Total
Information storage and processing											
Translation, ribosomal structure, and biogenesis	17	3	6	5	3	7	4	1	0	2	24
RNA processing and modification	11	2	4	5	0	1	0	1	0	0	12
Transcription	29	2	8	10	9	4	2	1	0	1	33
Chromatin structure and dynamics	3	1	1	1	0	1	0	1	0	0	4
Cellular processes and signaling											
Cell cycle control, cell division, chromosome partitioning	1	0	0	1	0	0	0	0	0	0	1
Defense mechanisms	31	7	8	3	13	14	1	3	10	0	45
Signal transduction mechanisms	41	8	17	6	10	32	7	4	19	2	73
Cell wall/membrane/envelope biogenesis	35	6	14	8	7	13	1	0	5	7	48
Cytoskeleton	10	2	4	2	2	4	0	1	2	1	14
Extracellular structures	1	0	0	0	1	3	1	1	1	0	4
Intracellular trafficking, secretion, and vesicular transport	1	0	0	0	1	2	0	1	1	0	3
Posttranslational modification, protein turnover, chaperones	41	18	7	9	7	21	5	8	5	3	62
Metabolism											
Energy production and conversion	11	0	4	7	0	2	1	1	0	0	13
Carbohydrate transport and metabolism	58	11	14	19	14	30	8	11	6	5	88
Amino acid transport and metabolism	38	8	12	15	3	5	1	3	1	0	43
Nucleotide transport and metabolism	8	2	2	3	1	0	0	0	0	0	8
Coenzyme transport and metabolism	7	2	2	2	1	1	0	0	0	1	8
Lipid transport and metabolism	82	14	22	25	21	14	7	4	2	1	96
Inorganic ion transport and metabolism	25	10	8	3	4	11	4	4	3	0	36
Secondary metabolites biosynthesis, transport and catabolism	97	28	28	29	12	20	6	7	5	2	117
Poorly characterized											
General function prediction only	119	34	37	27	21	113	18	29	33	33	232
Function unknown	32	7	12	8	5	73	14	16	28	15	105

Table 1.2 List of representative ams binding targets as shown by ChIP-qPCR that were differentially expressed in the ams microarray analysis (from Xu *et al*., 2010).

Genes Positively Regulated by AMS		Mei	PMI	Bi	PMII	Cluster Group
At1g59740	H⁺/oligopeptides symporter	−1.19	−2.18	−2.27	−0.86	4
At3g51590	Lipid transfer protein type 1	−3.68	−2.98	−4.24	0.00	4
At1g66850	Lipid transfer protein type 2	−4.08	−4.92	−4.10	−1.09	4
At1g75920	Predicted lipase	−3.12	−3.58	−4.06	0.00	4
At5g49070	Fatty acid elongase/3-ketoacyl-CoA synthase	−1.01	−0.66	0.00	0.00	5
At1g75790	Multicopperoxidase	−3.09	−3.36	−2.79	0.00	4
At1g67990	O-methyltransferase	−4.22	−3.90	−3.86	0.00	4
At4g00040	Chalcone synthase	−0.45	−1.30	0.00	0.00	5
At3g28740	Cytochrome P450 CYP2 subfamily	0.00	1.14	1.34	0.18	1
At1g13140	Cytochrome P450 CYP4/CYP19/CYP26 subfamilies	−1.77	−2.78	−1.57	−0.07	4
At5g17050	UDP-glucuronosyl and UDP-glucosyl transferase	−1.50	−1.08	−1.01	0.00	7
At3g13220	Transporter, ABC superfamily (Breast cancer resistance protein)	−2.57	0.00	−1.95	0.00	7
At1g73220	Synaptic vesicle transporter SVOP and related transporters (major facilitator superfamily)	−1.74	−3.09	−2.14	−1.60	4

Recently (Ma *et al*, 2012) analysed the transcriptome profiles of early anthers from sterile mutants of *ams* and found that 1,368 genes were differentially expressed in *ams* compared to wild type anthers; these changes involved genes affecting metabolism, transportation, ubiquitination and stress responses. Moreover, the lack of significant enrichment of potential AMS binding sites (E-box) in the promoters of these differentially expressed genes suggests both direct and indirect regulation. Combining *ams* transcriptome profiles with those of *spl/nzz* and *ems1/exs* anthers uncovered overlapping and distinct sets of regulated genes, including transcription factors and other proteins The investigation of expression patterns of major transcription factor families, such as bHLH, MYB and MADS, suggested that some closely related homologs of known anther developmental genes might also have similar functions (Ma *et al*, 2012) (Figure 1.8).

Despite the increasing knowledge of genes that contribute to anther development, the regulatory mechanisms controlling this process are still unclear. Careful analyses of transcriptome data, combined with genetic and phylogenetic information has revealed an elaborate regulatory network during early anther development that has expanded the understanding of the hierarchy of anther-development related genes, especially transcription factors (Ma *et al*, 2012) (Figure 1.8).
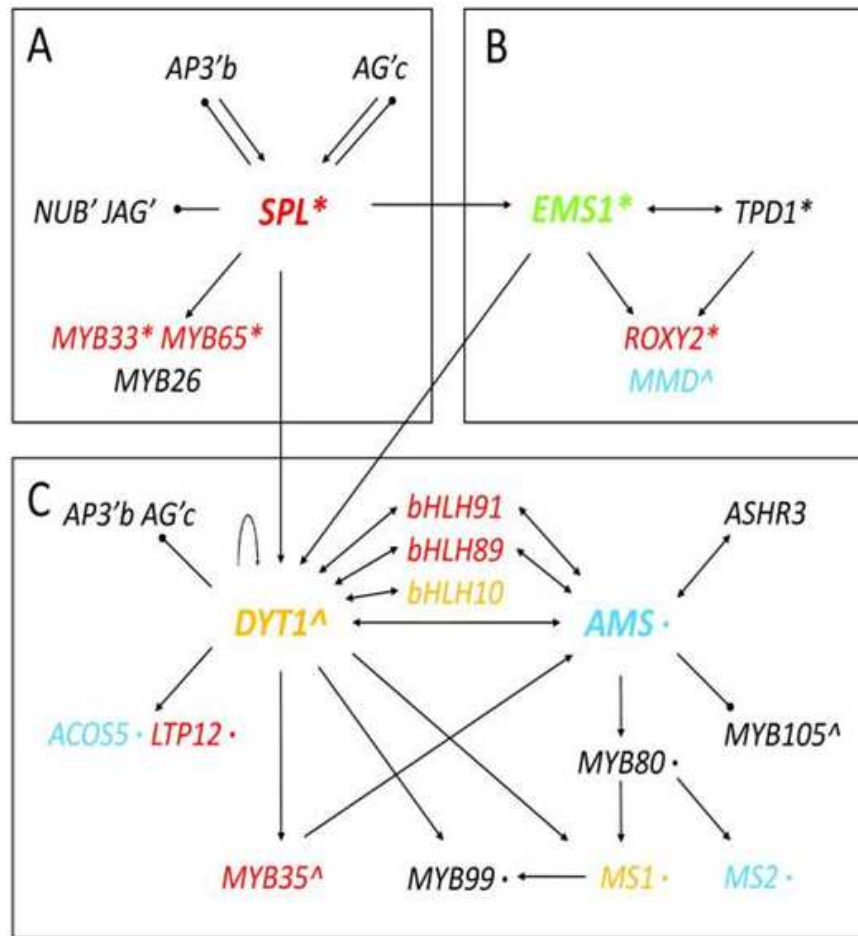
Figure 1.8 Gene regulatory network of anther development during early stages (from Ma *et al*, 2012). Gene regulation is represented by T-bars (negatively) and arrows (positively). The direct regulation confirmed by experiment is represented in bold line. Genes encoding proteins with interaction is represented by double arrows. Gene expression patterns in different tissues are shown by colours (blue for anther specific; red for anther preferential; green for reproductive-preferential and yellow for genes not included in ATH1 chip). Gene function in tapetum formation is marked by an apostrophe; in pollen wall formation by an asterisk; in callose dissolution by double asterisks; in stamen and petal formation by the letter b; in stamen and carpel formation by the letter C.

These results provide insight into the regulatory role of the AMS network during anther development. However, little information is available regarding possible

21

relationships between these genes and genetically defined regulators. At the moment more needs to be done to identify genes that may be interacting with MS1 in facilitating regulation of late tapetal function during pollen development.

1.7 AIMS AND OBJECTIVES

The overall objective of this research work was to conduct a molecular analysis of genes that are expressed in the anther tapetum and are critical for functional pollen development. A number of these are thought to be regulated by, or interact with MALESTERILITY1, or ABORTED MICROSPORE. The following detail the specific aims of the work in order to achieve these goals:

i) To identify regulatory targets of the tapetally expressed transcription factor ABORTED MICROSPORE (AMS)

ii) To identify mutants and genes that are involved in the regulation of the MALE STERILITY1 transcription factor.

iii) To investigate how MS1 is functioning to regulate gene expression in the tapetum by investigate the proteins that are interacting with MS1.

1.7.1 IDENTIFICATION OF AMS REGULATORY TARGETS

The regulatory role of AMS in the tapetum during anther and pollen development has not been fully defined, therefore genes were identified from DNA microarray analysis of the *ams* mutant that are down-regulated, or show altered expression, in

the *ams* mutant. These genes were selected for further functional analysis to identify those that are direct targets of the AMS transcription factor that are expressed in the tapetum during microspore development. SAIL/SALK insertional mutants in these putative targets were used for bioinformatics, genotypic, phenotypic and expression analysis. From this analysis a SALK mutant line in the ABC transporter (At3g13220) has been shown to be critical for viable pollen formation and confirmed as directly regulated by AMS (Xu *et al*., 2010) (Chapter 3).

1.7.2   IDENTIFYING FACTORS REGULATING THE EXPRESSION OF MS1

Fast-neutron mutagenesis was utilised to identify additional factors that may be regulating the expression of the tapetal transcription factor MS1. This was conducted by screening for novel male sterile mutants from a population of mutagenised seed carrying the MS1Prom:MS1-GFP functional fusion protein. Microscopic observation of the fluorescent reporter gene GFP showed changes in the stage specific expression of MS1-GFP in these mutants. Backcrossing of these male sterile mutants with the parental plants (carrying the MS1Prom:MS1-GFP fusion construct) and the *ms1* mutant was conducted to confirm the mutations and check for allelism to the *ms1* mutation. Gene mapping of one mutant has been conducted. Delayed and prolonged expression of MS1-GFP in the mutant may suggest that the mutant gene function as a negative regulator of MS1 activity during pollen development. These results will in the future provide insight into the regulation of the MS1 network (Chapter 4).

1.7.3 IDENTIFICATION OF THE PROTEINS THAT ARE INTERACTING WITH MS1

A stamen Yeast-2-Hybrid library was previously screened using the MS1 protein as bait to identify proteins that might be interacting with MS1 *in vivo*. A putative interacting protein, POB2, was identified, which appeared to be involved in ubiquitin-based proteolytic breakdown. This interaction was been subsequently confirmed by further Yeast-2-Hybrid analysis and bifluorescent gene complementation (BIFC) analysis. Additional attempts were also performed to verify this interaction *in vitro* and *in planta* by pull-downs of transiently expressed proteins in *E.coli* and *Nicotiana benthamiana* respectively (Chapter 5). The MS1 protein has been previously shown to have transient expression in the tapetum, therefore the putative interaction with POB2 may indicate that the mechanism of MS1 degradation in the tapetum is via a ubiquitin-proteasome system involving the POB2 protein.

CHAPTER 2

GENERAL MATERIALS AND METHODS

2. GENERAL MATERIALS AND METHODS

2.1 MAINTENANCE AND CULTIVATION OF *ARABIDOPSIS THALIANA* SEED STOCK

Arabidopsis thaliana ecotype Landsberg *erecta* (L*er*) and Columbia (Col) was used as wild type and was obtained from Nottingham Arabidopsis Stock Center NASC. Growth conditions consisted of 22 hours photoperiod under 111µmol s^{-1} m-^{1} warm white fluorescent light at 75% rh and 22+/- 2^{0}C. Plants were grown in 9cm to 12cm plastic pots filled with Levington M3 compost (Scotts Company UK). Intercept 5GR (21g in 75 L; Scotts Company UK) was added and mixed thoroughly before sowing. Seeds were sown at a density as required by the experiment. Cross pollination between plants was avoided by using transparent plastic sleeves (Zwapak Netherlands), which isolated each pot from its neighbours. Plants were watered until at least 90% of the siliques had dried completely and then the plants were allowed to dry slowly on dry benches for maximum viable seed production. Seeds were harvested after the whole plants had dried out and stored in moisture-porous paper bags in a dry atmosphere at room temperature.

2.2 MAINTENANCE AND CULTIVATION OF *NICOTIANA BENTHAMIANA PLANT MATERIAL*

N. benthamiana seeds (from University of Nottingham) were grown at 22ºC+/- 2^{0}C in pots containing compost immersed in 0.15g.l-1 of Intercept® solution

(Scotts, U.K.). Plants used for transient assays were grown for 5-6 weeks at which point the plants were infiltrated with appropriate gene constructs for protein expression (Kapila *et al*., 1997). The leaves were harvested three days after infiltration for protein analysis (Section 5.2.2 and 5.3.2).

2.3 IDENTIFICATION AND ANALYSES OF ARABIDOPSIS SALK AND SAIL MUTANT LINES

Using DNA microarray analysis, putative regulatory target genes of AMS and MS1 were identified in SAIL/SALK mutants or KO lines (*http://arabidopsis.info/BrowsePage*). Seeds of these Knock-Out (KO) lines were ordered from Nottingham Arabidopsis Stock Centre (NASC) (University of Nottingham) in Arabidopsis ecotype Columbia (Col) background (Chapter 3 Table 3.1). The KO lines were genotyped and phenotyped to identify homozygous sterile lines. The T-DNA in these lines not only disrupted the expression of the gene into which it was inserted but also acted as a marker for subsequent identification of the mutation (Chapter 3).

2.4 PLANT, BACTERIAL AND YEAST GROWTH CONDITIONS, MEDIA AND ANTIBIOTICS

Screening of *Arabidopsis* transgenic lines and onion epidermal strips for transient assays was conducted on 12cm Petri dishes containing 0.5X Murashige and Skoog Basal Medium (MS) (Table 2.1) containing antibiotics as appropriate. For yeast-2-hybrid analyses protease peptone salt amino acids were supplied by Oxoid Ltd

(Hampshire, U.K.). YPD medium, YPD agar medium, Minimal SD Base, Minimal SD agar base and DO supplements were sourced from BD Biosciences Clontech. Yeast extract powder and agar supplied by Formedium Ltd. Tryptone was provided by Melford (Watford U.K.). Luria broth medium was supplied by Invitrogen. Various bacterial and yeast strains were added to the growth medium with appropriate antibiotics for selection of positive colonies carrying the construct with the gene of interest (Table 2.2 and 2.3).

Table 2.1 Plant growth medium utilised for screening *Arabidopsis* transgenic lines and onion epidermal strips for transient assays.

MATERIAL	PURPOSE	GROWTH MEDIUM
A. thaliana	Mutant screen	0.5 X Murashige and Skoog Basal Medium (MS)
A. thaliana	Mutant screen	0.5 X Murashige and Skoog Basal Medium (MS)
Onion epidermal cells	BIFC complementation	0.5% X Murashige and Skoog Basal Medium (MS)

Table 2.2 Bacterial and Yeast growth media.

STRAIN	SPECIES	PURPOSE	GROWTH MEDIUM
DH5a	*Escherichia coli*	Cloning/Amplification	Luria Broth (LB) liquid/agar
BL21 (DE3)	*Escherichia coli*	Cloning/Amplification	Luria Broth (LB) liquid/agar
BL21 AI	*Escherichia coli*	Cloning/Amplification	Laural Broth (LB) liquid/agar
C58	*Agrobacterium tumefaciens*	Transient assays in tobacco	Laural Broth (LB) liquid/agar
C58	*Agrobacterium tumefaciens*	Floral dip transformatiion	Laural Broth (LB) liquid/agar
GV3101	*Agrobacterium tumefaciens*	Transient assays in tobacco	Laural Broth (LB) liquid/agar
MAV203	Yeast	Yeast-2-Hybrid analyses	YPD liquid/agar

Table 2.3 Antibiotics used for selection of constructs/lines as appropriate.

ANTIBIOTIC	SOLVENT	STOCK CONC.	WORKING CONC.	MANUFAC-TURER
Ampicillin (Amp)	H_2O	100 mg.ml-1	100 µg.ml-1	Melford
Carbenicillin (Carb)	H_2O	100 mg.ml-1	100 µg.ml-1	Melford
Chloramphenicol (Cm)	Ethanol	50 mg.ml-1	25 µg.ml-1	Duchefa Biochimie
Gentamycin (Gent)	H_2O	25 mg.ml-	1 25 µg.ml-1	Duchefa Biochimie
Rifampicin (Rif)	Methanol	10 mg.ml-1	100 µg.ml-1	Sigma-Aldrich
Spectinomycin (Spect)	H_2O	50 mg.ml-1	50 µg.ml-1	Sigma
Tetracyclin (Tet)	Ethanol	5 mg.ml-1	5 µg.ml-1	Sigma-Aldrich
Kanamycin (Kan)	H_2O	100 mg.ml-1	100 µg.ml-1	Melford
Hygromycin (Hyg)	H_2O	50 mg.ml-1	50 µg.ml-1	Sigma-Aldrich

2.5 PHENOTYPIC ANALYSIS

Transgenic plants were compared to the wild type and differences in terms of general plant growth, including a detailed study of reproductive organs and silique formation, were recorded. Plants that had short siliques and did not contain seeds,

or had reduced numbers of seeds were analysed for pollen viability by Alexander stain (Alexander, 1969) to establish whether viable pollen was present.

2.6 ALEXANDER STAIN AND MICROSCOPIC ANALYSIS

Alexander's stain (Alexander 1980) contains malachite green, which stains cellulose in pollen walls, and acid fuchsin, which stains the pollen protoplasm; this is able to distinguish between viable (red-purple) and non-viable (green-blue) pollen grains. Alexander staining was conducted on isolated anthers from closed buds and open flowers. Anthers were placed on glass slides to which a drop of Alexander's stain was added for 5min. Cover slips were placed onto the anthers and these were viewed under a (Zeiss) microscope to detect pollen viability. The images were captured by Nixon 1 camera connected to the (Zeiss) microscope.

2.7 GFP AND YFP MICROSCOPIC ANALYSIS

Plants were analysed for GFP expression using Zeiss and Nixon microscopes. Forceps and sharp surgical needles were used to collect anthers from the smallest to the largest closed buds under a dissecting microscope (Zeiss) at x20 magnification. Excised anthers were placed onto glass slides to which a drop of 10% (v/v) glycerol was added. Cover slips were placed onto the glass slides and the material was viewed under a microscope at magnification (Zeiss) to detect GFP expression under UV light (488-nm excitation filter and 560-nm excitation filter for chlorophyll autofluorescence). The images were captured by a Nixon 2 camera connected to the microscope (Zeiss). Apart from the Zeiss microscope the

Leica TCS SP2 AOBS confocal scanning microscope was also used to image the plant materials expressing YFP fluorescence proteins using a 100 mW multi-line Argon laser (458 nm, 476 nm, 488 nm, 496 nm and 514 nm) and a 1 mW He-Ne laser (543 nm) as excitation sources. The plant tissues were mounted on to the glass slide with a drop of 1 x PBS buffer (Sigma) and covered gently with the No. 1.5 cover slip (0.17 mm thick, Scientific Laboratory Supply). The slide was observed under UV light. The appropriate laser intensity was selected in the Leica Control Software (LCS) accordingly and the live image was acquired instantly through continuous scan mode. The scanner and detector was set to xyz scanning mode, 514x514 image size, 8 bit image and 400 Hz scan rate unless otherwise mentioned. The pinhole was set to AE unit 1 as default. The voltage applied to the photomultiplier tube (PMT, AKA "detector gain value") was adjusted experimentally to obtain the best signal-to-noise ratio. The z-position and electric-zoom were chosen accordingly and the series z-position scan and image maximal projection was carried out by using LCS. Images were acquired using the 10x/0.4 HC PL APO CS object lens and the 40x/0.7 HCX PL FLUOTAR object lens. The 63x/1.3 HCX PL APO CS and 40x1.25-0.75 HCX PL APO CS object lenses were used to obtained images where fluorescent proteins were targeted to the nucleus.

2.8 MOLECULAR BIOLOGY METHODS

2.8.1 GENOMIC DNA EXTRACTION

Genomic DNA was isolated from *Arabidopsis* leaf and floral tissue, 300µl of DNA extraction buffer (200mM Tris-HCl pH8.0, 250mM NaCl, 25mM EDTA, 0.5% SDS) was added to 1g of leaf material in a microcentrifuge tube and ground using a micropestle (www.fisher.co.uk). The homogenate was spun at ~17 900xg for 5min, the supernatant was then transferred to a fresh microfuge tube. One volume of isopropanol was added to precipitate the DNA and incubated at -20ºC for 15-20min before repeating the centrifugation. The supernatant was discarded and the pellet was washed with ice-cold 75% (v/v) ethanol and air-dried; 20µl of TE (10mM Tris-HCl, 1mM EDTA, pH8.0) was added to the pellet and this was left at room temperature for 10-15min, tapping the tube occasionally. High quality genomic DNA for cloning and sequencing purposes (section 2.9.9 and 2.9.11) was extracted using the DNeasy® Plant Mini Kit (Qiagen) according to the manufacturer's instructions. DNA quality and quantity were determined using a NanoDrop N-1000 fluorospectrometer (NanoDrop technologies USA). Genomic DNA samples were stored at -20ºC.

2.8.2 HIGH THROUGHPUT DNA ISOLATION

Using a multichannel pipette 750µl DNA extraction buffer (1M Tris-HCl (pH 7.5), 4M NaCl, 0.5M EDTA, 10% (w/v) SDS) was dispensed into each DNA collection microtube rack (96 well, Qiagen, Catalogue Number 19560). A small

leaf sample (~40 mg) from an individual plant was picked using forceps, cleaned in ethanol (and shaken to remove excess ethanol) and placed in the collection tube with DNA extraction buffer and the steel ball bearing (Diameter=0.3cm; www.Simplybearings.co.uk). The collection tubes were then sealed with the 96 well DNA collection microtube caps (Qiagen, Catalogue Number 19566) ensuring all wells did not leak (the plate was tipped upside down to check this). The samples were homogenised by placing them in a Mixer Mill 8000 (GLEN Creston LTD) for up to 5 minutes. The collection tube box (containing buffer, sample and ball bearings) was centrifuged at 5010 x g for 10 minutes. 400ul of the supernatant was transferred to a 1.1mL 96 deep well plate (Starlab, Catalogue Number E2896-0100) containing an equal volume of chilled isopropanol, without disturbing the pellet, and incubated at room temperature for 5-10 minutes to precipitate the DNA. The deep well plate was sealed with the 96 well storage microplate mat (Starlab, Catalogue Number E2896-1100) and centrifuged at 5010 x g for 15 minutes. The supernatant was discarded and the pellet washed with 150µl of 70% (v/v) ethanol. The ethanol was then discarded and the plate was air dried overnight at room temperature. The DNA was resuspended in 100ul of sterile water. The plate was sealed and stored at -20°C.

2.8.3   EXTRACT-N-AMP^{TM} PLANT PCR KIT

Using the Extract-N-Amp^{TM} Plant PCR Kit (Qiagen) DNA was extracted from 15 day-old *Arabidopsis* leaves of T-DNA insertion lines (Section 3.2) and control plants (wild type Columbia ecotype). Using forceps to handle leaf material, which had previously been rinsed in 70% (v/v) ethanol, leaf tissue disks measuring about

0.5-0.7cm diameter were collected for DNA isolation. Leaf disks were placed into a 2ml collection tube and incubated in 50ml of extraction solution for 10 min at 95°C. Afterwards an equal volume of dilution buffer was added to the extract in readiness for PCR. Aliquots of the diluted extracts were genotyped by PCR using the Extract-N-Amp PCR Reaction Mix, the LBa1 primer (Chapter 3 Table 3.3) and gene specific primers (Chapter 3 Table 3.4).

2.8.4 AGAROSE GEL ELECTROPHORESIS AND VISUALIZATION OF DNA

Confirmations of PCR products or restriction digests were checked by agarose gel electrophoresis. DNA samples were mixed with 5x loading buffer [0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 30% (v/v) glycerol] and loaded into agarose gels consisting of 0.8 -1% (w/v) agarose (Bioline USA) in 1xTAE (40mM Trisacetate, 1mM EDTA) and visualised using 0.5ng/ml ethidium bromide. 5ul of Hyperladder 1 (Bioline USA) was loaded next to the samples as a DNA size marker (unless otherwise stated). Samples were separated by electrophoresis in 1xTAE at 100V and visualized under UV light.

2.8.5 ISOLATION OF DNA FRAGMENTS FROM AGAROSE GELS

DNA fragments of the expected sizes were excised from the gel with a clean razor blade under a UV-illuminator. The extracted gel band was weighed and the DNA was purified using the QIAquick® Gel Extraction Kit (Qiagen) in accordance

with the manufacturer's instructions. Briefly, the gel was dissolved in 3 volumes (100μg of gel = 100μl) of buffer QG at 50^0C for ~10min. Then 1volume of isopropanol was added and the solution was applied to the supplied columns and spun at maximum speed (~17 900xg) for 1min to absorb the DNA onto the silica membranes. Impurities were then washed off the membrane using the ethanol-containing PE Buffer (supplied in the kit) by centrifugation as before. An additional centrifugation step was carried out to remove extra PE Buffer. Finally 30μl of RNAse free water was added to the membrane and incubated at room temperature for 1min before spinning at (~17 900xg) for 1min. The eluted DNA was stored at -20ºC.

2.8.6 POLYMERASE CHAIN REACTION PCR ANALYSIS

Typical PCR reactions comprised of 20μl total reaction mixture as shown in Table 2.4. The PCR reaction parameters are shown in Table 2.5.

Table 2.4 Typical PCR reagent mixture

REAGENTS	VOLUME (μl)	MANUFACTURER
Distilled Water	14.0	Sigma
PCR buffer	2.0	New England Biolabs USA
$MgCl_2$ (25mM)	1.0	Invitrogen, USA
dNTPs (10 mM each)	0.5	Invitrogen, USA
Forward primer 10 mM	0.5	Eurofins
Reverse primer 10 mM	0.5	Eurofins
Taq polymerase 5U/uL	0.5	University of Nottingham
DNA extract 10mM	2.0	
Total volume	20.0	

Table 2.5 PCR cycling parameters for genotypic analysis

STEP	TEMPERATURE	TIME	CYCLES
Initial Denaturation	94^0C	2 min	1
Denaturation	94^0C	30 sec	
Annealing*	X^0C	30 sec	22-35
Extension	72^0C	1 min 30 sec	
Final extension	72^0C	4 min	1

*An appropriate annealing temperature was chosen depending on the primers used and was typically 5°C below the melting temperature.

2.8.7   cDNA SYNTHESIS AND RT-PCR

cDNA was prepared using 5μg of total RNA in a 20 μl reaction containing SuperScript™ II Reverse Transcriptase (Invitrogen) (Table 2.6). This Reverse Transcriptase is an engineered version of Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) with reduced RNase H activity and increased thermal stability. The enzyme is purified to near homogeneity from *E. coli* containing the modified *pol* gene of M-MLV RT used to synthesize first-strand cDNA at higher temperatures than conventional M-MLV RT, providing increased specificity, higher yields of cDNA, and more full-length product of cDNA up to 12.3 kb. First-Strand cDNA Synthesis using SuperScript™ II RT used a 20-μl reaction volume for 5 μl of total RNA. The components added to the nuclease-free microcentrifuge tube for cDNA synthesis are given in Table 2.6. To remove RNA complementary to the cDNA, 1 μl (2 units) of *E. coli* RNase H was then added to the cDNA which was incubated at 37°C for 20 min. The cDNA extracted was used as a template for amplification in RT-PCR analyses. RT-PCR analyses were

conducted by using 10mM of cDNA template and the appropriate gene specific primer pairs for 28-35 cycles (see Table 2.7 and 2.8). RT-PCR results were normalized using *A. thaliana* actin-7 forward and reverse primer

Table 2.6 Components for First-Strand cDNA synthesis

REAGENT	VOLUME (μl)
Oligo(dT)12-18 (500 μg/ml)	1
Total RNA 5ug	5
dNTPs Mix (10 mM each)	1
5X First-Strand Buffer	4
DTT (0.1 M)	2
RNaseOUT™ (40 units/μl) (optional) [4.]	1
SuperScript™ II RT (200 units)	1
Sterile distilled water	5
Total volume	20

Table 2.7 RT-PCR reaction mixture

REAGENTS	VOLUME (μl)
Taq buffer (Invitrogen) (x10)	3.0
Gene specific Forward primer (10 mM)	0.6
Gene specific Reverse primer (10 mM)	0.6
dNTPs mix(2.5 mM each)	0.6
cDNA (10ng/uL) Template	1.0
Taq polymerase 5U/uL	0.2
Sterile water	24.0

Table 2.8 RT-PCR cycling parameters

STEP	TEMPERATURE	TIME	CYCLES
Initial Denaturation	94^0C	2 min	1
Denaturation	94^0C	30 sec	
Annealing**	X^0C	30 sec	
Extension	72^0C	1 min 30 sec	28-35
Final extension	72^0C	4 min	1

**An appropriate annealing temperature was chosen depending on the primers used (typically 5°C below the melting temperature

2.8.8 HEAT-SHOCK TRANSFORMATION OF BACTERIAL COMPETENT CELLS

As appropriate different strains of competent *E.coli* cells (Invitrogen; stored at - 80°C) were used for transforming the DNA constructs; 50µl of competent cells were placed on ice and 50 ng of circular DNA added, and incubated for 10 min. The cells were then heat shocked at 42^oC for 45 seconds and then returned to the ice for 2 minutes. Tubes had 250µl of S.O.C medium was added to the reaction mixture and were then incubated for 1 hour at 37^oC. 100µl of the resulting culture was spread onto LB plates with appropriate antibiotic added and left to grow overnight at $37^°C$ on LB medium. Positive transformant were selected by colony PCR (2.9.10).

2.8.9   GATEWAY™ CLONING PROCEDURE

The Gateway™ Cloning System (http://www.invitrogen.com) is increasingly important in the field of biology, particularly genetics and molecular biology. The cloning itself, like other methods, allows for functional analysis of genes, protein expression, and cloning or subcloning of DNA fragments using various plasmid vectors (Table 2.9). The PCR insert with *att*B sites was used as the typical substrate for BP recombination cloning into the pDONR™ vectors to create an Entry clone. An entry clone is a plasmid containing the gene of interest and flanked by Gateway *attL* recombination sites (Figure 2.1). To carry out BP reaction the components in Table 2.10 were added to a 1.5 ml tube at room temperature, mixed and incubated at 25°C for 1 hour. Then 1 µl of Proteinase K (2µg/µl) was added to each sample to terminate the reaction.

The reaction was vortexed briefly and incubated at 37°C for 10 minutes. Heat-shock transformation of 50 µl of the appropriate competent *E. coli* cells with 1 µl of the reaction was carried out. 250µl of SOC or LB medium was added to the reaction, which was spread onto LB plates containing the appropriate antibiotic for selection. After overnight incubation an efficient BP recombination reaction typically produced >1500 colonies. Once positive clones were identified by colony or plasmid DNA PCR (Section 2.8.10) the plasmid DNA was purified (Section 2.9.11) and used for creating the expression clone by Gateway LR reaction (Figure 2.1).
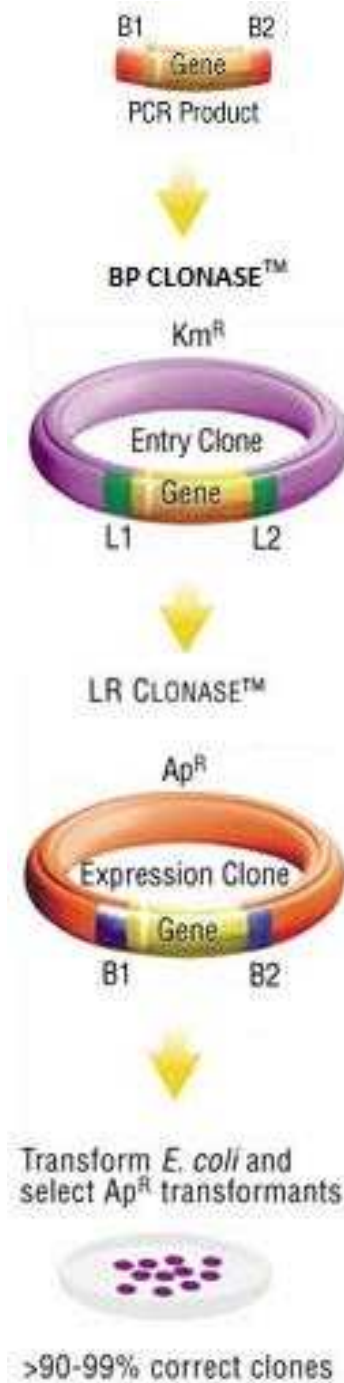
Figure 2.1 Overview of the Gateway method (http://www.invitrogen). PCR product was inserted into the entry clone (purple cycle) by Gateway BP reaction. The entry clone was then used to express the gene in a destination vector facilitated by the *att* recombination sites used with enzyme clonase mixes via the Gateway LR reaction (for more information refer to Gateway manual on http://www.invitrogen).

41

Table 2.9 Plasmid vectors used for Gateway cloning system.

PLASMID OF VECTOR	DESCRIPTION	FEATURES	SOURCE
pDONR™221	Donor vector	Recombinational donor/master vector with M13 F and R primer sites; kanamycin resistance; recombinational cloning	Invitrogen
pDONR™201	Donor vector	Recombinational donor/master vector; kanamycin resistance; recombinational cloning	Invitrogen
pENTRTM-D/TOPO	Donor vector	Directional cloning vector, hT7 promoter, with M13 F and R primer sites; kanamycin resistance.	Invitrogen
pDEST15 [BE]	Destination vector	T7 Promoter, Glutathione-S-transferase Tag (GST) on (C-terminal)	Invitrogen
pDEST17 [BE]	Destination vector	T7 promoter, 6xHis Tag (N-terminal)	Invitrogen
pDEST 24 [BE]	Destination vector	T7 promoter, GST Tag (C-terminal)	Invitrogen
pDEST 22 [Y]	Destination vector	AHD1 promoter, GAL4 DNA activation domain (C-terminal)	Invitrogen
pDEST 32 [Y]	Destination vector	AHD1 promoter, GAL4 DNA binding domain (N-terminal)	Invitrogen
pGWB5 [BE]	Destination vector	35S promoter, GFP Tag (C-terminal)	Invitrogen
pGWB8 [BE]	Destination vector	35S promoter, 6xHis Tag (C-terminal)	Invitrogen
pGWB17 [BE]	Destination vector	35S promoter, 4xMyc Tag (C-terminal)	Invitrogen
pG005 [BIFC]	Destination vector	BIFC gateway binary vector (Hygromycin), 35S promoter, GW, YFP N-terminus	donated by Don Grierson, Silin Zhong (UoN)
pG006 [BIFC]	Destination vector	BIFC gateway binary vector (Hygromycin), 35S promoter, GW, YFP C-terminus	donated by Don Grierson, Silin Zhong (UoN)

[Y], (Yeast-2-Hybrid analysis) [BE], (Bacterial expression) ***, (BIFC analysis) [BIFC]

Table 2.10 Gateway BP reaction mixture

MATERIAL	VOLUME (µl)
PCR product (=10 ng/µl; final amount ~15-150 ng)	1-7
Donor vector (150 ng/µl)	1
10 X TE-buffer pH 8.0	1
BP Clonase™ II enzyme mix (100 ng/µl)	1
Total volume	10

The LR reaction was carried out to transfer the sequence of interest to one or more destination vectors (Table 2.11) in simultaneous reactions, making the technology high throughput. The LR Reaction involves a recombination reaction between attL and *attR* sites (Figure 2.2). As the Destination vector is designed to include the *attR* sites, the Entry clones have *attL*'s on both sides of their gene of interest. The sequences on *attL* sites then match up with the sequences on the destination vector, which contains attR recombination sites. Gateway manual on http://www.invitrogen). Table 2.11 shows the reaction mix required to carry out the LR reaction. 1 µl of entry vector plasmid DNA was transferred to a 1.5 ml microfuge tube containing 1.5 µl TE (pH 8.0), 0.5 µl of plasmid DNA (10 ng/µl) of the appropriate destination vector, 1µl of 5x LR Clonase plus buffer and LR Clonase enzyme (directly from the freezer) and mixed by pipetting. The mixture was incubated at 25 ºC overnight. The reaction was halted by adding 0.5 µl proteinase K (2µg/µl) and incubated at 37ºC for 10 min. To perform heat shock transformation 50 µl of library efficiency® DH5α™ competent cells of *E. coli*

strain was added and the reaction mixture was left on ice for 30min. The mixture was taken to 42°C for 1 minute and back on ice for 3min; 250 μl of LB medium was added to the reaction mixture and incubated for 1 hour. The bacteria culture was plated on LB medium containing 100μg/ml of ampicillin as a selective antibiotic. The plates were incubated at 37ºC for overnight and screening for clones were done by PCR using Destination vector adapter and gene specific primers.
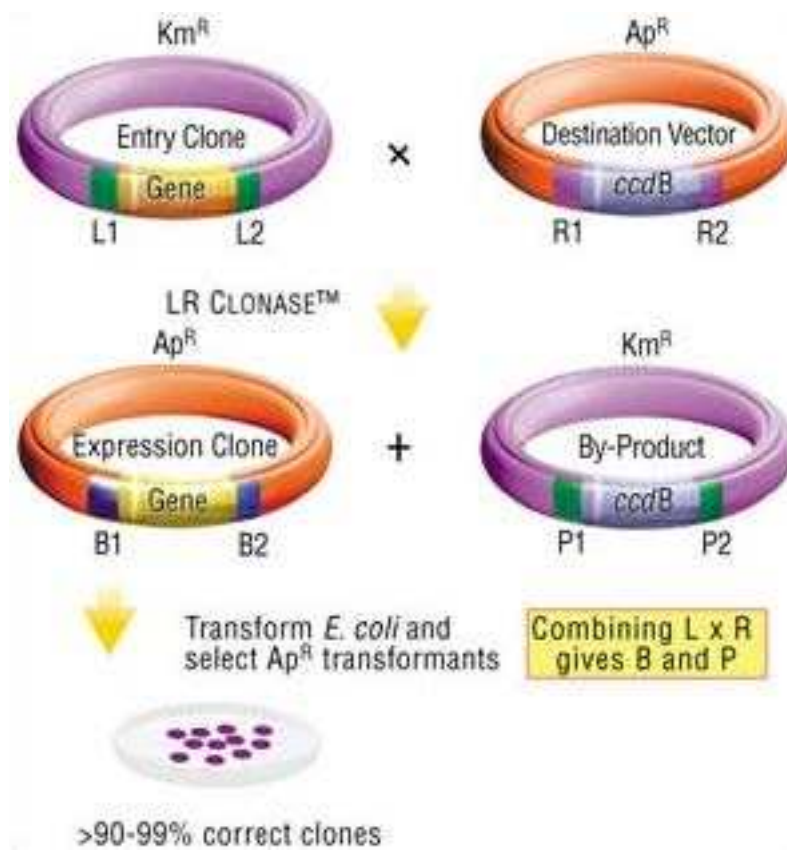
Figure 2.2 LR recombination reactions (http://www.invitrogen). Entry clone with *attL* recombines with a destination vector with *attR* to form a new expression clone with *attB* and a by-product with *attP* (for more information refer to (for more information refer to Gateway manual on http://www.invitrogen).

Table 2.11 Gateway LR reaction mixture

MATERIAL	VOLUME (µl)
Entry vector DNA (100 ng/µl)	1.0
Gateway destination vector DNA (pDEST 22) (200 ng/µl)	0.5
10 X TE-buffer pH 8.0	1.5
5X LR Buffer	1
LR Clonase(100 ng/µl)	1
Total volume	5

2.8.10 ANALYSIS OF TRANSFORMANTS

Analysis of transformants clones was done by colony or plasmid DNA PCR using the destination vector adapter and gene specific primers. Colony PCR involved inoculating the PCR reaction mix with the colony material and amplifying for 30 cycles (Table 2.12). General PCR conditions (Table 2.5) were applied except for the annealing temperature which was specific for each amplification.

2.8.11 PLASMID DNA PURIFICATION

Positive clones containing the gene of interest were inoculated into 10ml of LB medium with appropriate antibiotic and grown overnight at 37^0C for plasmid DNA extraction and purification. The purified plasmid was later utilised for

various analyses, including yeast-2-hybrid analysis, transient expressions, BIFC and *E.coli* expression. Plasmid DNA from *E.coli* constructs (e.g. D-TOPO-POB2 and pDEST22/MS1-No-PHD) was purified using the QIAprep Spin Miniprep Kit from 1-5 ml overnight culture grown in LB (Luria-Bertani) medium. The culture was centirfuged at 500 x g for 10 minutes. The pelleted bacterial cells were re-suspending in 250 µl of Buffer P1 and transferred to a microfuge tube. 250 µl of Buffer P2 was added and gently mixed by inverting the tube 4-6 times for less than 5 minutes. Then 350 µl of Buffer N3 was added and the tube was inverted immediately, but gently 4-6 times. The mixture was centrifuged for 10 minutes at 10,000 x g. A QIAprep spin column was then placed in a 2-ml collection tube and the supernatant applied to the QIAprep column by pipetting. The column was then centrifuged for 30-60 seconds. The flow-through was discarded and the QIAprep spin column was washed by adding 0.5 ml of Buffer PB and centrifuged 30-60 seconds. The flow-through was discarded and the QIAprep spin column washed by adding 0.75 ml of Buffer PE and centrifuged for 30-60 seconds. The flow-through was discarded and the QIAprep spin column centrifuged for an additional 1 minute to remove residual wash buffer. To elute the DNA the QIAprep column was placed in a clean 1.5-ml microfuge tube and 30 µl of sterile H_2O was added to the centre of the QIAprep column. The column was left to stand for 1 minute, and centrifuged for 1 minute at 10,000 x g. PCR using appropriate primers were used to confirm the presence of the gene of interest. The plasmid DNA was then stored at -20ºC until required.

2.8.12 DNA SEQUENCING REACTION

Sequencing reactions were performed using 100-200ng template DNA and big dye terminator v1.1 cycle sequencing kit (Applied Biosystems, USA) following the manufactures' instructions. Electrophoresis and data analysis were performed on the ABI 3130 Analyser at Genomics Facility, Queens Medical Centre, University of Nottingham. Sequence analysis was conducted using the software MacVector (MacVector, INC, USA).

Table 2.12 Expression Colony PCR reaction mix

REAGENTS	VOLUME (µl)
Distilled Water	7.6
10x PCR buffer	1
dNTPs (10 mM each)	0.2
MgCl$_2$ 25mM	0.5
Forward primer (10 mM)	0.25
Reverse primer (10 mM)	0.25
Taq enzyme (5U/uL)	0.2
Total volume	10

2.9 PROTEIN METHODS

2.9.1   TOTAL PROTEIN EXTRACTION

Plant tissue from *N. benthamiana* was ground in ice-cold homogenization buffer [25mM Tris-HCl, 150mM NaCl, 1mM EDTA, 10% (v/v) glycerol, 5mM DTT,

0.1% (v/v) Triton X-100, Complete® EDTA-free protease inhibitor tablet cocktail] (http://www.sigmaaldrich.com) using a mortar and pestle, and clarified by centrifugation (12,000g, 4ºC, 15min). The supernatant was transferred to a clean microfuge tube and protein concentrations were determined using the Bradford assay (described in 2.9.3). The total protein fraction was then mixed with 1x SDS loading buffer [25mM Tris-HCl pH6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.001% (w/v) bromophenol blue] and boiled for 5min before being separated by SDS-PAGE (Section 2.10.3).

2.9.2   PROTEIN QUANTIFICATION USING THE BRADFORD ASSAY

2μl of protein extract were mixed with 18μl sterile water and 900μl diluted (5-fold) Bradford reagent (Bio-Rad.com). The mixture was then transferred to a cuvette and incubated at room temperature for 5min. The absorbance of the samples was read at a wavelength of 595nm, against a blank sample (no protein extract added). The concentration of the samples was calculated based on a standard curve obtained by measuring the absorbance of a series of BSA concentrations (200ng/μl, 400ng/μl and 600ng/μl) (Bio-Rad.com).

2.9.3 SEPARATION AND VISUALIZATION OF RECOMBINANT FUSION PROTEINS BY DODECYL SULFATE-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

Protein samples mixed with 1x SDS loading buffer [25mM Tris-HCl pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.001% (w/v) bromophenol blue] were heated to 100ºC for 5min and separated according to their size using SDS-PAGE. A pre-stained protein marker (New England Biolabs) was run alongside the samples and used as a reference. SDS-PAGE gels consist of an upper stacking gel and lower separating gel. The stacking gel [132mM Tris-HCl pH6.8, 4% (w/v) acrylamide, 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.15% (v/v) TEMED] was used for loading and concentrating the protein samples. A 10% (w/v) SDS-PAGE separating gel [0.38M Tris-HCl pH8.8, 10% (w/v) acrylamide, 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.07% TEMED] fractionated the proteins according to their molecular weight. The gels were loaded into a Mini-PROTEAN II cassette (Bio-Rad.com) filled with SDS running buffer [25mM TrisHCl pH8.5, 190mM glycine and 1% (w/v) SDS]. Proteins were first electrophoresed at 80V until they reached the end of the stacking gel, after which the voltage was increased to 150V for 1 hour.

2.9.4 COOMASSIE BLUE STAINING

To visualize proteins on the SDS-PAGE gels, the gels were immersed in Coomassie stain solution (50% (v/v) methanol, 13% (v/v) acetic acid, 0.2% (w/v) Coomassie Brilliant Blue R250) for 2 hours at room temperature with gentle

agitation. The Coomassie stain solution was then poured off and the gel was distained in 20% (v/v) methanol, 5% (v/v) acetic acid and 75% (v/v) distilled H$_2$O.

2.9.5   WESTERN BLOTTING AND IMMUNOLABELLING

PVDF membrane (Bio-Rad.com) was activated by soaking in methanol for 1min before use. After the SDS-PAGE run, proteins were transferred onto PVDF membrane using the Mini-PROTEAN Trans-Blot transfer cassette (Bio-Rad.com) filled with cold transfer buffer [25mM Tris-HCl, 190mM glycine, 20% (v/v) methanol]. Electro-transfer was carried out at 150V for 1h or at 30V overnight. The membrane was then blocked in 5% (w/v) non-fat dried milk dissolved in TBS-T [15mM Tris-HCl pH7.6, 150mM NaCl, 0.1% (v/v) Tween 20] with gentle agitation, to reduce non-specific binding. Then primary antibodies (Tables 2.13 and 2.14) Diluted in TBS-T containing 3% (w/v) non-fat dried milk, were added to the membranes and incubated for 1-3 hours. Membranes were then washed several times with TBS-T for a total of 20min before incubation for 45 min with secondary antibodies. The exact concentrations of the primary and secondary antibodies used in this study are described in Tables 2.13 and Table 2.14. Following the incubation, the membranes were washed five times with TBS-T for a total of 25min.

Table 2.13 Primary antibodies used in this study. All dilutions were made in TBST [15mM Tris-HCl pH7.6, 150mM NaCl, 0.1% (v/v) Tween 20] + 5% (w/v) milk.

ANTIBODY	SOURCE	DILUTION	MANUFACTURER
anti-GFP	rabbit polyclonal	1:5000	Abcam
anti-His	rabbit polyclonal	1:500	Abcam
anti-GST	rabbit polyclonal	1:10000	Agrisera

Table 2.14 Secondary antibodies used in this study. All dilutions were made in TBST [15mM Tris-HCl pH7.6, 150mM NaCl, 0.1% (v/v) Tween 20] + 5% (w/v) milk.

ANTIBODY	SOURCE	DILUTION	MANUFACTURER
goat anti-mouse	horseradish peroxidase conjugated	1:10000	Sigma
goat anti-rabbit	horseradish peroxidase conjugated	1:15000	Sigma
goat anti-rat	horseradish peroxidase conjugated	1:10000	Sigma
rabbit anti-goat	horseradish peroxidase conjugated	1:5000	Sigma

2.9.6   IMMUNODETECTION

Immunodetection was carried out using Immobilon (Millipore) for chemiluminescent detection in accordance with the manufacturer's instructions. The membrane, washed from any unbound HRP-conjugated secondary antibody, was incubated in 1:1 Luminol Reagent and Peroxide Solution for 5min at room temperature. Excess substrate was drained and the membrane was covered in cling film and placed in an X-ray cassette overnight. The membrane was then exposed to X-ray film (Kodak) and the film developed using the X-OMAT developing system.

CHAPTER 3: ABORTED MICROSPORES

(AMS) REGULATORY NETWORKS

3. ABORTED MICROSPORES (AMS) REGULATORY NETWORKS

3.1 INTRODUCTION

Arabidopsis thaliana ABORTED MICROSPORES (AMS) gene encodes a basic helix-loop-helix (bHLH) transcription factor that is required for tapetal cell development and post-meiotic microspore formation (Sorensen *et al*., 2003) However, the regulatory role of AMS in anther and pollen development has not been fully defined (see Chapter 1 Section 1.6). To help determine the role of AMS on pollen and anther development microarray analysis was previously conducted using wild type and *ams* mutant buds from four developmental stages (meiosis, pollen mitosis I, Bicellular and Pollen Mitosis II) (Wilson unpublished data; Xu *et al*., 2010). Genes that were down-regulated or showed significant changes in *ams* buds were selected for further analysis of their role in anther and pollen development; a subset of 13 of these were chosen based on their potential role in pollen wall development for further detailed analysis. SAIL/SALK insertional mutants (Table 3.1) (http://arabidopsis.info/BrowsePage) of these targets were selected for bioinformatics, genotypic, phenotypic and expression analyses.

3.2 BIOINFORMATIC ANALYSES OF AMS REGULATORY TARGETS

Bioinformatics is the branch of biology that deals with the storage, analysis, and interpretation of experimental data (Primrose and Twyman, 2006). The instruments of bioinformatics include computers, databases, statistical tools and algorithms (Primrose and Twyman, 2006). Data mining is defined as the analysis

of data in a database using tools that look for trends or anomalies without knowledge of the meaning of the data. With the current increase in the number of databases, bioinformatics has facilitated the extraction of information and identification of relationships between datasets by using a number of software tools which search sequence databases and measure similarity between sequences to distinguish biologically significant relationships from chance similarities.

The discovering of sequence homology to a known protein or family of proteins often provides the first clues about the function of a newly sequenced gene (Altschul *et al*., 1990). Algorithms are essentially used for data analysis, interpretation of sequence data, and the development and implementation of different types of information (Primrose and Twyman, 2006). Furthermore, bioinformatics has been used for a vast array of other important tasks, including analysis of gene variation and expression, analysis and prediction of gene structure and function, detection of gene regulation networks, simulation environments for whole cell modelling, and presentation and analysis of molecular pathways in order to understand gene interactions (Chen, 2002; Obayashi *et al*., 2011; Wu *et al*., 2011).

One of the most commonly used bioinformatics tools is the Basic Local Alignment Search Tool (BLAST) (Camacho *et al*., 2009). BLAST finds regions of local similarity between sequences. There are different types of BLAST programs and these include blast-n which compares nucleotides against nucleotides, blast-p that compares amino acids and t-blast-n that compares translated sequences to nucleotides. BLAST uses a number of software tools to

search sequence databases for some measure of similarity between sequences to distinguish biologically significant relationships from chance similarities. BLAST then calculates the statistical significance of the matches to infer functional and evolutionary relationships between sequences as well as help identify members of gene families (Altschul *et al*. 1990).

Microarray analysis of RNA extracted from *ams* mutant buds was used to identify genes that may be regulated by AMS. SAIL/SALK insertional mutants with T-DNA potentially within these genes (Table 3.1A-3.1C) were ordered form the Nottingham Arabidopsis stock centre (NASC). Basic description of these genes and scientific literature associated with them were all retrieved from the Ensembl and TAIR databases (Table 3.2).

Table 3.1A Bioinformatics analyses of possible AMS target genes down-regulated or showed significant changes in *ams* buds.

PROTEIN FAMILY	GENE ID	GENE DESCRIPTION AND POSSIBLE FUNCTION	REFERENCES
ABC transporter protein (WBC27)	AT3G13220	Belongs to the ABC transporter family protein identical to putative white-brown complex homolog protein 27 (WBC27) [Arabidopsis thaliana]. It contains InterPro domain ABC-2 type transporter, InterPro domain ABC transporter related and contains InterPro domain AAA+ ATPase, core. Its molecular function suggests ATPase activity, coupled to transmembrane movement of substances. It is expressed during petal differentiation and expansion, expanded cotyledon stage and in seeds, flowers and leaves.	Xu *et al.* (2010) Choi *et al.*, (2011) Dou *et al.*, (2011) Quilichini *et al.*, (2010)
Lipid Transfer Protein (LTP12)	AT3G51590	Encodes a member of the lipid transfer protein family. Proteins of this family are generally small (~9 kD), basic, expressed abundantly and contain eight Cys residues. The proteins can bind fatty acids and acylCoA esters and can transfer several different phospholipids. They are localized to the cell wall. The LTP12 promoter is active exclusively in the tapetum during the uninucleate microspore and bicellular pollen stages.	Arondel *et al.*, (2000) http://www.ncbi.nlm.nih.gov/gene/824322 http://www.arabidopsis.org/servlets/TairObject?id=36903&type=locus
3-Ketoacyl-Coa Synthase 21	AT5G49070	Encodes KCS21, a member of the 3-ketoacyl-CoA synthase family involved in the biosynthesis of VLCFA (very long chain fatty acids).Involved in very-long-chain fatty acid metabolic process, development. Has acyltransferase activity and is expressed during petal differentiation and expansion, anthesis, 4 leaf senescence stage, bilateral stage, expanded cotyledon stage, four leaves visible, eight leaves visible, ten leaves visible, twelve leaves visible. It is expressed in seed, embryo, root, flower, stamen, sepal, petal, leaf, petiole, stem, cauline leaf, leaf whorl.	Millar *et al.*, (1999) http://www.arabidopsis.org/servlets/TairObject?id=132082&type=locus

Table 3.1B Bioinformatics analyses of possible AMS target genes down-regulated or showed significant changes in *ams* buds.

PROTEIN FAMILY	GENE ID	GENE DESCRIPTION AND POSSIBLE FUNCTION	REFERENCES
Protease inhibitor LTP	AT1G66850	Encodes protease inhibitor/seed storage/lipid transfer protein (LTP) family protein; similar to protease inhibitor/seed storage/lipid transfer protein (LTP) family protein [*Arabidopsis thaliana*] (TAIR:AT5G38160.1); similar to Lipid transfer protein [Brassica rapa] (GB:BAA25680.1); contains InterPro domain Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (InterPro:IPR003612); contains InterPro domain Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140); contains InterPro domain Plant lipid transfer protein and hydrophobic protein, helical; (InterPro:IPR013770)	Xing et al., (2007) http://www.arabidopsis.org/servlets/TairObject?id=30426&type=locus
(CYTOCHROME P450)	AT3G26740	Gene function is unknown however, according to TIGR gene annotation and GenBank records its described as a protein coding gene in which transcripts are differentially regulated at the level of mRNA stability at different times of day controlled by the circadian clock. mRNAs are targets of the mRNA degradation pathway mediated by the downstream (DST) instability determinant.	Perez-Amador *et al.,* 2001 http://www.arabidopsis.org/servlets/TairObject?id=38508&type=locus
ACA1, CA^{2+}ATPase	AT1G27770	Encodes a chloroplast envelope Ca2+-ATPase with an N-terminal autoinhibitor. *ATPases* are a class of enzymes that catalyze the decomposition of adenosine triphosphate (ATP) into adenosine diphosphate (ADP) and a free phosphate ion	Huang et al., (1993) Malmstrom et al., (1997) http://www.arabidopsis.org/servlets/TairObject?id=29850&type=locus
Anthocyanidin 3-0-glucosyl transferase; UDP-GLUCOSYL TRANSFERASE 78D2,	AT5G17050	Encodes UDP-glucoronosyl/UDP-glucosyl transferase family protein which is involved in metabolic processes and its molecular function has UDP-glycosyltransferase activity, transferase activity, transferring glycosyl groups, quercetin 3-O-glucosyltransferase activity, flavonol 3-O-arabinosyltransferase activity. It is expressed during petal differentiation and expansion stage, anthesis, bilateral stage, expanded cotyledon stage, mature embryo stage. It is expressed in seed, embryo, flower, carpel, sepal, petal, pedicel and leaf whorl	Peng *et al.*, (2007) http://www.arabidopsis.org/servlets/TairObject?id=131116&type=locus

Table 3.1C Bioinformatics analyses of possible AMS target genes down-regulated or showed significant changes in *ams* buds.

PROTEIN FAMILY	GENE ID	GENE DESCRIPTION AND POSSIBLE FUNCTION	REFERENCES
Cytosolic GADPH (C subunit)	AT3G04120	Encodes cytosolic GADPH (C subunit) involved in the glycolytic pathway but also interacts with H_2O_2 potentially placing it in a signalling cascade induced by ROS. It is involved in gluconeogenesis, glycolysis, response to heat, response to oxidative stress, response to stress, response to sucrose stimulus, response to hydrogen peroxide, fruit development, seed development required for glycolys and response to salt stress as well as cadmium ion. It has glyceraldehyde-3-phosphate dehydrogenase activity, glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity. It is located in the cytosol, mitochondrial envelope, mitochondrion chloroplast, membrane, nucleus, plasma membrane, apoplast and male gametophyte. It is expressed during seedling growth mature pollen stage, germinated pollen stage, petal differentiation and expansion stage, 4 anthesis, 4 leaf senescence stage, C globular stage, D bilateral stage, expanded cotyledon stage, mature embryo stage, two leaves visible, four leaves visible, six leaves visible, eight leaves visible, ten leaves visible, twelve leaves visible.	Jamet *et al*., (2008) http://www.arabidopsis.org/servlets/TairObject?id=40612&type=locus
bHLH protein	AT2G31210	Basic helix-loop-helix (bHLH) family protein; similar to basic helix-loop-helix (bHLH) family protein [Arabidopsis thaliana] (TAIR: AT2G31220.1). It contains InterPro domain Basic helix-loop-helix dimerisation region bHLH and InterPro domain Helix-loop-helix DNA-binding. It is involved in regulation of transcription. It's located in the nucleus with a molecular function in DNA binding. It also has transcription factor activity. It is expressed during petal differentiation and expansion stage and in flower, sepal, leaf whorls.	Riechmann *et al*., (2000) http://www.arabidopsis.org/servlets/TairObject?id=31917&type=locus
bHLH protein	AT1G06170	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein; Functions in DNA binding, sequence-specific DNA binding transcription factor activity. It is involved in regulation of transcription. It is located in the nucleus and expressed in 6 plant structures during anthesis, petal differentiation and expansion stage, expanded cotyledon stage, bilateral stage. It contains Helix-loop-helix DNA-binding domain (InterPro: IPR001092)	Riechmann *et al*., (2000) http://www.arabidopsis.org/servlets/TairObject?id=31224&type=locus
MBK21	AT3G12760	Similar to Calcium ion binding protein [*Arabidopsis thaliana*] (TAIR: AT1G15860.1 and SM10 [*Nicotiana tabacum*] (GB: ABI49160.1). It contains InterPro domain UBA-like; (InterPro: IPR009060); contains InterPro domain protein of unknown function DUF298; (InterPro: IPR005176); contains InterPro domain Ubiquitin-associated; (InterPro: IPR000449	Biswas *et al*., (2007) Wang *et al*., (2008) Ascencio-Ibanez et al (2008)

Table 3.2 Websites utilized for online analysis of gene sequences

ANALYSIS	WEBSITES USED
Gene model and structure (Ensembl)	http://plants.ensembl.org/Arabidopsis_thaliana
Gene finding and sequence viewer (TAIR)	http://www.arabidopsis.org/servlets/sv
Gene expression analysis using Genevestigator (TAIR)	https://www.genevestigator.com/
Description of the molecular changes resulted from the lesions (SIGnAL)	http://signal.salk.edu/
Multiple sequence alignment and design primer design (MacVector)	http://www.macvector.com/

3.3 LOCALISATION OF T-DNA INSERTIONS, PRIMER DESIGN AND MOLECULAR BIOLOGY ANALYSIS OF SAIL/SALK MUTANTS

To construct T-DNA insertional maps of genes in Table 3.1, their sequence information was retrieved from TAIR (http://www.arabidopsis.org/servlets/sv), stored and analysed using MacVector software. MacVector is a comprehensive application that provides sequence editing, primer design, internet database searching, protein analysis, sequence confirmation, multiple sequence alignment,

phylogenetic reconstruction, coding region analysis, and a wide variety of other functions (http://www.macvector.com/). MacVector was used for multiple sequence alignment and gene specific primer design. Primers were designed for PCR genotyping and gene specific RT-PCR (expression analysis) (Table 3.3 and 3.4). MacVector was used to produce maps which consisted of the 3' and 5' UTR ends of the gene of interest, their upstream and downstream genes, the left and right primer and T-DNA insertion site (Figure 3.1).
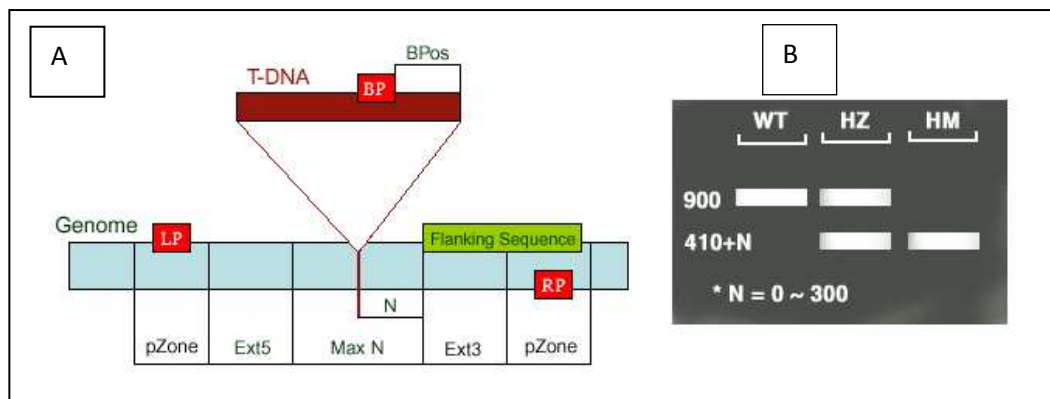
Figure 3.1: **(A)** illustrates a guide to identifying T-DNA insertion localisation and map. N-Difference of the actual insertion site and the flanking sequence position, usually 0 - 300 bases. MaxN- Maximum difference of the actual insertion site and the sequence, default 300 bps. Ext5, Ext3-regions between the MaxN to pZone, reserved not for picking up primers. LP, RP - Left, Right genomic primer. BP T-DNA border primer LB - the left T-DNA border primer. BPos - The distance from BP to the insertion site. LB - Left border primer of the T-DNA insertion: (LBb1.3 and LBa1) **(B)** Guide for the PCR genotyping and gel analysis of the insertion lines for the identification of wild type (WT), heterozygous (HT) and homozygous (HM) lines provided by the SALK institute genomic analysis laboratory (SIGnAL) (http://signal.salk.edu/tdnaprimers.2.html).

Table 3.3 Primers used for PCR genotyping of *Arabidopsis* T-DNA insertion lines.

POSSIBLE FUNCTION	GENE ID	PRIMER KO CODE	FORWARD PRIMER (5 ' - 3')	REVERSE PRIMER (5 ' - 3')
ABC transporter	AT3G13220	SALK_062317F SALK_062317R	TTTTCCAGAATTGGCTTTGTG	GACTTCCACCAGAGAAGACCC
LTP12	AT3G51590	SALK_052271F/R	GATCAGACTCTACTGCCA	TTGTACCAATTGGCTCAG
3-Ketoacyl-Coa synthase 21	AT5G49070	SALK_089611F/R	GGCAACAAGTCCTTTGATTTG	CTTATTGTTAGCACCGAAGCG
Protease inhibitor LTP	AT1G66850	SALK_018008 F/R	ATGATTGGTTCACGTCGTCTC	AAGTGGGATAAGCAACATTGC
Cytochrome p450	AT3G26740	SAIL_743_725F SAIL_743_1615R	AAGGGGGATGCAATGGATATTGT	CAATCGAGACGACGAGACAGTGA
ACAI CA2 =ATPase	AT1G27770	SALK_107728_6572 AT1G27770_7521R	GATCTCCAGTGGGAAGGAAAG	TCGACCGTTGAGTGTAATTCC
Anthocyanidin 3-0-glucosyl transferase	AT5G17050	SALK_049338	CTCTTCGTTATTTTCCTCCGG	TCAAACCCATCTTTCGTGAAG
Cytosolic GADPH (C subunit)	AT3G04120	SALK_129085F SALK_129085R	CCTTGACATTGTCTCCAACG	AAACTCCCATCACCTCACATG
bHLH	AT2G31210	SALK_022241F SALK_022241R	ACTGAGAGCAAGCAAGATTGC	TTCTGTTTGATGGAGTTTGG
bHLH	AT1G06170	SALK_090231F SALK_090231R	CCAAAGTCCAGTCCAGAGTACC	GAAACTCCCATTAGGAGGCTG
MBK21	AT3G12760	SALK_143248L_R SALK_143248R_F	GTGCGAGTGAGAATGAAGGAG	TGTAAAATGTTTCAGGGGCA
N/A	T-DNA insertion	LBb1.3	ATTTTGCCGATTTCGGAAC	N/A
N/A	T-DNA insertion	LBa1	TGGTTCACGTAGTGGGCCATCG	N/A

Table 3.4 Gene specific primers used for RT-PCR analysis in T-DNA insertion lines.

POSSIBLE FUNCTION	GENE ID	PRIMER KO CODE	FORWARD PRIMER (5 ' - 3')	REVERSE PRIMER (5 ' - 3')
ABC transporter	AT3G 13220	At3g13220_624F At3g13220_934R	CAAAGAGCTCG GCCTCGAAA	TGCCGTAAA AGGCAGGGT GT
LTP12	AT3G 51590	At3g51590_53F At3g51590_355R	TCCCATCCCCA ACAGAGTCA	CGGCAGTCG ATATACTGT CGCAAT
3-Ketoacyl-Coa synthase 21	AT5G 49070	At5g49070_194F At5g49070_627R	ACCGCGTCCCA ATGTCAACT	AGCCAATGA GCCTCGGTG AG
Protease inhibitor LTP	AT1G 66850	At1g66850_2F At1g66850_301R	TGAAGATCGTG ACATTGGTACT CG	TGGGATAAG CAACATTGC AAGC
Cytochrome p450	AT3G 28740	SAIL_743_725F SAIL_743_1615R	AAGGGGGATG CAATGGATATT GT	CAATCGAGA CGACGAGAC AGTGA
ACAI CA2 =ATPase	AT1G 27770	SALK_107728_6572 AT1G27770_7521R	GATCTCCAGTG GGAAGGAAAG	TCGACCGTT GAGTGTAAT TCC
Anthocyanidin 3-0-glucosyl transferase	AT5G 17050	At5g17050_778F At5g17050_111R	GTGCACGATCC TCACGGTTG	CGGTACACC TGCCGACAC AC
Cytosolic GADPH (C subunit)	AT3G 04120	At3g04120_351F At3g04120_725R	GGGTGGTGCCA AGAAGGTTG	ACGGTTGGA CACGGAAA GA
bHLH	AT2G 31210	AT2g31210_62F AT2g31210_42R	GTGCGGCGGA AACAACTTTT	TGAAGCAAG AACCCGCCA TT
bHLH	AT1G 06170	AT1G06170_16F AT1G06170_47R	CATTCCACCTC AGCCCACAA	GCGGCGACA GAGGAAGA TGT
MBK21	AT3G 12760	AT3G12760_10F AT3G12760_48R	GGCACCTTGAA GCAGCGTTT	TCTCTTTCG CCCACCCAA AA

MacVector was also used to verify gene specific primers designed by T-DNA Primer Design Tool powered by Genome Express Browser Server (GEBD) (http://signal.salk.edu/tdnaprimers.2.html). T-DNA Primer Design Tool returned insertion location information and the estimated T-DNA confirmation product size. The PCR reaction was set up by using the three primers (LB: Left border primer of the T-DNA insertion, LP: Left genomic primer and RP: Right genomic primer) (Figure 3.1) and two separate reactions were performed for wild type sequences and T-DNA insert confirmation. Sections 3.4.1 to 3.4.12 show the T-DNA insertion localization maps for all SALK KO lines, genotypic analyses and expression profiles from Genevestigator

3.4 ANALYSIS OF SALK INSERTION LINES

3.4.1 T-DNA INSERTIONAL MAP, GENOTYPIC AND EXPRESSION ANALYSIS OF AT3G13220 : SALK_062317 T-DNA KO

Initially there was very little information on the ATP-binding cassette transporter G26 (ABCG26) protein; however more recently a number of publications have provided more information as to the role of ABCG26. It is expressed in the anther, leaves, flowers, seeds and sepals during expanded cotyledon stage and petal differentiation stage. Generally, ABC transporters belong to the ATP-Binding Cassette (ABC) superfamily, which uses the hydrolysis of ATP to energise diverse biological systems. The ATP-Binding Cassette (ABC) superfamily forms one of the largest of all protein families with a diversity of physiological functions (Saurin *et al*., 1999). As reported by Kuromori *et al*., (2011), AT3G13220

encodes an ATP-binding cassette transporter G26 (ABCG26) involved in tapetal cell and pollen development. It is required for male fertility and pollen exine formation. It is involved in ATP catabolic processes, pollen development, pollen exine formation, pollen maturation and transmembrane transport. It was subsequently found to be located in the plasma membrane and has ATPase activity, coupled to transmembrane movement of substances. SALK_062317 is an insertional mutant with a T-DNA putatively in At3g13220; the T-DNA insertion and primer positions are shown in Figure 3.2.
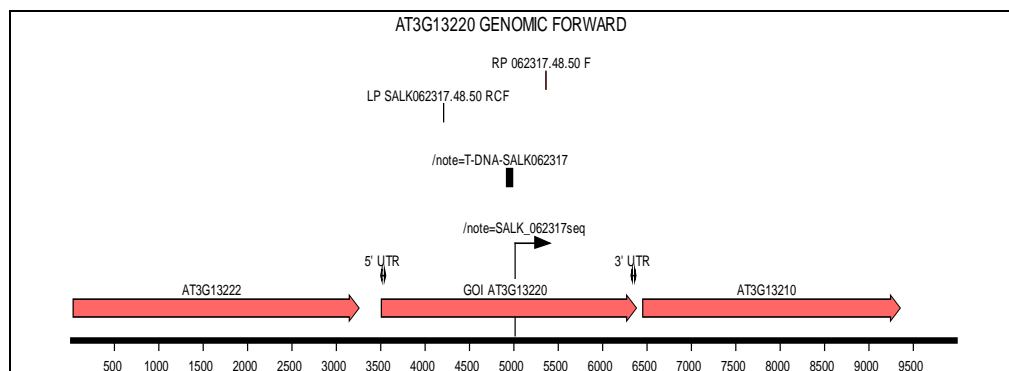
Figure 3.2: T-DNA localisation map of SALK_062317 (AT3G13220). Red bold arrows represent the direction of the coding sequence of AT3G13220, its upstream gene (AT3G13222) and downstream gene (AT3G13210). Black thick vertical line show the sequenced T-DNA flanking region, black arrowhead shows the approximate T-DNA insertion point, and black vertical thin lines show the Left primer (LP) and Right Primer (RP) used for AT3G13220 expression analysis.

Genotyping identified three positive homozygous plant lines (Figure 3.3). The gene specific RT-PCR amplification showed no expression of At3g13220 in two

of the homozygous plant lines (Figure 3.4). This result means that gene expression was completely silenced because of the T-DNA insertion and as a result a significant change in its expression was seen compared to the wild type (Col) cDNA. Actin RT–PCR was used as a housekeeping control for detecting for gene expression by RT-PCR (Figure 3.4). The wild type (Col) cDNA and genomic DNA (GN) were used as positive controls. Phenotypic analysis of the KO line was conducted (section 3.4.2).
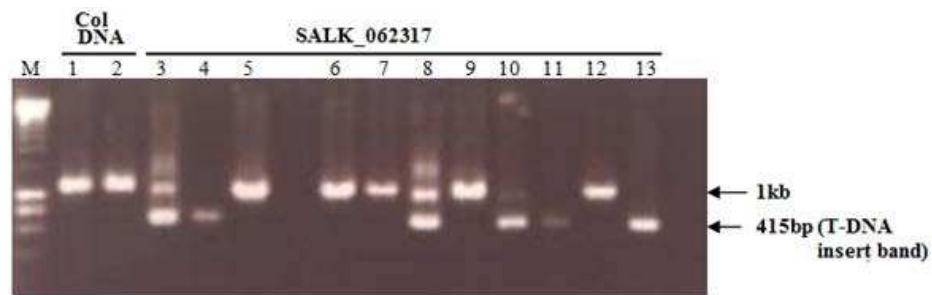
Figure 3.3 Genotypic analysis of *Arabidopsis* wildtype Columbia ecotype and the SALK_062317 (AT3G13220) line (two genes specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular marker. Wildtype Col genomic DNA was used a positive control (numbered 1). Out of total number of twelve SALK_062317 KO plants analysed five segregated as wildtype (numbered 5, 6, 7, 9 and 12), three as heterozygous (numbered 3, 8 and 10) and three as homozygous with 415bp size of the T-DNA (numbered 4, 11 and 13). M=Molecular marker.
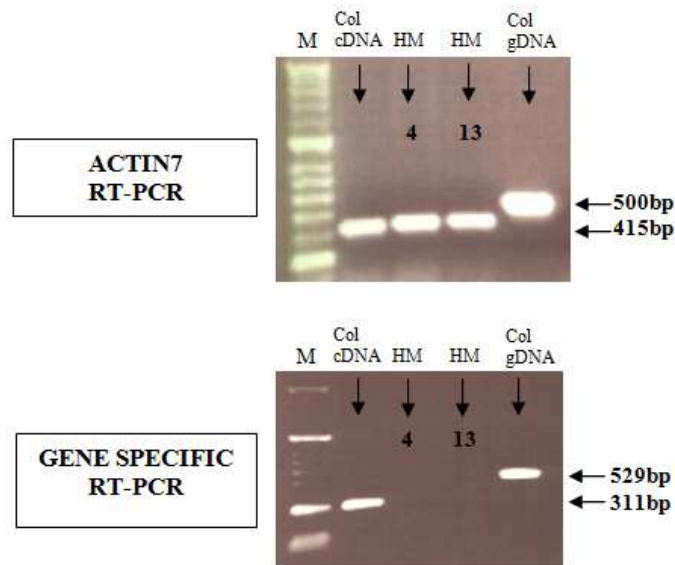
Figure 3.4 Gene expression analyses of *Arabidopsis* wildtype Columbia ecotype and mutant inflorescences of At3g13220. RT-PCR was carried out using 5µg of total RNA extracted from *Arabidopsis* inflorescence tissue at 6 weeks old (Primers used refer to Table 3.4). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder II (Bio-line.com) as molecular weight marker Expression analysis was compared to ACTIN7. The gene specific RT-PCR amplification showed the two homozygous mutants (numbered 4 and 13) as completely down regulated. (M=Molecular marker, Col cDNA= Columbia cDNA, HM=Homozygous mutant cDNA and Col gDNA= Columbia genomic DNA).

3.4.1.1 PHENOTYPIC AND MICROSCOPIC ANLAYSIS OF AT3G13220 : SALK_062317 T-DNA INSERTIONAL MUTANT

Phenotypic characteristics of the mutant plants were carried out by visual and microscopic observations of general plant growth characteristics and of the

reproductive organs. Genevestigator expression analysis, gene function and description provided an indication of the location of gene expression and other gene families to which they were identical or similar to in other plant species. Vegetative plant development and flower formation proceeded normally in the SALK_062317 lines compared to the wild type. However, the SALK_062317 homozygous plants exhibited sterility as demonstrated by a lack of silique expansion (Figure 3.5C and 3.5D) compared to the wildtype Columbia ecotype (Figure 3.5A and 3.5B). In addition to male sterility phenotype SALK_062317 homozygous plants also showed late petal abscission as flowering was prolonged compared to the wildtype. Abscission is an active process and has a variety of roles during plant development. Plant parts such as pollen, fruits, seeds, and leaflets may be shed in response to developmental cues to guarantee efficient dispersal or propagation of the plant. Hormones such as ethylene and auxin have long been associated with regulating abscission, and levels of ethylene have been shown to be higher within abscission zone tissues. Identification and characterization of delayed floral abscission phenotype is another aspect that could be studied further in this mutant. Mature SALK_062317 mutant flowers contained shrivelled anthers that failed to dehisce (Figure 3.6C and 3.6D) compared to wild-type flowers, which exhibited dehisced anthers containing pollen (see arrows in Figure 3.6A and 3.6B). Female fertility did not appear effected since seed set could be achieved by fertilisation using wild type pollen. Immature wild-type anthers contained newly released microspores (Figure 3.7A) in contrast to the equivalent stage of immature SALK_062317 mutant anthers (Figure 3.7B) that contained abnormal immature microspores, which lacked the defined shape of the wild-type microspores.

Pollen viability tests using Alexander's stain (Alexander, 1969) on pollen from open flowers showed viable pollen stained pink in wildtype anthers (Figure 3.7C) and reduced viable pollen in SALK_062317 anthers (Figure 3.7D). This result showed that down-regulation of expression of the *ABC transporter (WBC27)* had a significant effect on pollen development with no viable pollen was produced in the absence of *WBC27* expression.
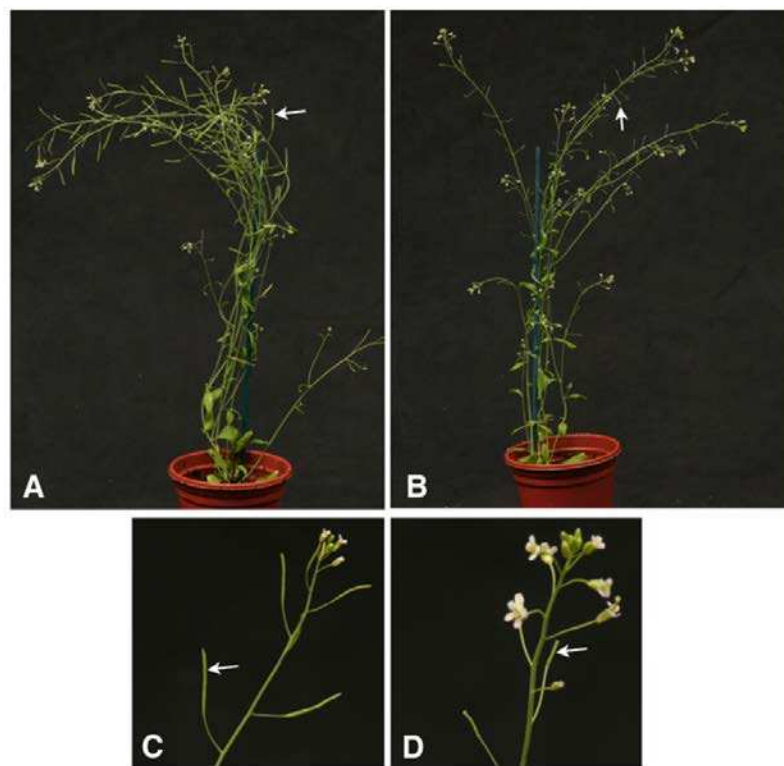
Figure 3.5: Phenotypic analysis of *Arabidopsis* wildtype Columbia ecotype and mutant inflorescences of SALK_062317 in AT3G13220 **(A)** and **(C)** Columbia wildtype plant with normal silique development (arrows). **(B)** and **(D)** The SALK_062317 homozygous mutant, showing reduced silique elongation,

Figure 3.6 Analyses of *Arabidopsis* wildtype Columbia ecotype and mutant inflorescences of SALK_062317 (AT3G13220). **(A)** and **(B)** show wild-type flowers containing dehisced anthers (Arrows). **(C)** and **(D)** show mature SALK_062317 mutant flowers containing shrivelled anthers that fail to dehiscence. Magnification in (**A**) and (**C**) = X10. Magnification in (**B**) and (**D**) = (X40).
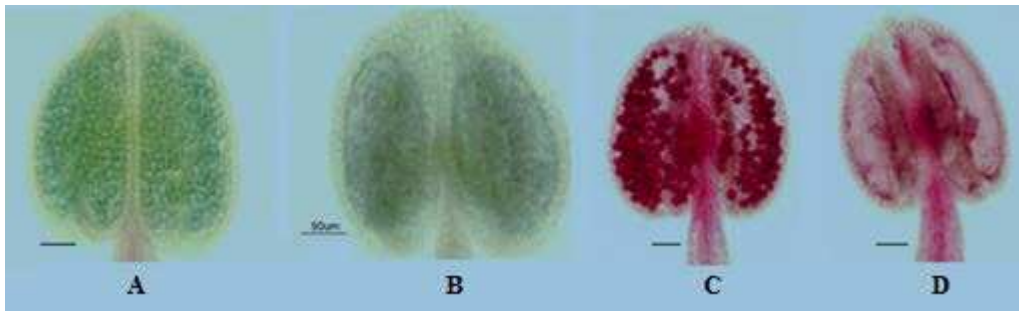
Figure 3.7 Analysis of pollen viability in *Arabidopsis* wildtype Columbia ecotype and mutant inflorescences of SALK_062317 (AT3G13220) using Alexander staining. **(A)** Immature wild-type anther containing newly released microspores. **(B)** Immature SALK_062317 mutant anther containing abnormal immature microspores, which lack the defined shape of the wild-type microspores. Excessive amounts of material are visible in the locule, suggesting abnormal tapetal development. **(C)** Alexander-stained mature wild-type anther showing viable pollen grains. **(D)** Alexander-stained mature SALK_062317 mutant anther lacking viable pollen. Bars = 50 µm.

3.4.2 T-DNA INSERTIONAL MAP, GENOTYPIC AND EXPRESSION ANALYSIS OF AT3G51590 (LIPID TRANSFER PROTEIN 12): SALK_052271 T-DNA KO

At3g51590 encodes a member of the lipid transfer protein family. Proteins of this family are generally small (~9 kD), basic, expressed abundantly and contain eight Cys residues. The proteins can bind fatty acids and acylCoA esters and can transfer several different phospholipids. They are localized to the cell wall. The LTP12 promoter is active exclusively in the tapetum during the uninucleate microspore and bicellular pollen stages (http://www.ncbi.nlm.nih.gov/gene/824322). Genevestigator expression profile indicates that maximal expression of the gene during plant development is in floral organs (https://www.genevestigator.com) (Figure 3.8).
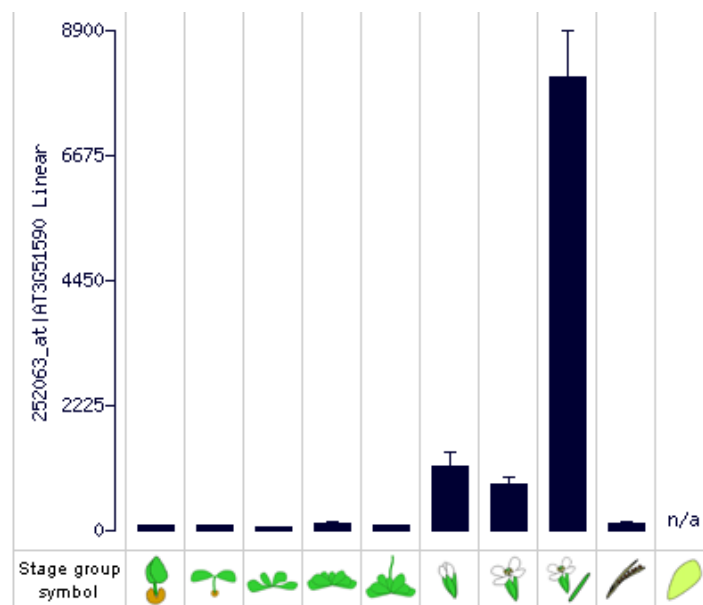
Figure 3.8 Genevestigator expression profile indicates that maximal expression of AT3G51590 during plant development is in floral organs (https://www.genevestigator.com).

SALK_052271 T-DNA insertion and primer positions are shown in Figure 3.9. Genotyping of SALK_052271 T-DNA KO plants identified 5 homozygous lines alongside 2 Columbia wildtypes (Figure 3.10 lanes 1-2). Actin RT–PCR was used as background control for detecting for gene specific RT-PCR (Figure 3.10). Gene specific RT-PCR revealed the expression level of the gene in the homozygous plants was reduced, but expression still occurred (Figure 3.11). Therefore, despite homozygous insertions in SALK_052271, expression of At3g51590 were still seen, although this was reduced compared to wild type (Figure 3.11, sample 4). None of the 5 homozygous KO lines had a male sterility, or altered vegetative phenotype when compared to the wild type (Col). The lack of the mutant phenotype could be due to gene redundancy or position of the T-DNA insertion in the 5' end of the gene allowing the gene to be expressed.

Figure 3.9 T-DNA localisation map of SALK_052271 (AT3G51590). Purple bold arrow represent the direction of the coding sequence for the gene of interest AT3G51590, its upstream gene (AT3G51600) and downstream gene (AT3G51680) are represented by bold green arrows pointing in the direction of

the coding sequence. Black thick vertical line shows sequenced T-DNA flanking region, black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotyping and expression analysis (refer to Table 3.3 and 34 for detailed primer information).
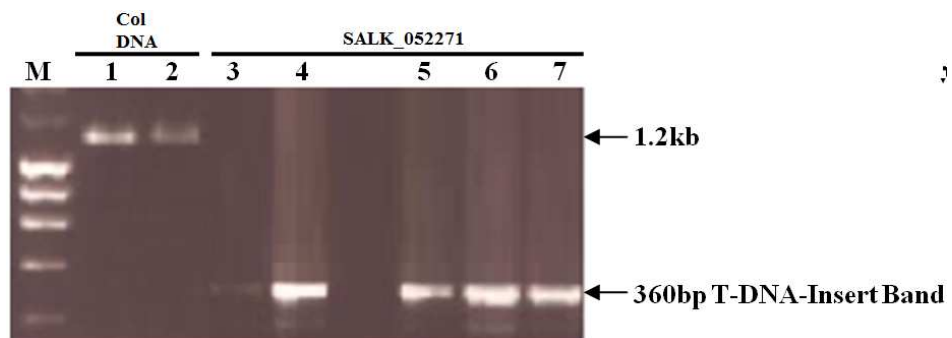
Figure 3.10 Genotyping analysis of *Arabidopsis* wildtype Columbia ecotype and SALK_052271 (AT3G51590) mutant (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular marker. Two Wildtype Col genomic DNA were used as positive controls (numbered 1-2). Five lines numbered 3-7 were identified as homozygous with 415bp size of the T-DNA insert. (M=Molecular marker, Col cDNA= Columbia cDNA, HM=Homozygous mutant cDNA and Col gDNA= Columbia genomic DNA).
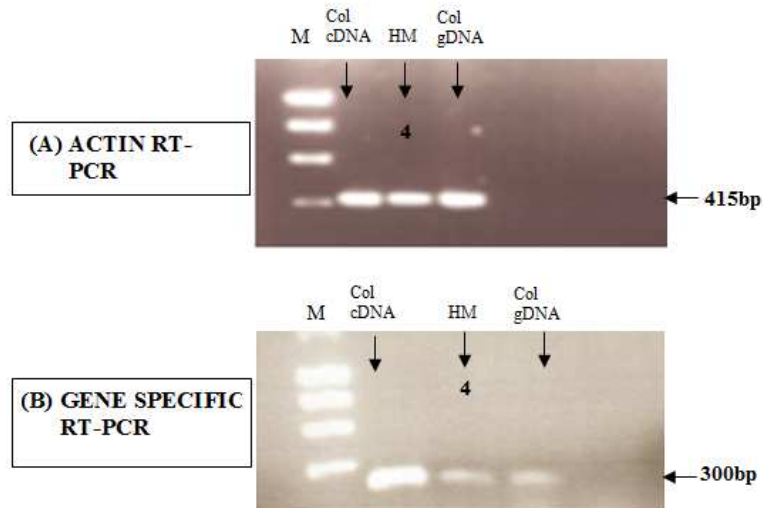
Figure 3.11 Gene expression analyses by RT-PCR of AT3G51590 in inflorescences from *Arabidopsis* wildtype Columbia ecotype and mutant inflorescences of SALK_052271 (For primers used refer to Table 3.4). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. (**A**) Expression analysis was normalized to ACTIN7. (**B**) Gene specific RT-PCR amplification of SALK_052271 (sample 4) alongside the wildtype Columbia ecotype cDNA and genomic DNA positive controls. (M= Molecular marker, Col cDNA=Wildtype Columbia cDNA, HM=Homozygous mutant cDNA (sample 4). Col gDNA= Columbia genomic DNA. The size of cDNA product is predicted to be the same as genomic DNA because of lack of introns in the region amplified.

3.4.3 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT5G49070 (3-KETOACYL-COA SYNTHASE 21): SALK_089611 T-DNA

AT5G49070 encodes KCS21, a member of the 3-ketoacyl-CoA synthase family involved in the biosynthesis of VLCFA (very long chain fatty acids).

The Genevestigator profile (Figure 3.12) shows that the gene is expressed throughout plant growth, but is highest in floral organs (https://www.genevestigator.com).

Figure 3.12 Genevestigator expression profile indicates that maximal expression of AT5G49070 during plant development is in floral organs (https://www.genevestigator.com).

SALK_089611 T-DNA insertion and primer positions are shown in Figure 3.13. Genotyping (Figure 3.14) of SALK_089611 T-DNA KO plants identified all 8 plants as segregating wildtype lines with no male sterility phenotype and therefore no further analysis was carried out on this SALK line. The lack of a mutant phenotype could have been due to gene redundancy; further analysis could be done on this line by using mutant combinations if they belong to a small gene

family, or by using RNAi to silence multiple gene members by targeting conserved regions in the gene family.
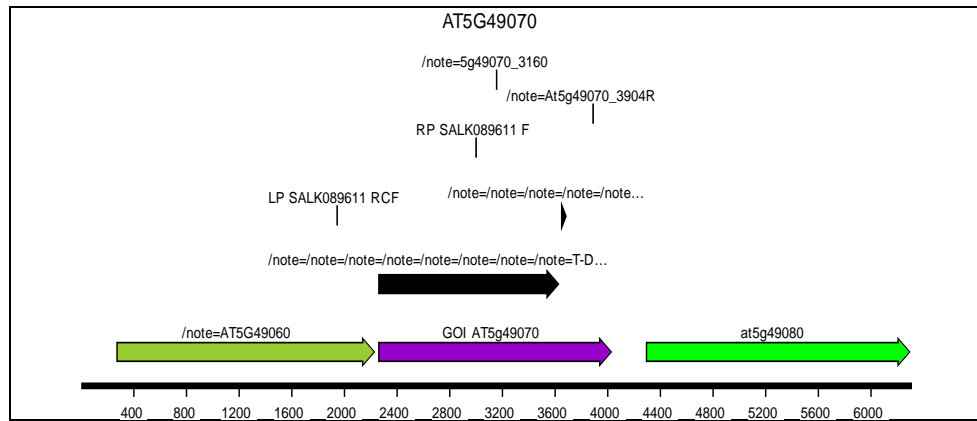
Figure 3.13 T-DNA localisation map of SALK_089611 (AT5G49070). Purple bold arrow represent the direction of the coding sequence for the gene of interest AT5G49070, its upstream gene (AT3G49060) and downstream gene (AT3G49080) are represented by bold green arrows pointing in the direction of the open reading frame. Black bold arrow shows sequenced T-DNA flanking region, black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic and gene expression analysis.
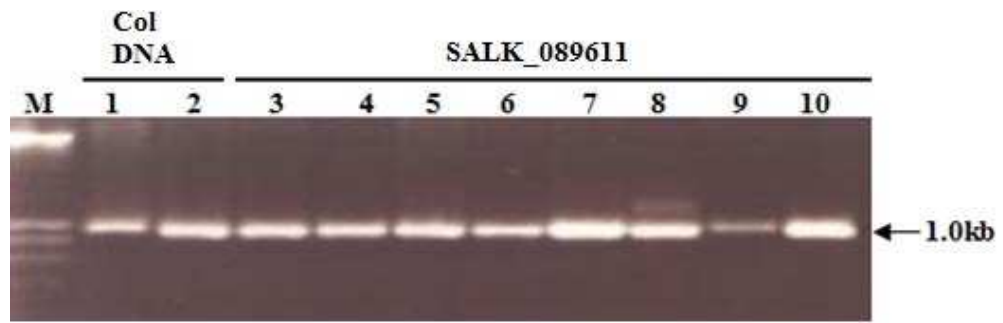
Figure 3.14 Genotyping of SALK_089611 (AT5G49070) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Wildtype Columbia ecotype was used as positive control (numbered 1-2). Genotyping identified all eight SALK_089611 lines segregating as wildtype (numbered 3-10). (M=Molecular marker, Col cDNA= Columbia cDNA, and Col gDNA= Columbia genomic DNA).

3.4.4 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT1G66850 (PROTEASE INHIBITOR/SEED STORAGE/LIPID TRANSFER PROTEIN): SALK_018008 KO

AT1G66850 encodes a protease inhibitor/seed storage/lipid transfer protein (LTP) family protein. The Genevestigator profile shows that the gene is expressed throughout plant growth but is highest in floral organs (Figure 3.15). SALK_018008 T-DNA insertion and primer positions are shown in (Figure 3.16). A total of 11 SALK_018008 plants were analysed and 3 of them segregated as wildtype and 6 were heterozygous (Figure 3.17). Heterozygous lines could be grown further to allow analysis of recessive alleles after segregation.
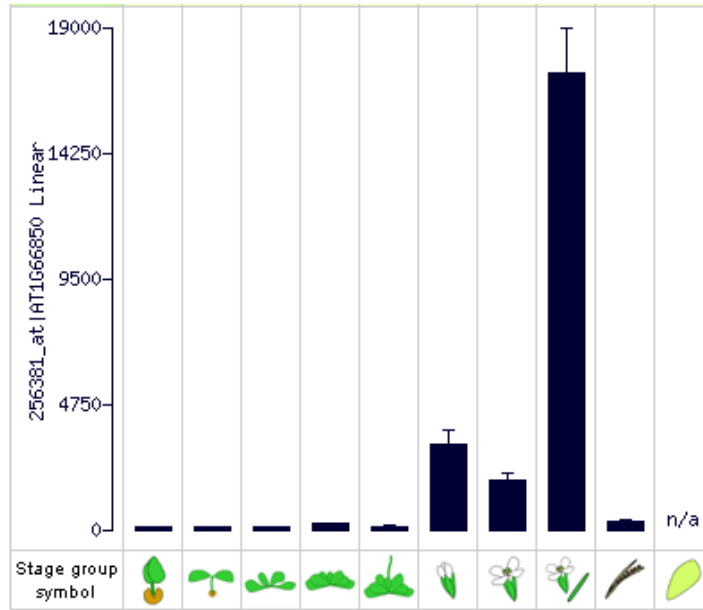
Figure 3.15: Genevestigator expression profile indicates that maximal expression of AT5G66850 during plant development is in floral organs (https://www.genevestigator.com).
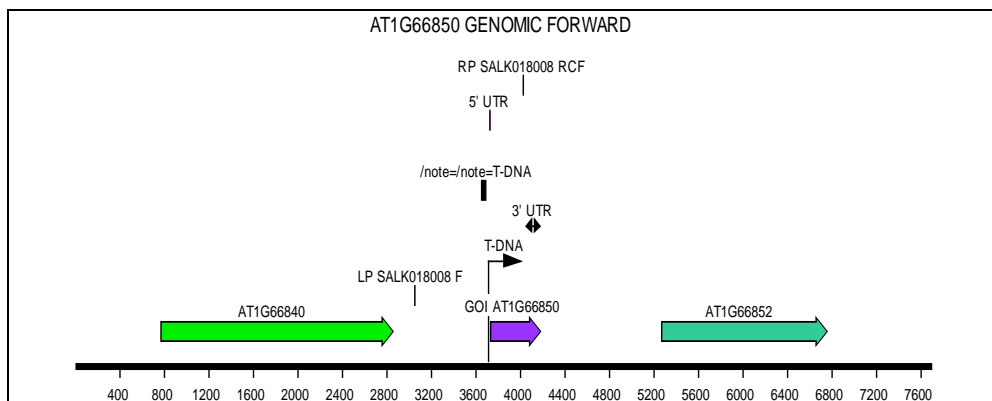
Figure 3.16: T-DNA localisation map of SALK_018008 AT5G66850. Purple bold arrow represent the direction of the coding sequence for the gene of interest At1g66850, its upstream gene (At1g66840) and downstream gene (At1g66852) are represented by bold green arrows pointing in the direction of the coding sequence . Black bold arrow shows sequenced T-DNA flanking region, black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis.

Figure 3.17 Genotyping analysis of Arabidopsis wildtype Columbia ecotype and SALK_018008 (AT5G66850) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Wildtype Col genomic DNA was used a positive control (numbered 1) and Out of 11 SALK_018008 lines analysed four segregated wildtype (numbered 3, 10, 11 and 12), seven were heterozygous (numbered 2, 4, 5, 6, 7, 8 and 9). M=Molecular marker, WT Col DNA=Wildtype Columbia genomic DNA.

3.4.5 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT3G28740 (CYTOCHROME P450): SAIL_743

AT3G28740 encodes a member of the cytochrome p450 family protein. Genevestigator expression profile shows that the gene is expressed throughout plant growth but is highest during the seedling stage (Figure 3.18). SAIL_743 T-DNA insertion and primer positions are shown in Figure 3.19. A total of 14 plants were genotyped and all of them segregated as wildtype when compared to Columbia (Figure 3.20). Consequently no further analysis was carried out on this line. In future more plants from this mutant could be screened for homozygous individuals exhibiting male sterile phenotypes.
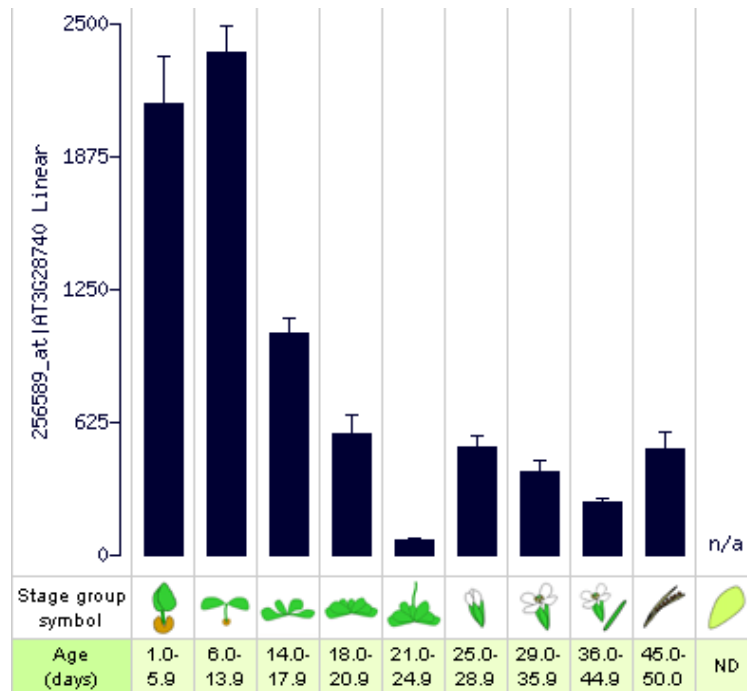
Figure 3.18 Genevestigator expression profile indicates that maximal expression of At3g28740 during plant development is during seedling stages (https://www.genevestigator.com.
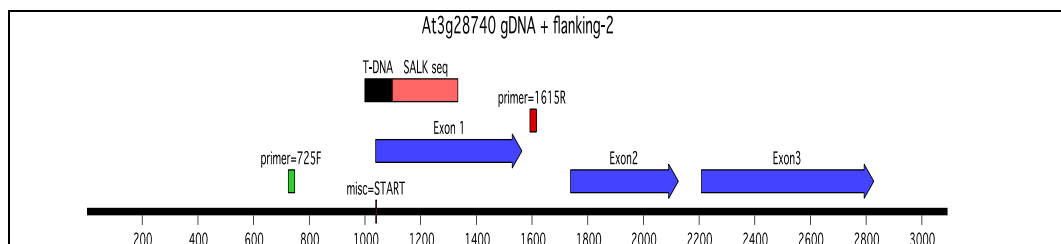
Figure 3.19 T-DNA localisation map in SAIL_743 (AT3G28740). Purple arrows represent exons and the direction of the coding sequence for the gene of interest AT3G28740. Black horizontal box shows approximate T-DNA insertional position. Red horizontal box shows sequenced T-DNA flanking region, green and red vertical box show Left primer (LP) and Right Primer (RP) respectively used for genotypic analysis.
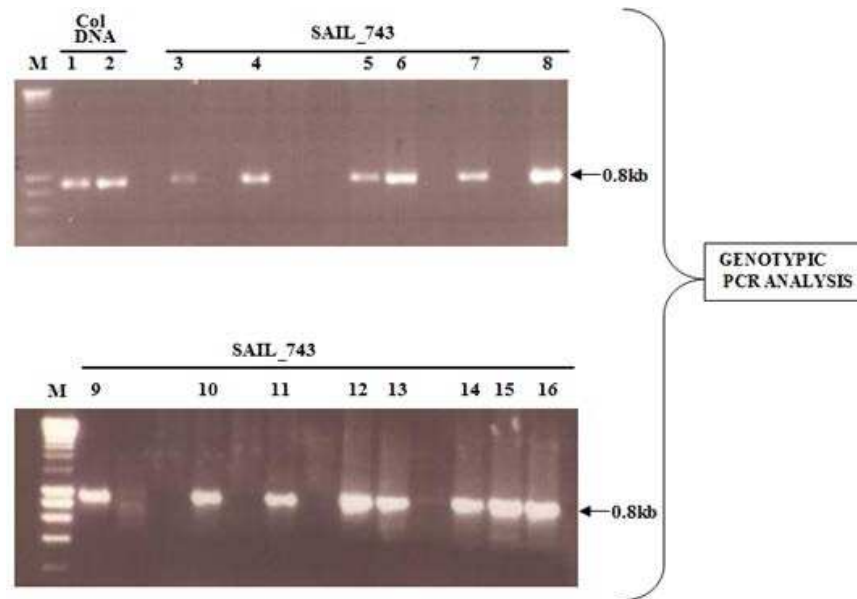
Figure 3.20 Genotyping analysis of SAIL_743 (AT3G28740) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker**.** Genotyping PCR identified all 14 SAIL_743 as segregating wildtype (numbered 3-16). Wildtype Columbia ecotype was used as positive control (numbered 1-2). M=Molecular marker, WT Col DNA=Wildtype Columbia genomic DNA.

3.4.6 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT1G27770 AUTO-INHIBITED CA2+-ATPASE 1/ PLASTID ENVELOPE ATPASE 1: SALK_107728

At1g27770 encodes a chloroplast envelope Ca2+-ATPase with an N-terminal auto inhibitor. The Genevestigator profile shows that the gene is expressed significantly throughout plant growth (Figure 3.21). SALK_107720 T-DNA insertion and primer positions are shown in (Figure 3.22). Genotypic analysis

showed that 3 plants were heterozygous and 11 were homozygous (Figure 3.23). Homozygous lines showed no male sterile phenotype. This could have been due to position of the T-DNA insertion in the 3' end of the gene allowing it to be expressed. An alternative approach would be to analyse At1g27770 with the T-DNA insertion in the exon, or by using alternative mutagenesis options e.g. RNAi gene silencing.
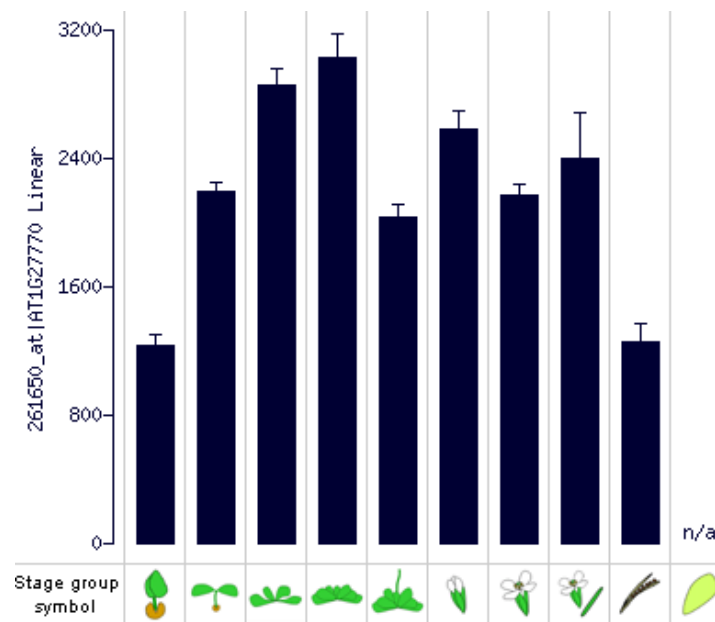
Figure 3.21 Genevestigator expression profile indicates that expression of AT1G27770 occurs throughout plant development (https://www.genevestigator.com).
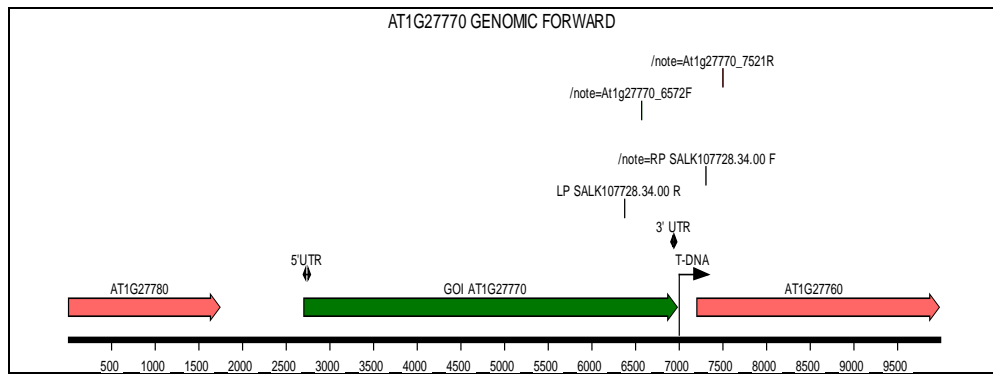
Figure 3.22 Figure 3.16: T-DNA localisation map of SALK_107728 (AT1G27770). Green bold arrow represent the direction of the coding sequence for the gene of interest AT1G27770, its upstream gene (AT1G27780) and downstream gene (AT1G27760) are represented by bold red arrows pointing in the direction of the coding sequence. Black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis.
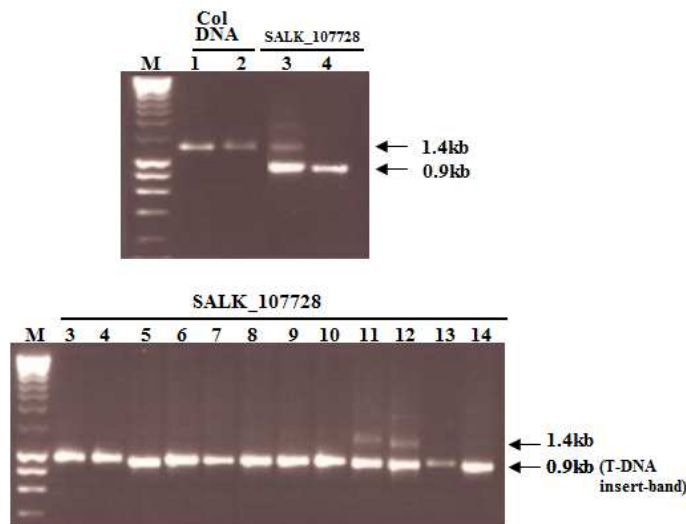
Figure 3.23 Genotyping analysis of SALK_107728 (AT1G27770) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Genotyping PCR identified three lines as heterozygous (numbered 3, 11 and 12) and eleven as homozygous with 900bp size of the T-DNA insert. Wildtype Columbia ecotype was used as positive control (numbered 1-2). M=Molecular marker.

3.4.7 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT5G17050 ANTHOCYANIDIN 3-O GLUCOSYLTRANSFERASE: SALK_049338

AT5G17050 encodes an Anthocyanidin 3-O-glucosyltransferase which specifically glucosylates the 3-position of the flavonoid C-ring. The Genevestigator profile indicates that the gene is expressed throughout plant development (Figure 3.24). SALK_049338 T-DNA insertion and primer positions are shown in Figure 3.25. All 10 plants genotyped were identified as heterozygous (Figure 3.26). In future allowing this generation to segregate would provide homozygous lines which could then be analysed for male sterile phenotype.
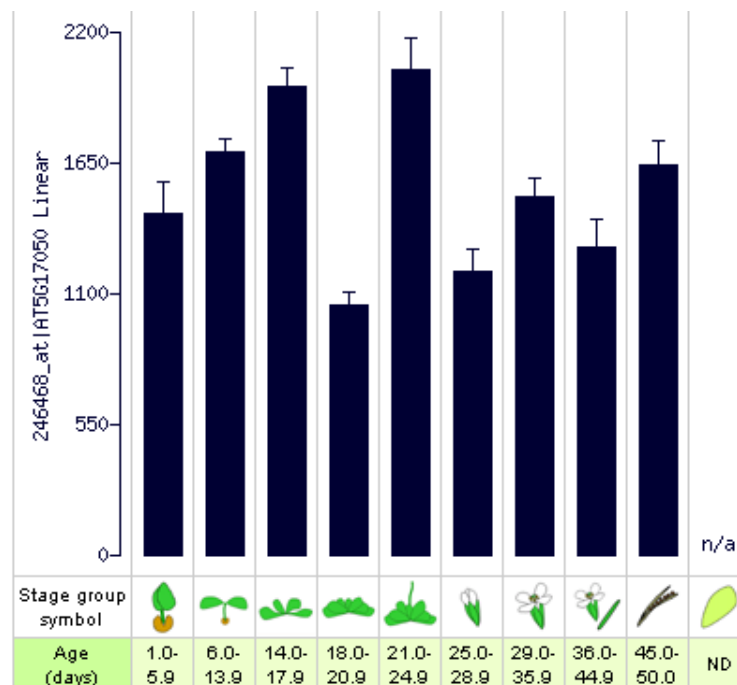
Figure 3.24 Genevestigator expression profile indicates that expression of AT5G17050 occurs throughout plant development (https://www.genevestigator.com).

84
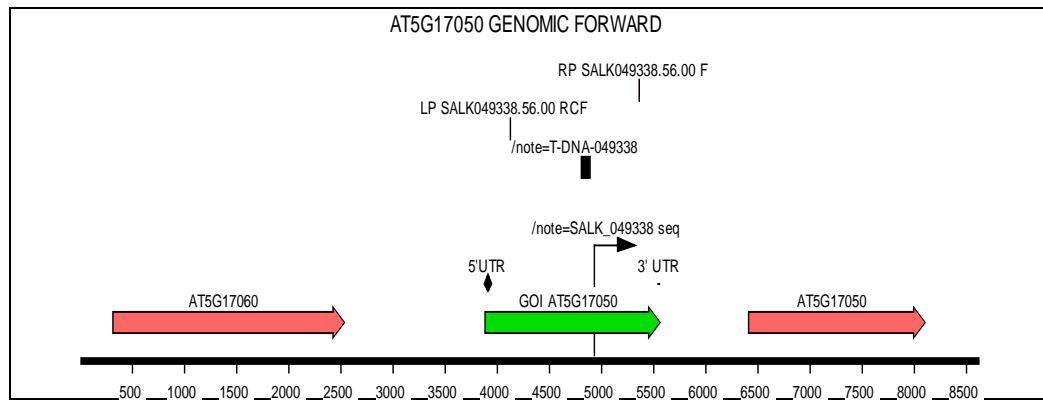
Figure 3.25 T-DNA localisation map of SALK_049338 (AT5G17050). Green bold arrow represent the direction of the open reading frame for AT5G17050, the upstream gene (AT5G17060) and downstream gene (AT5G17080) are represented by red arrows pointing in the direction of the coding sequence. Black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotype analysis.
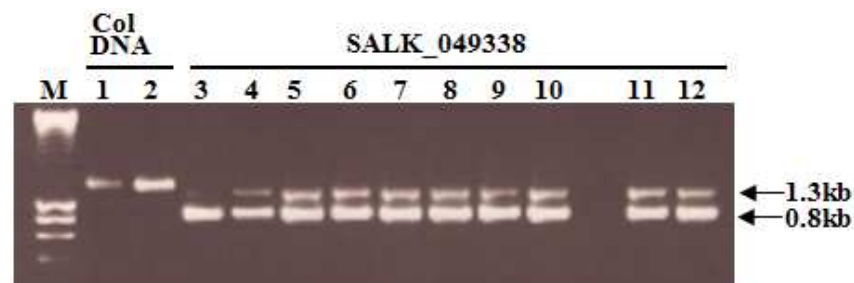
Figure 3.26 Genotyping analysis of SALK_049338 (AT5G17050) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker**.** Genotyping PCR identified all SALK_049338 lines as heterozygous (numbered 3-12). Wildtype Columbia ecotype was used as positive control (numbered 1-2). M=Molecular marker.

3.4.8 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT3G04120 (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE-C- SUBUNIT): SALK_129085

AT3G04120 encodes for a Cytosolic GADPH (C subunit) involved in the glycolytic pathway, but also interacts with H_2O_2 potentially placing it in a signalling cascade induced by ROS (http://www.arabidopsis.org/servlets/). The Genevestigator profile showed that the gene is expressed throughout plant growth (Figure 3.27). SALK_129085 T-DNA insertion and primer positions are shown in Figure 3.28. Genotypic analysis identified 9 plants as wildtype and 1 homozygous line (Figure 3.29). The homozygous line exhibited no male sterile phenotype. This could have been due to the position of T-DNA insertion in the 3' end of the gene. In future AT3G04120 could be analysed with a T-DNA insertion in the exon region, or by using alternative mutagenesis options.
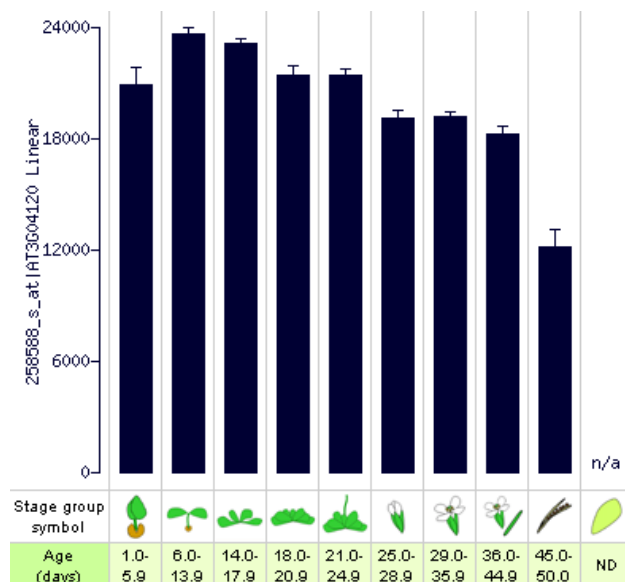
Figure 3.27 Genevestigator expression profile indicates that expression of AT3G04120 occurs throughout development (https://www.genevestigator.com).
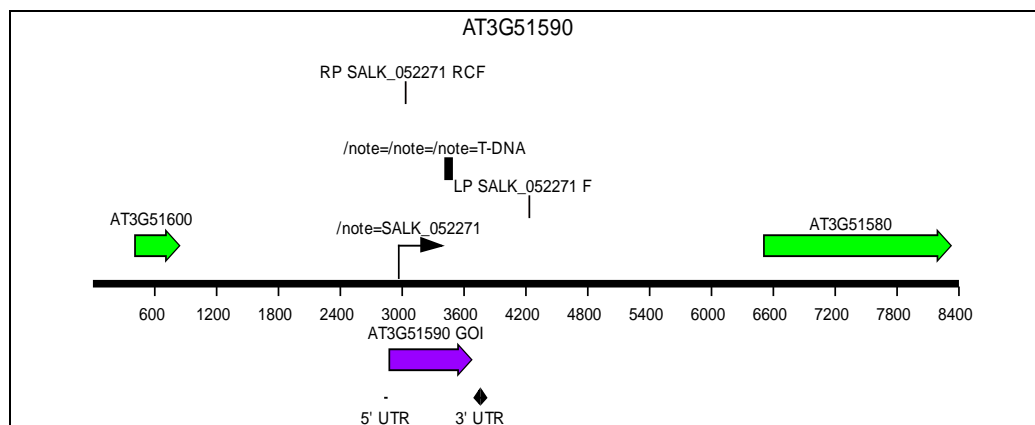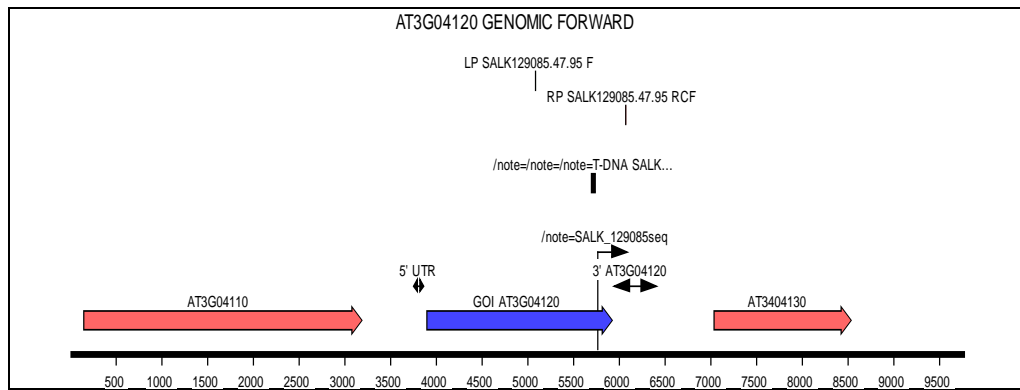
Figure 3.28 T-DNA localisation map of SALK_129085 (AT3G04120). Blue bold arrow represents the direction of the coding sequence for the gene of interest (AT3G04120), the upstream gene (AT3G04110) and downstream gene (AT3G04130) are represented by red arrows pointing in the direction of the coding sequence. Black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis.

Figure 3.29 Genotyping analysis of SALK_129085 (AT3G04120) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Genotyping PCR identified 9 as segregating wildtype (numbered 3-11) and 1 homozygote line with 800bp T-DNA insert (numbered 12). Wildtype Columbia ecotype was used as positive control (numbered1-2). M=Molecular marker.

3.4.9 T -DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT2G31210 (BASIC HELIX-LOOP-HELIX (BHLH) FAMILY PROTEIN): SALK_022241

AT2G31210 encodes for a putative bHLH transcription factor, it contains an InterPro domain Basic helix-loop-helix dimerisation region bHLH; (InterPro: IPR001092) and contains InterPro domain Helix-loop-helix DNA-binding. Genevestigator expression profile shows that the gene is expressed throughout plant growth but is strongly expressed in the floral organs (Figure 3.30).
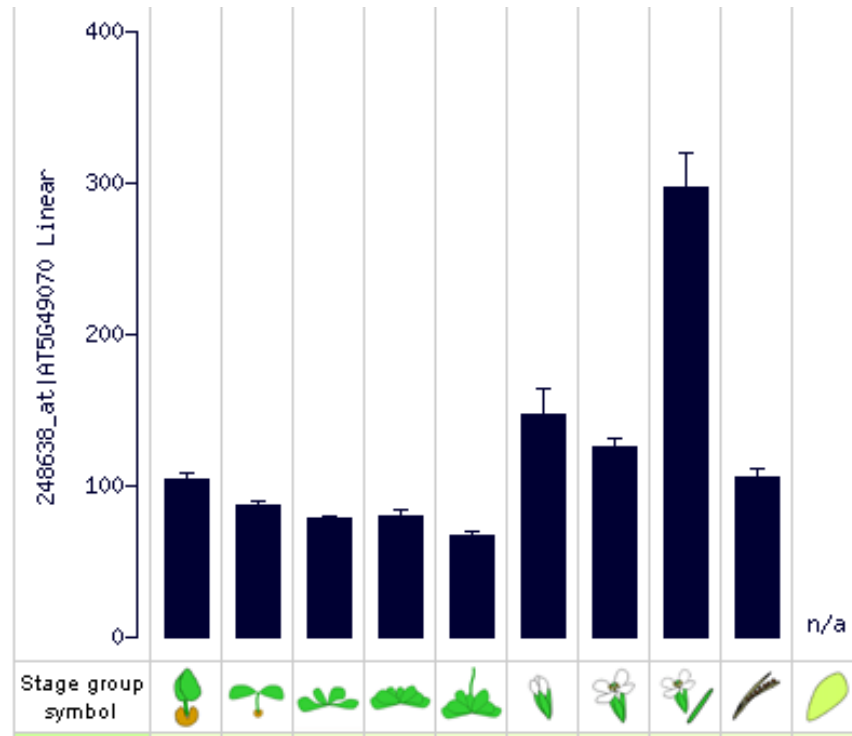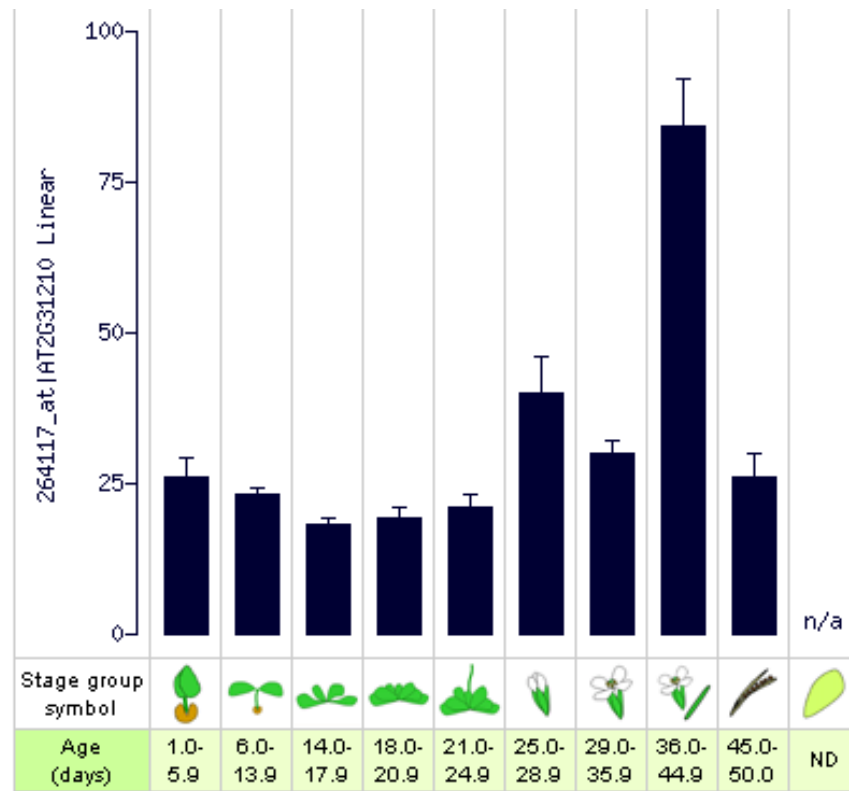

Figure 3.30 Genevestigator expression profile indicates that maximal expression of AT2G31210 during plant development is in floral organs (https://www.genevestigator.com).

SALK_022241 T-DNA insertion and primer positions are shown in Figure 3.31. Genotypic analysis showed that all SALK_022241 plants analysed were identified as heterozygous (Figure 3.32). The heterozygous lines lacked a male sterility phenotype; future analysis of potential phenotypes in homozygous lines from the segregating population of this line would be beneficial.
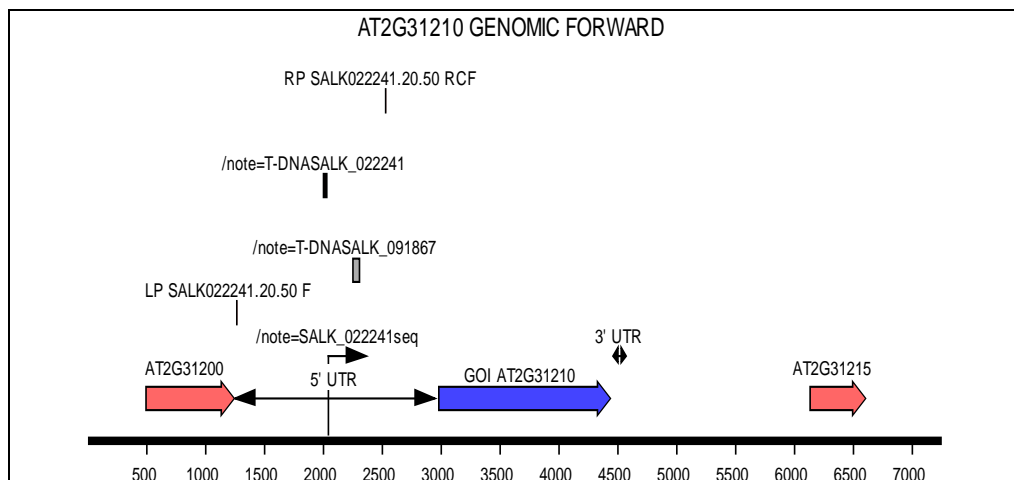
Figure 3.31 T-DNA localisation map of SALK_022241 (AT2G31210). Blue arrow represent the direction of the coding sequence for the gene of interest (AT2G31210), the upstream gene (AT2G31200) and downstream gene (AT2G31215) are represented by red arrows pointing in the direction of the coding sequence . Black arrowhead shows the approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis.
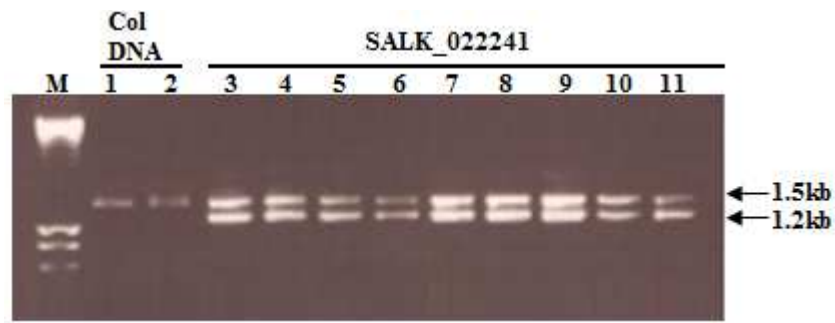
Figure 3.31 Genotyping analysis of SALK_022241 (AT2G31210) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Genotyping PCR identified all 9 SALK_022241 KO lines as heterozygous (numbered 3-11). Wildtype Columbia ecotype was used as positive control (numbered 1-2). M= Molecular marker.

3.4.10 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT1G06170 (BASIC HELIX-LOOP-HELIX): SALK_090231

AT1G06170 encodes for a putative basic helix-loop-helix (bHLH) DNA-binding superfamily protein, which functions in DNA binding, as a sequence-specific DNA binding transcription factor activity. The Genevestigator profile showed that the gene is expressed throughout plant growth but expression is highest in floral organs (Figure 3.32). SALK_090231 T-DNA insertion and primer positions are shown in Figure 3.33. Genotypic analysis identified all five lines segregating as wildtype (Figure 3.34); no further work was carried out on this line.
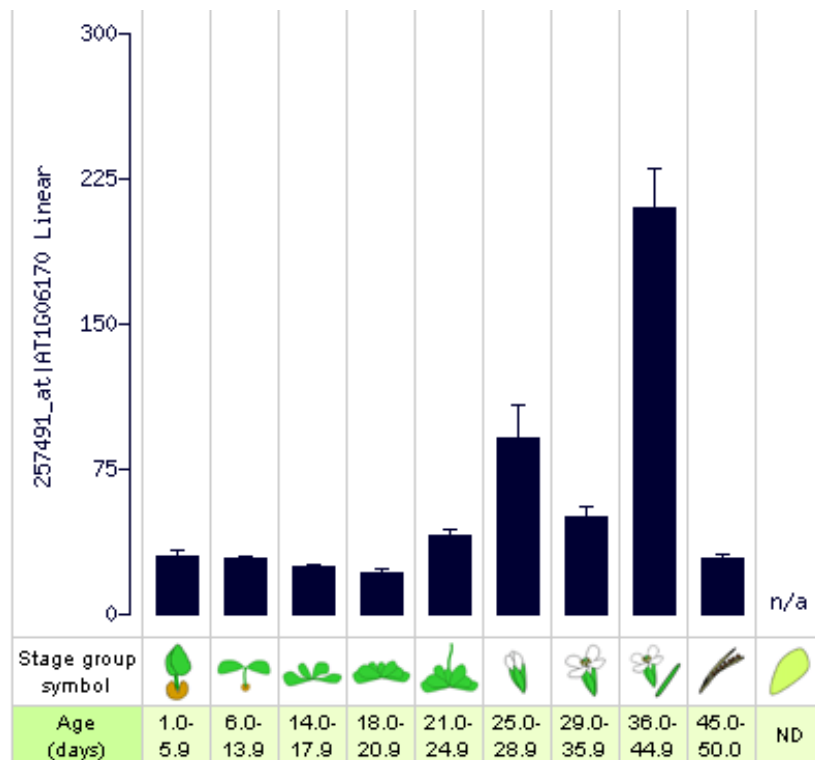
Figure 3.32 Genevestigator expression profile indicates that maximal expression of AT1G06170 during plant development is in floral organs (https://www.genevestigator.com).
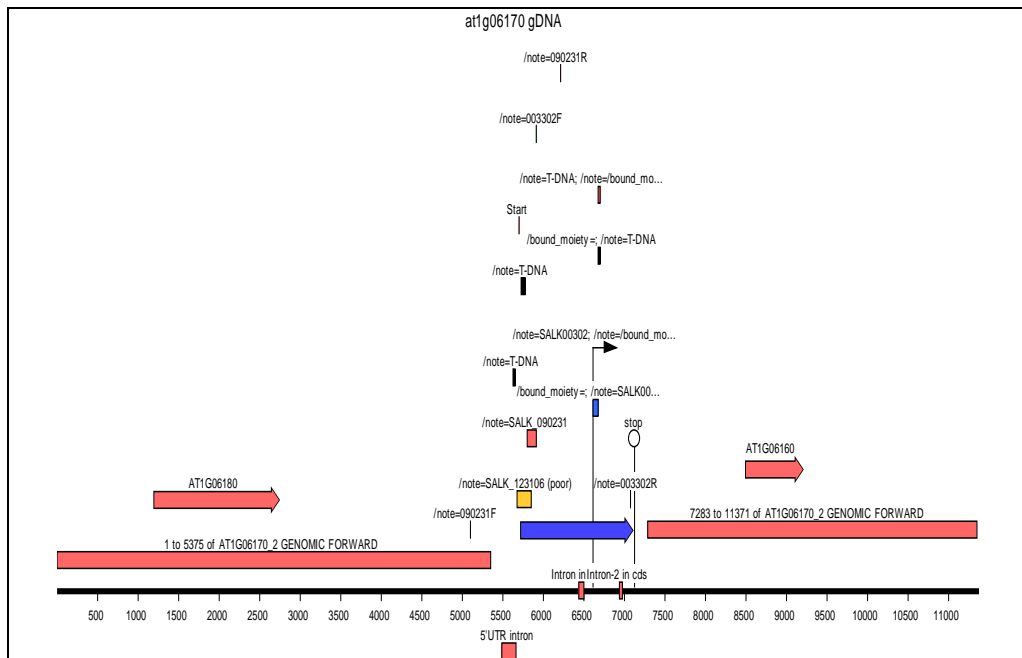
Figure 3.33 T-DNA localisation map of SALK_090231 (AT1G06170). Blue arrow represents the direction of the coding sequence for the gene of interest AT1G06170. Black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis. Black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis.

Figure 3.34 Genotyping analysis of SALK_090231 (AT1G06170) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Genotyping PCR analysis identified all five SALK_090231 AT1G06170 KO lines segregating as wildtypes (Numbered 2-6). Wildtype Columbia ecotype was used as a positive control (numbered 1). M= Molecular marker.

3.4.11 T-DNA INSERTIONAL MAP AND GENOTYPIC ANALYSIS OF AT3G12760 CALCIUM ION BINDING (MBK21): SALK_143248

AT3G12760 is similar to calcium ion binding [Arabidopsis thaliana] (TAIR: AT1G15860.1 and SM10 [*Nicotiana tabacum*] (GB: ABI49160.1) proteins and contains an InterPro Ubiquitin-associated domain. The Genevestigator profile shows that the gene is expressed throughout plant growth (Figure 3.35); SALK_143248 T-DNA insertion and primer positions are shown in Figure 3.36. Genotypic analysis identified all 12 SALK_143248 plants segregating as wild type (Figure 3.37) and consequently no more work was carried out on this SALK line.
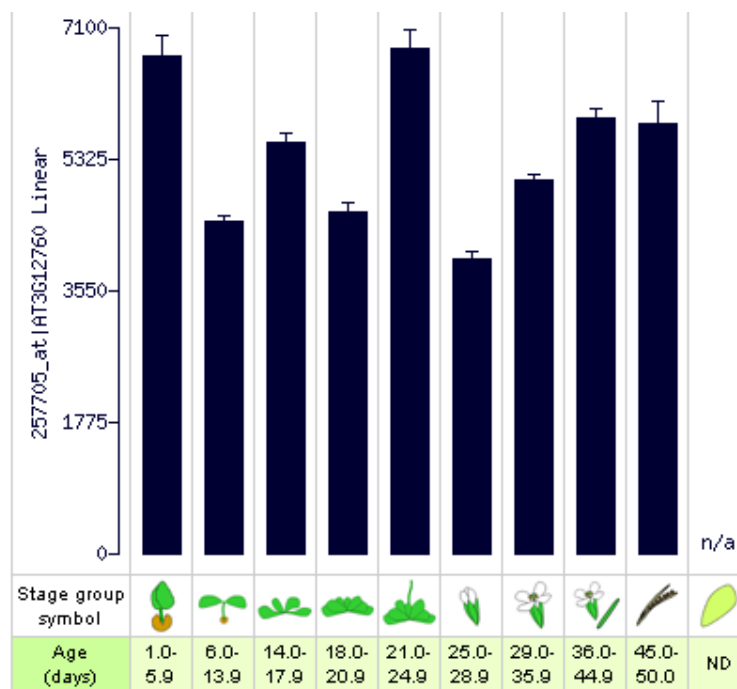
**Figure 3.35: Genevestigator expression profile** indicates that expression of At3g12760 is throughout plant development is in floral organs (https://www.genevestigator.com).
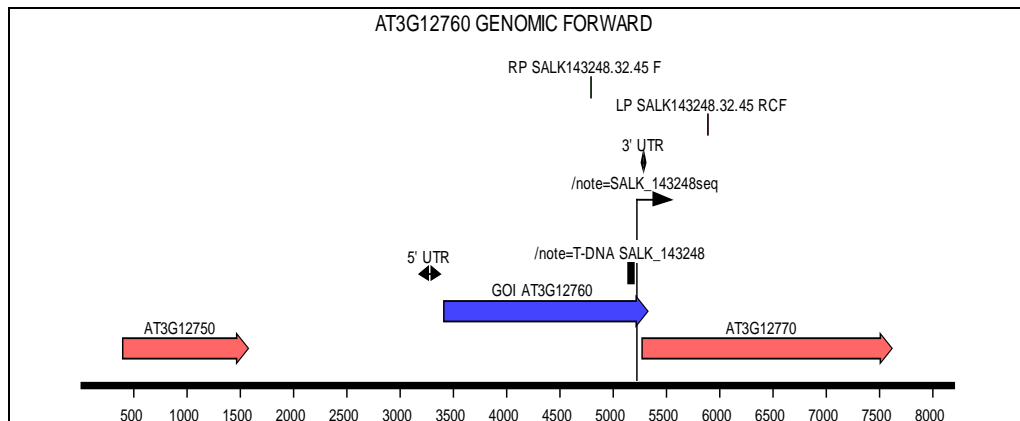
Figure 3.36: T-DNA localisation map of SALK_143248 (AT3G12760). Blue arrow represents the direction of the open reading frame for AT5G17050, the upstream gene (AT5G17060) and downstream gene (AT5G17080) are represented by red arrows pointing in the direction of the coding sequence. Black arrowhead shows approximate T-DNA insertion point and direction, black vertical thin lines show Left primer (LP) and Right Primer (RP) used for genotypic analysis.
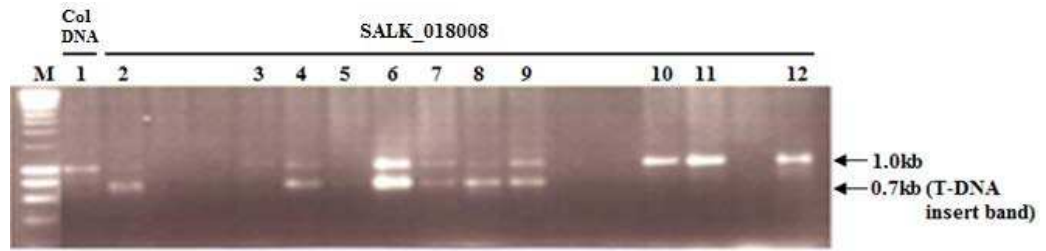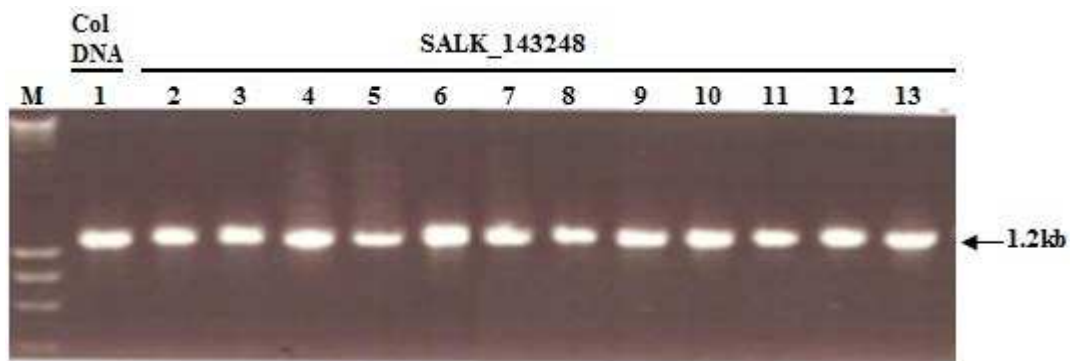

Figure 3.37 Genotyping analysis of SALK_143248 (AT3G12760) (Two gene specific and the LBb1.3 primers were used see Table 3.3). The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker**.** Genotyping PCR identified all 12 SALK_143248 lines segregating as wildtype (numbered 2-13). Wildtype Columbia ecotype was used as positive control (numbered 1). M= Molecular marker.

3.5 SUMMARY OF SALK INSERTIONAL MUTANT ANALYSIS

Eleven insertional mutants of putative targets of AMS were analysed by genotyping and phenotyping, this data is summarised in Table 3.5 and table 3.6. However except for At3g13220 (ABC transporter protein) homozygous lines the rest of the SALK KO lines analysed had normal growth characteristics compared to the wild type Columbia, regardless of some of them being homozygous lines.

Table 3.5 Summary of analysis of insertional mutants of possible regulatory targets of AMS.

Gene	AGI Code	Insertional mutant line	Location of T-DNA insert	Homozygous lines/ Total plants analysed	Expression of gene in mutant	Phenotype
ABC transporter protein (WBC27)	AT3G13220	SALK_062317	Exon	3/15	No	Male sterility in homozygous line
Lipid Transfer Protein (LTP12)	AT3G51590	SALK_052271	5'UTR	5/5	Yes	Fully fertile
3-Ketoacyl-Coa Synthase 21	AT5G49070	SALK_089611	Exon	0/8	N/A	Fully fertile
Protease inhibitor LTP	AT1G66850	SALK_018008	5'UTR	0/9	N/A	Fully fertile
(CYTOCHROME P450)	AT3G28740	SAIL_743	Exon	0/14	N/A	Fully fertile
ACA1, CA^{2+}ATPase	AT1G27770	SALK_107728	3'UTR	11/14	N/A	Fully fertile

Table 3.6 Summary of analysis of insertional mutants of possible regulatory targets of AMS

Gene	AGI Code	Insertional mutant line	Location of T-DNA insert	Homozygous lines/ Total plants analysed	Expression of gene in mutant	Phenotype
Anthocyanidin 3-0-glucosyl transferase; UDP-GLUCOSYL TRANSFERASE 78D2,	AT5G 17050	SALK_049 338	Exon	0/10	N/A	Fully fertile
Cytosolic GADPH (C subunit)	AT3G 04120	SALK_129 085	Exon	1/10	N/A	Fully fertile
bHLH protein	AT2G 31210	SALK_022 241	5'UTR	0/9	N/A	Fully fertile
bHLH protein	AT1G 06170	SALK_090 231	Exon	0/5	N/A	Fully fertile
MBK21	AT3G 12760	SALK_143 248	3'UTR	0/12	N/A	Fully fertile

3.6 DISCUSSION AND CONCLUSIONS

The work described here was specifically focused on characterization of genes that have been proposed to be regulated by AMS based upon previous microarray data analysis (Wilson and Zhang groups UoN and SJTU). During these studies, morphological observations of male fertility on SALK mutants and wildtype anthers coupled with molecular biology analyses showed how the effect of T-DNA insertion can alter gene expression levels. Out of 11 SALK insertional lines analysed only AT3G13220 (ABC transporter protein) mutant exhibited complete

silencing in the homozygous line. The remaining SALK insertion mutant lines had no clear phenotype and in particular no effect on male fertility despite some of them being homozygous. For example homozygous insertions in SALK_052271 still showed expression of At3g51590, although this was reduced compared to wild type (Figure 3.11, line 4). None of the 5 homozygous KO lines had a male sterility, or vegetative phenotype when compared to the wild type (Col). This maybe a consequence of the low level of expression still remaining in the homozygous mutant, or may be a reflection of it functioning as a member of a multigene family that exhibit redundancy or highly specialized functions. Other related family members that may be acting redundantly with this gene include AT2G38540/LTP1, AT2G38530/LTP2, AT5G59320/LTP3, AT5G59310/LTP4, AT3G51600/LTP5, AT3G08770/LTP6, AT2G15050/LTP7, AT2G18370/LTP8, AT2G15325/LTP9, AT5G01870/LTP10, AT4G33355/LTP11, AT3G51590/LTP12, AT5G44265/LTP13 and AT5G62065/LTP14 (http://www.arabidopsis.org).

In addition other homozygous insertions that did not show any male sterility or vegetative phenotype when compared to the wild type (Col) are SALK_107720 (AT1G27770) (Figure 3.23) and SALK_129085 (AT3G04120) (Figure 3.29). This result may be a reflection of them functioning as members of multigene families that exhibit redundancy or highly specialized functions. Related family members, which may be acting redundantly with AT1G27770 and possibly directly connected with AT1G27770 on the network include AT5G66210, AT4G12720, AT4G12720, AT2G47060, AT4G02200, AT5G47910, AT1G22510 AND AT5G46470 (http://atted.jp/data/locus/At1g27770.shtml). Other related

family members which may be acting redundantly with AT3G04120 and possibly directly connected with AT3G04120 on the network include AT2G39730, AT3G60750, AT4G13930, AT4G39330 and ATCG00490 (Lozano-Juste *et al.*, 2011). Further studies on AT3G13220 have been performed by Xu *et al.* (2010) who determined that AT3G13220 (*WBC27*) is a direct target of ABORTED MICROSPORES (AMS), which directly binds the promoter regions of AT3G13220 and also 12 other genes (Figure 3.38). This was performed through ChIP-PCR followed by an electrophoretic mobility shift assay (EMSA) with AMS and the promoter region (−213 to −37) of the ABC transporter *WBC27* (At3g13220); these data confirmed that AMS directly regulates the expression of the ABC transporter *WBC27* (At3g13220). Since the ABC (AT3g13220) is essential for pollen wall development and is directly regulated by AMS as reported by Xu *et al.*, 2010, it also supports a role for the AMS pathway in the transport of pollen wall materials (Figure 3.39). In addition to studies performed by Xu *et al.*, (2010), Quilichini *et al.*, (2010) further suggested that the ABC transporter ABCG26 (At3g13220) has a function in pollen exine development and the later degeneration of pollen.

Further studies by Choi *et al.*, (2011) and Dou *et al.*, (2011) have demonstrated that tapetal expression of the ABC transporter ABCG26 is essential for normal exine formation in microspores, and therefore for normal viable pollen production. In these studies, the development of *abcg26* mutant pollen was observed closely, and severely defective exine structure was identified as the possible cause of the later degeneration of most *abcg26* pollen. In line with this explanation, Choi *et al.*, (2011) found aberrant internal structures in tapetum cells

at the early bicellular stage of the *abcg26* mutant, which most likely originate from failure to deliver precursors of the exine structure to the outside of the cells. Taken together, the results of these concurrent independent studies strongly support the suggestion that ABCG26 is involved in the transportation of pollen-wall precursors from the tapetum to the anther locules and that sporopollenin precursors are the most likely substrates of ABCG26/WBC27 (Choi *et al*., 2011) and that AMS controls exine formation at least partly via regulation of *WBC27* expression (Xu *et al*., 2010),
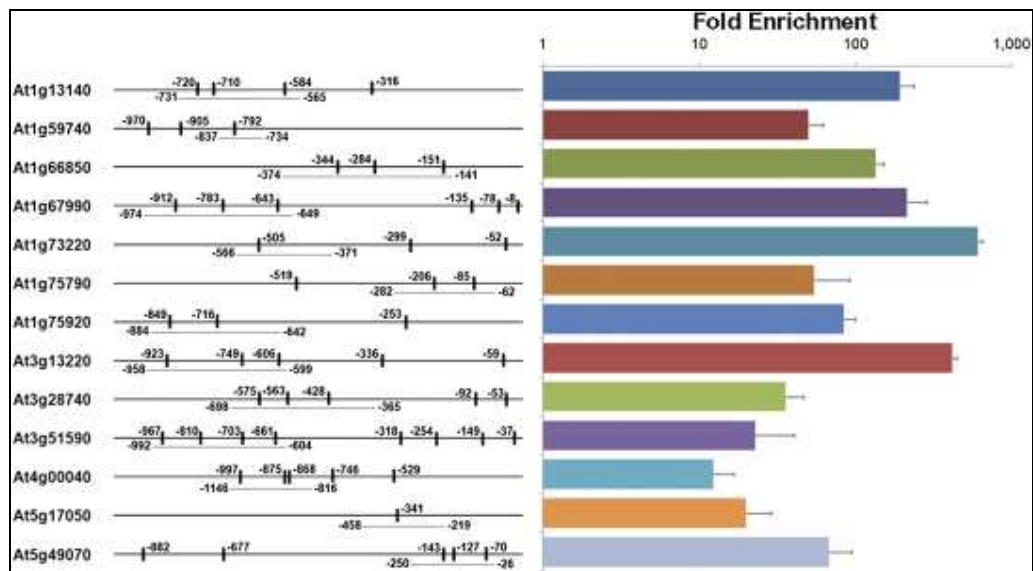
Figure 3.38: qChIP-PCR analysis of the enrichment of AMS regulatory targets and the predicted E-Boxes in their promoter regions. Fold enrichment calculations from qPCR assays in three independent ChIP experiments. The predicted E-boxes are indicated as vertical black lines in the promoter regions of the AMS targets, and the PCR amplicons used for ChIP-qPCR containing the E-boxes are underlined. Fold enrichment data were analysed to calculate the fold change between the ChIP (anti-AMS immune serum) and no-antibody control, and all of

the targets tested were present at higher amounts in the sample with anti-AMS serum than in the no-AMS antibody preimmune precipitation (Xu *et al*., 2010).
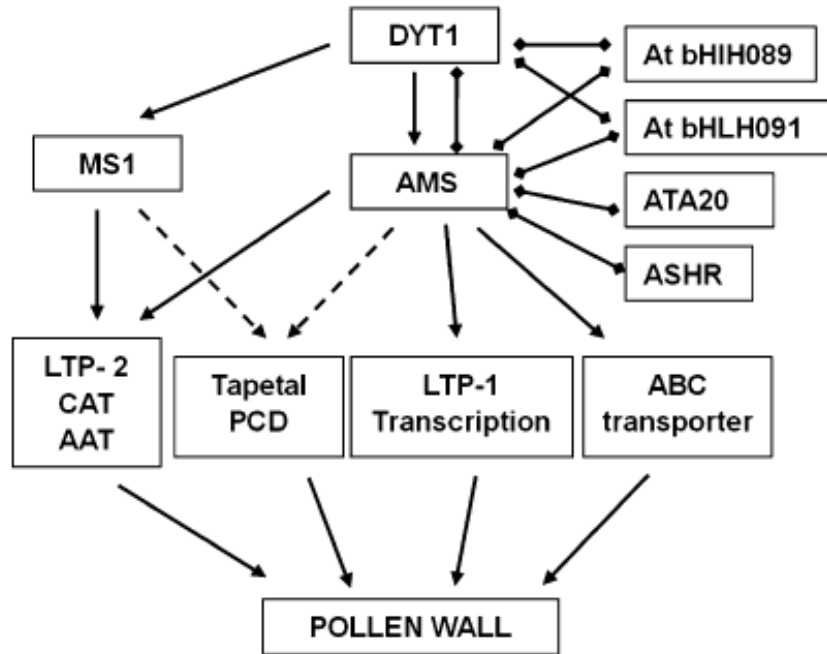
Figure 3.39 Model for the Role of AMS during Anther and Pollen Development. AMS regulates a number of direct and downstream regulatory targets related to tapetal PCD and pollen wall formation and interacts with a number of proteins associated with anther and pollen development. This model provides an indication of the regulatory network for anther development. LTP, lipid transfer protein; CAT, carbohydrate transport and metabolism; AAT, amino acid transport and metabolism. Arrows indicate regulation; lines with diamond ends indicate interaction (Xu *et al*., 2010).

The finding of a large number of candidate anther regulatory genes, and other genes with unknown functions, indicates that much remains to be discovered concerning this important floral organ. More work is required to get a fuller

understanding of the functional importance of the AMS regulatory pathway in *Arabidopsis* anther and pollen development. These results will provide insight into the regulatory role of the AMS network during anther development.

CHAPTER 4: FAST NEUTRON SEED

MUTAGENESIS TO IDENTIFY FACTORS

REGULATING MS1 EXPRESSION

4. FAST NEUTRON SEED MUTAGENESIS TO IDENTIFY FACTORS REGULATING MS1 EXPRESSION

Fast-neutron mutagenesis has long been used for forward genetic studies in plants (Li and Zhang, 2002). Forward Genetics involves the use of mutagenesis to identify genes of interest that are associated with a given process that are defined by a particular phenotypic change in a mutant. This is in contrast to reverse genetics where gene sequences are selectively manipulated and associated phenotypic changes investigated. Forward genetics has typically entailed the generation of random mutations in an organism, and then through a series of breeding of subsequent generations to isolate individuals with an aberrant phenotype associated with mutations at a single locus. Today there are many different means of generating aberrations and linking them to the observed phenotypes. Fast-neutron radiation is a highly efficient mutagen that produces deletion mutations that are easily detectable by PCR; this method is therefore extremely powerful to enable forward and reverse genetic screens in most plant species. In this study, focus was paid to identifying factors that regulate expression of the *MS1* gene. The approach that was used was to identify male sterile mutants in which the expression of *MS1* was altered. This was visualised using the *MS1Prom:MS1-GFP* fusion protein to analyse the expression of MS1 (Yang *et al*., 2007). During this work novel male sterile mutants were identified in which the MS1-GFP fluorescent translational fusion protein showed changes in stage specific expression. Backcrossing of these male sterile mutants with the parental plants (carrying the MS1Prom:MS1-GFP fusion construct) and the *ms1* mutant was performed to confirm the mutations, remove additional mutations in

the lines and check for allelism to the *ms1* mutation. Crosses were also conducted to enable mapping of the mutations.

4.1 METHODS

4.1.1 GENERATION OF MUTAGENISED SEED

Seeds of *Arabidopsis* Landsberg *erecta* ecotype carrying the MS1Prom:MS1-GFP (Yang *et al*., 2007) functional fusion was mutagenised at the HAS KFKI Atomic Energy Research Institute using fast neutrons to identify novel male sterile mutants and genes that are impacting on the expression of MS1. Irradiation geometry at BIF: 2Y/Cd (rotated) was selected to increase dose homogeneity. The averaged water kerma dose rate for 2Y/Cd geometry was 437mGy/min (+/-3.0%) and all the seeds were exposed together to a nominal dose of 60 Gy 137 minute net beam time. The inhomogeneity of dose distribution within one sample was less than 1%. The overall dose uncertainty for the seeds was less than 5% and the total surface gamma activity of the seeds after irradiation was 100 BGND [1BGND (background) ~ 90 nGy/h]. The total surface gamma activity of the seeds at the time of collection was <2 BGND. Pools (124) of mutagenised M1 seeds were grown for 6 weeks and each pool consisted of 25 plants; 300 seeds of each pool from the M1 population were subsequently grown and from 6 weeks onwards this resultant M2 generation was phenotypically and microscopically analysed for any abnormities in the morphology and functionality of the floral organs (section 4.2.2).

4.1.2   PHENOTYPIC ANALYSIS OF M2 MUTANT POPULATION

The M2 generation was analysed for any possible phenotype associated with male sterility (lack of silique elongation and seed set, or abnormal floral structure) compared to the wild type L*er* ecotype. Plants that had short siliques and did not contain seeds, or had reduced numbers of seeds were further analysed for pollen viability by Alexander staining (Alexander, 1969) to establish whether viable pollen was present. Any changes observed in the M2 plants were compared to the wild type L*er* ecotype and differences were recorded (section 4.3.1).

4.1.3   PHENOTYPIC AND MICROSCOPIC ANALYSIS OF F2 MUTANT GENERATION

In order to rescue the M2 male sterile plants and also to test if the sterility was a result of dysfunctional pollen, the M2 male sterile lines previously analysed for any possible phenotype were back-crossed with the *MS1-GFP* parental line. The resulting F1 generation was allowed to self to form an F2 segregating population. To ensure that the lines were free from additional mutations and to 'clean up' the desired mutant line, the mutant plants were further backcrossed and then selfed, at least three times to the MS1-GFP parental line whilst selecting for the male sterile phenotype in the segregating generations. To identify any changes in the stage specific expression of MS1, microscopic observation of the fluorescent reporter gene GFP (detailed method in Chapter 2 Section 2.7) were conducted for all of the male sterile lines. Four male sterile mutants plants (2Y, 8A 12T and 17L) exhibited altered MS1-GFP expression (section 4.3.2).

4.1.4 COMPLEMENTATION TEST FOR ALLELISM

Four male sterile plants (2Y, 8A 12T and 17L) with altered MS1-GFP expression were identified and crossed to the *ms1* mutant to determine whether their mutation was a consequence of mutation of the *MS1* gene. Heterozygous plants of the mutant lines, which were fertile, were used to cross with the *ms1* homozygous mutant. Table 4.1 shows the predicted segregation expected based upon complementation, or novel mutations (section 4.3.3).

Table 4.1 Possible results expected from the cross between *ms1* mutant and heterozygous new male sterile MS1Prom:MS1-GFP lines.

Possible genotype in the mutagenised MS1Prom:MS1-GFP line	*ms1ttg* genotype	Complementation result	Comments
MS1MS1 *msxmsx*	*ms1ms1* *MSXMSX*	*ms1MS1* *MSXmsx*	All plants would be fertile as both copies of the *MS1* gene in the new male sterile MS1Prom:MS1-GFP line will be normal and the new mutation will be complemented by the wild type version in *ms1ms1*.
ms1MS1	*ms1ms1*	*ms1MS1* *ms1ms1*	50 % of plants will be male sterile and 50% will be normal as the mutation in the MS1Prom:MS1-GFP mutant line is in the *MS1* locus and there will be no complementation of the phenotype.

4.1.5   MAPPING THE NOVEL MUTATION

Map-based cloning also called positional cloning is an approach that identifies the underlying genetic cause of a mutant phenotype by looking for linkage to markers whose physical location in the genome is known (Jander *et al*., 2002). The major strength of this approach is the ability to tap into a nearly unlimited resource of natural and induced genetic variation without prior assumptions or knowledge of specific genes (Jander *et al*., 2002). Essentially, we are looking at all of the genes in the genome at the same time to find the ones that affect the male sterility phenotype. It is a process of discovery that makes it possible to find mutations anywhere in the genome, including intergenic regions and the 40% of Arabidopsis genes that do not resemble any gene with known or inferred function (The Arabidopsis Genome Initiative, 2000). We focused on male sterility as our mutant phenotype of interest to identify novel male sterile mutants and genes that may be impacting on the expression of MS1.

To generate an approximate map position for the new male sterile mutant (2Y), a segregating mapping population was generated from a cross between 2Y (in the *Ler* background) and the *Columbia* (*Col*) ecotype.   The F1 generation was allowed to self and then the subsequent F2 segregating generation was screened for the presence of male sterile mutants for first-pass mapping. Leaves were harvested and used for DNA analysis for the entire male sterile plants in this segregating population. The DNA samples were analysed using 16 PCR markers corresponding to each chromosomal arm (Figure 4.1) (Table 4.2) to define genetic linkage to one or more of the markers that contained the gene of interest.

The primers were initially tested on genomic DNA from *Arabidopsis* wild type *Ler* and *Col* ecotypes to confirm the product sizes they would amplify and the presence of polymorphisms. The male sterile lines were compared to *Col* and *Ler* and scored for their genotype at each marker to identify recombinant around the mutant locus. The data were then analysed for distortion of segregation associated with the mutant locus. Mapping-resolution was principally determined by the size of the mapping population and marker density; therefore 600 F2 plants were grown for male sterility screening and segregation analysis. From the 600 plants grown male sterile lines were generated and PCR analysis was performed on these lines with 16 molecular markers (Table 4.2) (section 4.3.4).
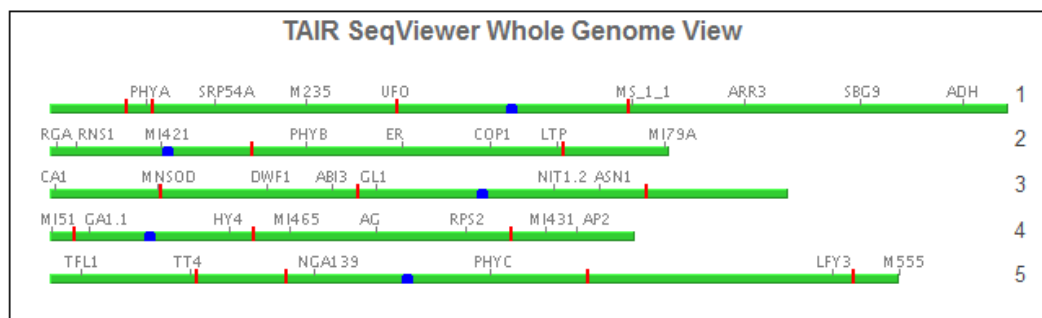
Figure 4.1 Map positions of PCR markers on the five Arabidopsis chromosome arms. Molecular markers (PCR primers; 16) are marked in red vertical lines. Green lines represent the 5 Arabidopsis chromosomes. Blue circles indicate the centromeric region.

Table 4.2 Molecular markers (PCR primers) used to map the gene of interest on its chromosomes (www.arabidopsis.org/).

PCR MARKER	GENE ID	PRIMER PAIRs (5'-3') FOWARD PRIMER (F) REVERSE PRIMER (R)	EXPECTED SIZE (BP) COL	EXPECTED SIZE (BP) LER	ANNEAL-ING TEMP. (^{0}C)
Chromosome 1	At1g07810	(F)= Gttcacggacaaagagcctgaaat (R)= aagcagtcaatattcaggaaggg	>300	<300	55
Chromosome 1	At1g09940	(F)= tcatgacgtgaagaagaagaaaa (R)= catatcgctgctactaattttaaacaa	>200	<200	55
Chromosome 1	At1g30930	(F)= tcaatgggatcgaaactggt (R)= actgaaaagcgagccaaaag	239	142	55
Chromosome 1	At1g49610	(F)= acattttctcaatccttactc (R)= gagagcttctttatttgtgat	<150	>150	55
Chromosome 2	At2g14890	(F)= gaaactcaatgaaatccactt (R)= tgaacttgttgtgagctttga	<200	>200	55
Chromosome 2	At2g39010	(F)= tcgtctactgcactgccg (R)= gaggacatgtataggagcctcg	150/151	130/135	55
Chromosome 3	At3g11220	(F)= ggattagatggggatttctgg (R)= ttgctcgtatcaacacacag	193	174	55
Chromosome 3	At3g26605	(F)= ccccgagttgaggtatt (R)= gaagaaattcctaaagcattc	<200	>200	53
Chromosome 3	At3g50820	(F)= gttcattaaacttgcgtgtgt (R)= tacggtcaagattgagtgattc	190	215	55
Chromosome 4	At4g01710	(F)= ggttaaaaattagggttacga (R)= agatttacgtggaagcaat	>150,	<150	55
Chromosome 4	At4g29860	(F)=gcccagaggaaagaagagcaaact (R)=tgggaattcatgagagaatatgtggg	492	404	55
Chromosome 4	At4g10360	(F)= gccaaacccaaaattgtaaaac (R)= tagagggaacaatcggatgc	268	188	55
Chromosome 5	At5g14320	(F)= Ggcctaagaaccaaatcaaaacaa (R)= cgtgatgaagtctccaagtacatg	225	271	55
Chromosome 5	At5g22545	(F)= Tagtgaaacctttctcagat (R)= Ttatgttttcttcaatcagtt	100	>130	50
Chromosome 5	At5g42600	(F)= Cagacgtatcaaatgacaaatg (R)= Gactactgctcaaactattcgg	165	145	55
Chromosome 5	At5g63640	(F)= Gagcatttcacagagacg (R)= atcactgttgtttaccatta	200	>200	50

4.2 RESULTS

4.2.1 PHENOTYPIC AND MICROSCOPIC ANALYSIS OF M2 MUTAGENISED SEED POPULATION

Phenotypic and microscopic analysis of plants from the M1 generation generally showed normal growth characteristics compared to wild type L*er* plants. The M1 mutant generation generally produced long siliques with full seed set, indicating that the plants from this M1 generation were fertile (Figure 4.2). Visual observation of M2 plants from different pools identified lines with male sterile phenotypes with short siliques (Figure 4.3: Table 4.3A - 4.3B). Microscopic observation of open flowers from these male sterile plants revealed that the stigma appeared normal, but no pollen was visible on the stigmatic surface (Figure 4.4B) compared to wild type *Ler* stigmatic surface that had visible pollen grains on its surface (Figure 4.4A). The anthers of male sterile plants were empty of pollen (Figure 4.4D), compared to wild type *Ler* which had released pollen onto the stigma, and had anthers that contained viable pollen that stained pink with Alexander stain (Alexander, 1969) (Figure 4.4C).

Figure 4.2 Fast-Neutron Mutagenised M1 generation carrying MS1Prom:MS1-GFP functional fusion. Each pool consisted of 25 plants and 124 pools were harvested. The entire M1 plants showed normal growth characteristics when compared to wild type L*er* plants, they produced long siliques with full seed set an indication that the plants from this M1 generation were all fertile.

Figure 4.3 Silique phenotype of wildtype L*er* and an M2 mutant. (**A**) The wild type L*er* ecotype was fertile with long siliques and full seed set. (**B**) M2 Male sterile plants carrying MS1Prom:MS1-GFP.

Table 4.3A Analysis of M2 Fast Neutron Mutagenised population. Male sterile plants were identified from the various pools analysed, these exhibited short siliques compared to the wild type *Ler* ecotype. (--- = negative pollen viability).

POOL ID NUMBER	TOTAL NUMBER OF PLANTS GROWN	ID NUMBER FOR SINGLE MALE STERILE PLANTS REPRESENTED BY EACH POOL	PHENOTYPIC DESCRIPTION	POLLEN VIABILITY
1	300	1A	Short siliques	---
1	300	1B	Short siliques	---
1	300	1C	Short siliques	---
1	300	1D	Short siliques	---
2	300	2C	Short siliques	---
2	300	2D	Short siliques	---
2	300	2Y	Short siliques	---
3	300	3A	Short siliques	---
3	300	3B	Short siliques	---
5	300	5A	Short siliques	---
5	300	5B	Short siliques	---
6	300	6A	Short siliques	---
6	300	6B	Short siliques	---
7	300	7A	Short siliques	---
7	300	7B	Short siliques	---
8	300	8A	Short siliques	---
8	300	8B	Short siliques	---
8	300	8C	Short siliques	---
9	300	9A	Short siliques	---
9	300	9B	Short siliques	---
10	300	10A	Short siliques	---
11	300	11A	Short siliques	---
11	300	11B	Short siliques	---
12	300	12B	Short siliques	---
12	300	12T	Short siliques	---
13	300	13B	Short siliques	---
14	300	14A	Short siliques	---
14	300	14B	Short siliques	---
14	300	14C	Short siliques	---
15	300	15A	Short siliques	---
15	300	15B	Short siliques	---
15	300	15C	Short siliques	---
16	300	16A	Short siliques	---
16	300	16B	Short siliques	---
16	300	16C	Short siliques	---

Table 4.3B Analysis of M2 Fast Neutron Mutagenised population. Male sterile plants were identified from the various pools analysed, these exhibited short siliques compared to the wild type *Ler* ecotype. (--- = negative pollen viability).

POOL ID NUMBER	TOTAL NUMBER OF PLANTS GROWN	ID NUMBER FOR SINGLE MALE STERILE PLANTS REPRESENTED BY EACH POOL	PHENOTYPIC DESCRIPTION	POLLEN VIABILITY
17	300	17L	Short siliques	---
17	300	17C	Short siliques	---
18	300	18A	Short siliques	---
18	300	18B	Short siliques	---
19	300	19A	Short siliques	---
20	300	20A	Short siliques	---
23	300	23A	Short siliques	---
23	300	23B	Short siliques	---
24	300	24A	Short siliques	---
25	300	25A	Short siliques	---
25	300	25B	Short siliques	---
26	300	26A	Short siliques	---
27	300	27A	Short siliques	---
27	300	27B	Short siliques	---
27	300	27C	Short siliques	---
28	300	28D	Short siliques	---
30	300	30A	Short siliques	---
31	300	31A	Short siliques	---
32	300	32A	Short siliques	---
32	300	32B	Short siliques	---
33	300	33A	Short siliques	---
37	300	37A	Short siliques	---
38	300	38A	Short siliques	---
43	300	43A	Short siliques	---
45	300	45A	Short siliques	---
45	300	45B	Short siliques	---
45	300	45C	Short siliques	---
47	300	47A	Short siliques	---
50	300	50A	Short siliques	---
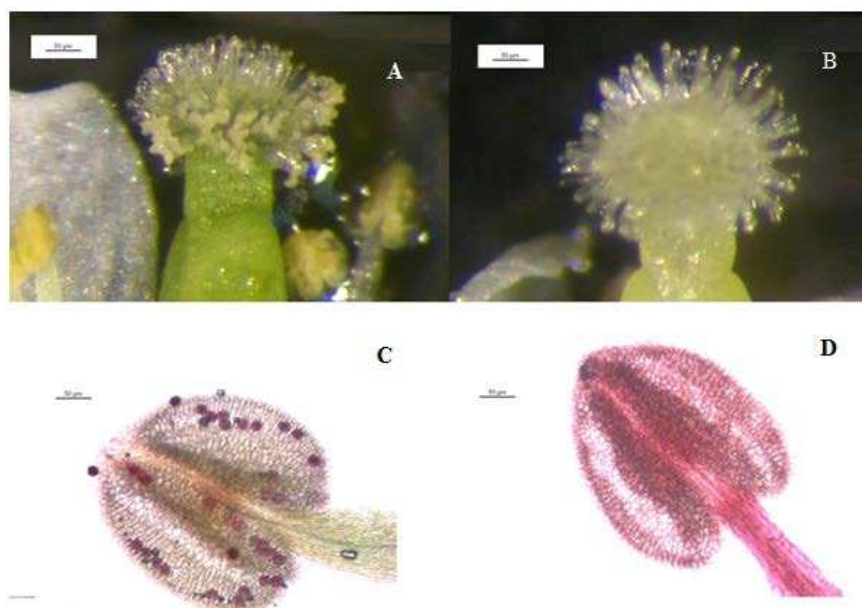
Figure 4.4 Phenotypic analysis of open flowers of Wild type *Ler* and M2 male sterile plants carrying MS1Prom:MS1-GFP functional fusion**. (A)** Wild type *Ler* flower with pollen grains on the stigmatic surface and on the dehisced anthers. (**B**) Pistil from an open flower of an M2 male sterile line carrying the *MS1Prom:MS1-GFP* functional fusion, no pollen is visible on the stigma. (**C**) Anther from wild type plant showing viable pollen after Alexander staining. (**D**) Anther from M2 male sterile plant after Alexander staining, but with no intact pollen detected. The figures are representative of the typical phenotype seen in the male sterile mutants. Bar=50 μm.

4.2.2 PHENOTYPIC AND MICROSCOPIC ANALYSIS OF F2 MUTAGENISED SEED POPULATION

The aim of the mutagenesis was to identify mutations, which were effecting the expression of MS1-GFP. Plants were analysed by fluorescent microscopy and

114

then detailed confocal microscopy for MS1-GFP expression. Out of 64 male sterile lines identified only 4 plants including 2Y, 8A, 12T and 17L (Figure 4.5) exhibited altered MS1-GFP expression (Figure 4.6 and 4.7). In the wild type, MS1-GFP is typically transiently expressed in the tapetum, from the callose release to the free microspore stage (Figure 4.6B). However in the mutants with altered MS1-GFP expression this was extended beyond microspore release and maintained during pollen mitosis 1 (Figure 4.6C-F and 4.7C-F). Mutant lines 2Y, 8A, 12T and 17L were subjected to allelism tests to determine whether they carried novel mutations or further allelic mutations of the *MS1* gene (Table 4.3).

Figure 4.5 Phenotype of F2 male sterile mutant lines. The male sterile lines all exhibited short siliques, increased flower number and altered MS1-GFP expression pattern compared to the wild type MS1-GFP parental line.
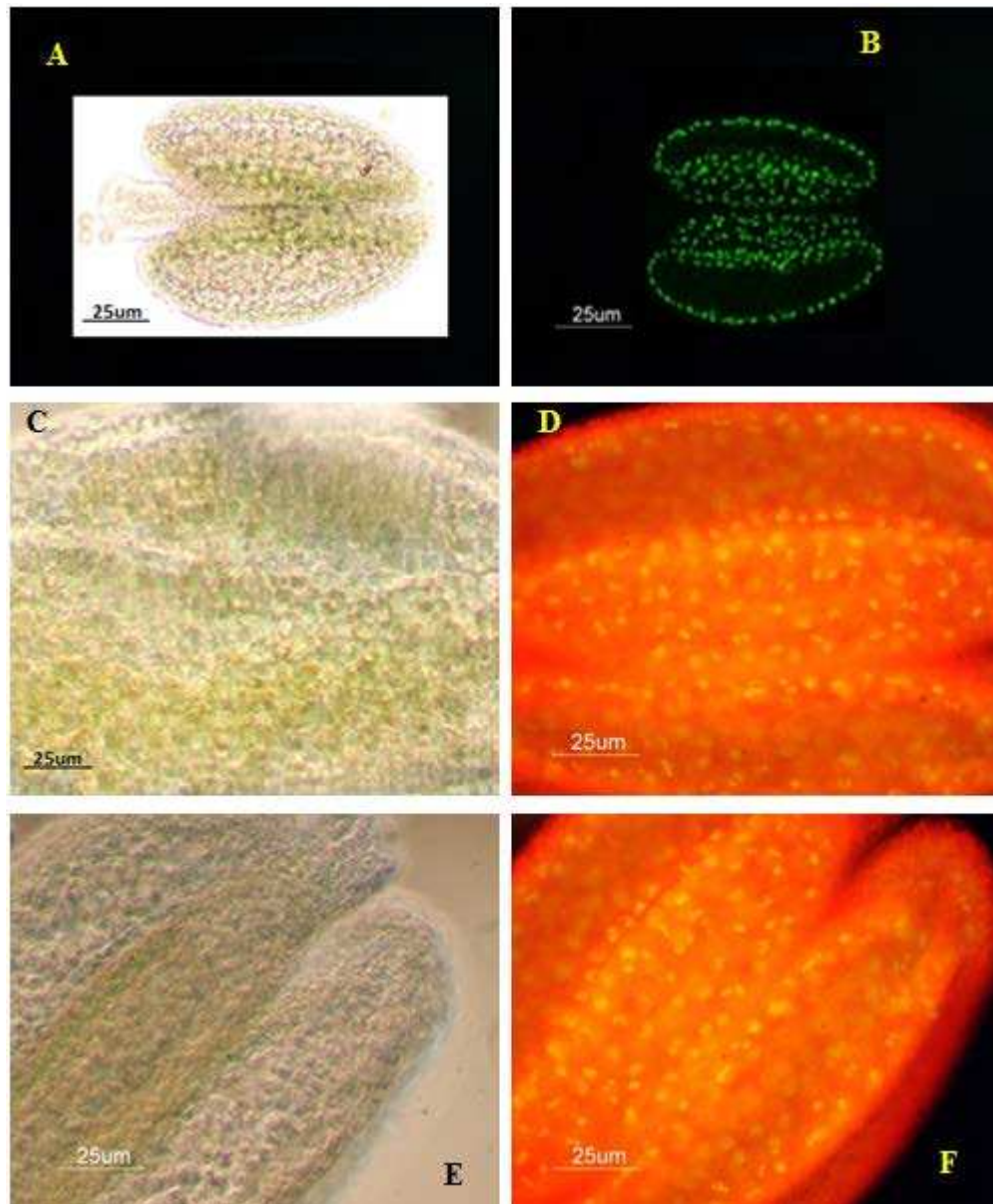
Figure 4.6 visualisation of MS1-GFP expression in *Arabidopsis* wild type Landsberg *erecta* ecotype and male sterile mutants (**A-B**: wild type) (**C-D**: 2Y mutant) (**E-F**: 8A mutant): (**A**) Representative wild type plant carrying the MS1Prom:MS1-GFP functional fusion protein in a bright field image of a wild type anther during tetraspore stage when microspores are released from the tetrads. (**B**) Functional MS1-GFP expression in the tapetum of wild type anthers. (**C**) Bright field image of 2Y Male sterile mutant anther. (**D**) 2Y mutant showing extended expression of MS1-GFP extended beyond microspore release. (**E**) Bright field image of the 8A male sterile mutant anther. (**F**) 8A mutant showing extended expression of MS1-GFP beyond microspore release.
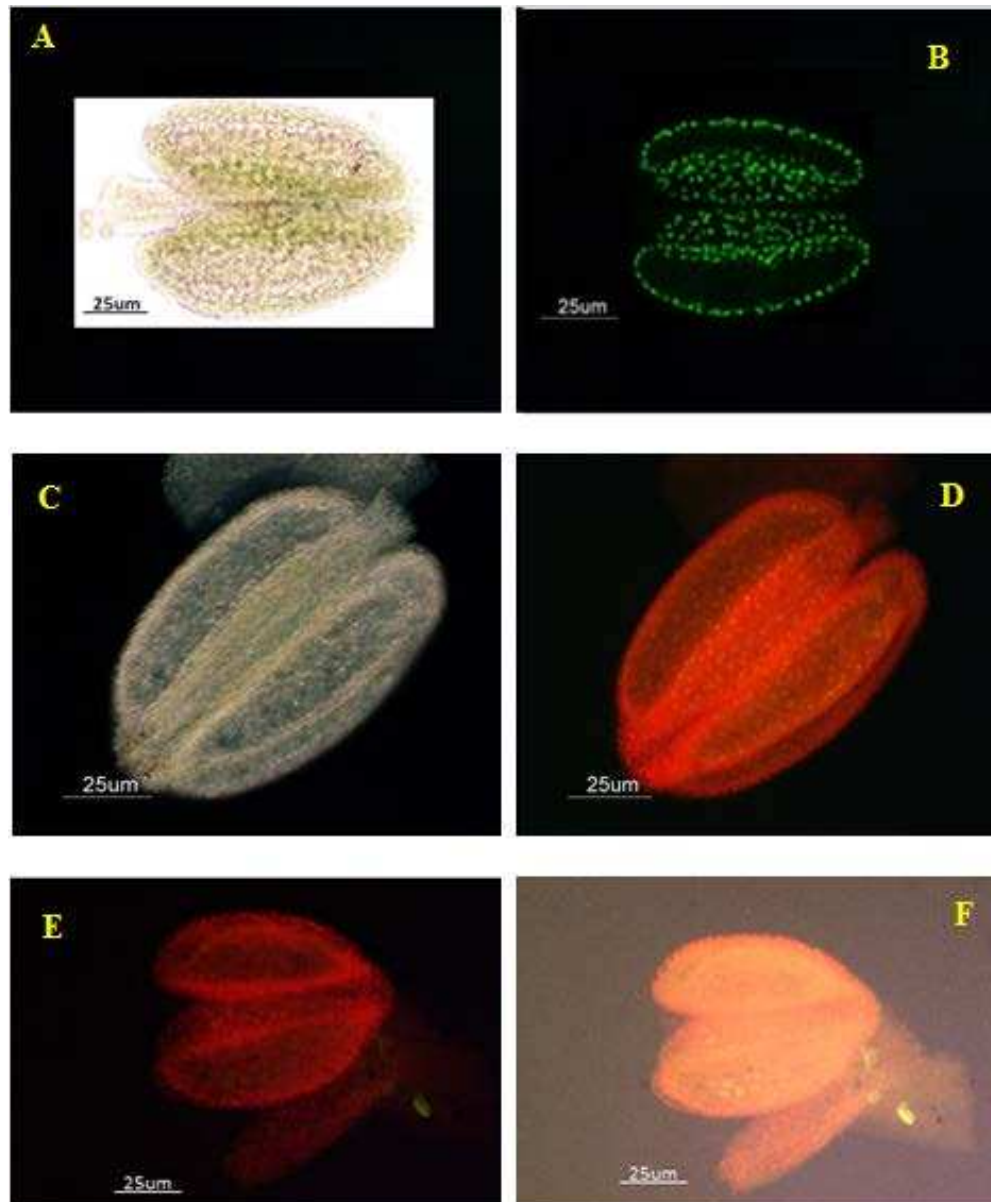
Figure 4.7 Visualisation of MS1-GFP expression in *Arabidopsis* wild type Landsberg *erecta* ecotype and male sterile mutants (**A-B**: wild type) (**C-D**: 12T mutant) (**E-F**: 17 mutant) (**A**) Representative bright field image of a wild type anther that is carrying the MS1Prom:MS1-GFP functional fusion protein during tetraspore stage when the microspores are released from the tetrads. (**B**) Functional MS1-GFP expression in the tapetum of wild type anthers. (**C**) Bright field image of anther from the 12T male sterile mutant. (**D**) 12T male sterile mutant showing extended expression of MS1-GFP beyond the microspore release stage. (**E**) 17L Male sterile mutant anther showing chlorophyll fluorescence. (**F**) 17L mutant male sterile mutant showing extended expression of MS1-GFP beyond microspore release.

4.2.3 ALLELISM TESTS

The four mutants that were showing altered expression of MS1-GFP protein were crossed to the *ms1* mutant to determine whether they were novel mutants, or allelic versions of the *ms1* mutation. Based upon this complementation analysis (Table 4.4) line 17L showed a clear 1:1 segregation for sterility indicating that no complementation of the *ms1* mutation had occurred and that the mutation in this line corresponded to the *MS1* locus. Lines 8A and 12T also showed some segregation of the male sterile phenotype; this was slightly lower than expected however suggests that the mutations in these lines are probably also in the *MS1* locus. Whereas crosses between line 2Y and the homozygous *ms1ttg* mutant did not show any male sterile progeny, suggesting that this line carries a novel mutation that effects expression of *MS1*, but is not within the *MS1* locus. Line 2Y was therefore used for molecular mapping analysis to identify the location of the novel mutation (Section 4.3.4). In addition to mapping analysis, further phenotypic observation of floral organs of 2Y were made, which confirmed the male sterility phenotype (short siliques with shrivelled anthers devoid of pollen) (Figure 4.8 and 4.9).

Table 4.4 Allelism tests for novel mutations and *ms1*. Crosses were conducted between heterozygous lines of the new mutants in the MS1-GFP background and homozygous *ms1ttg* plants. Chi-square analysis indicated a p-value of <0.05 for the four lines indicating that the observed deviation from the null hypothesis was significant.

MS1Prom:MS1-GFP line	Total number of plants analysed	Number of fertile plants	Number of male sterile plants	Probability of the chi-squared distribution	level of significance	Possible genetic outcome
17l	22	11	11	0.026564	P >0.05= statistically non-significant	MS1Prom:MS1-GFP line heterozygous for mutation in *MS1* gene and hence no gene complementation occurred. :- *MS1* allele
12T	22	13	9	0.011276	P >0.05= statistically non-significant	UNKNOWN, possible that genetic complementation has not occurred; mutation may be in *MS1* locus :- probable *MS1* allele
8AN	22	15	7	0.004701	P >0.05= statistically non-significant	UNKNOWN, possible that genetic complementation has not occurred; mutation may be in *MS1* locus :- probable *MS1* allele
2Y	22	22	0	0.0002	P <0.05= statistically significant	Mutation in MS1Prom:MS1-GFP line was complemented by *ms1ttg* mutant. :- novel mutation
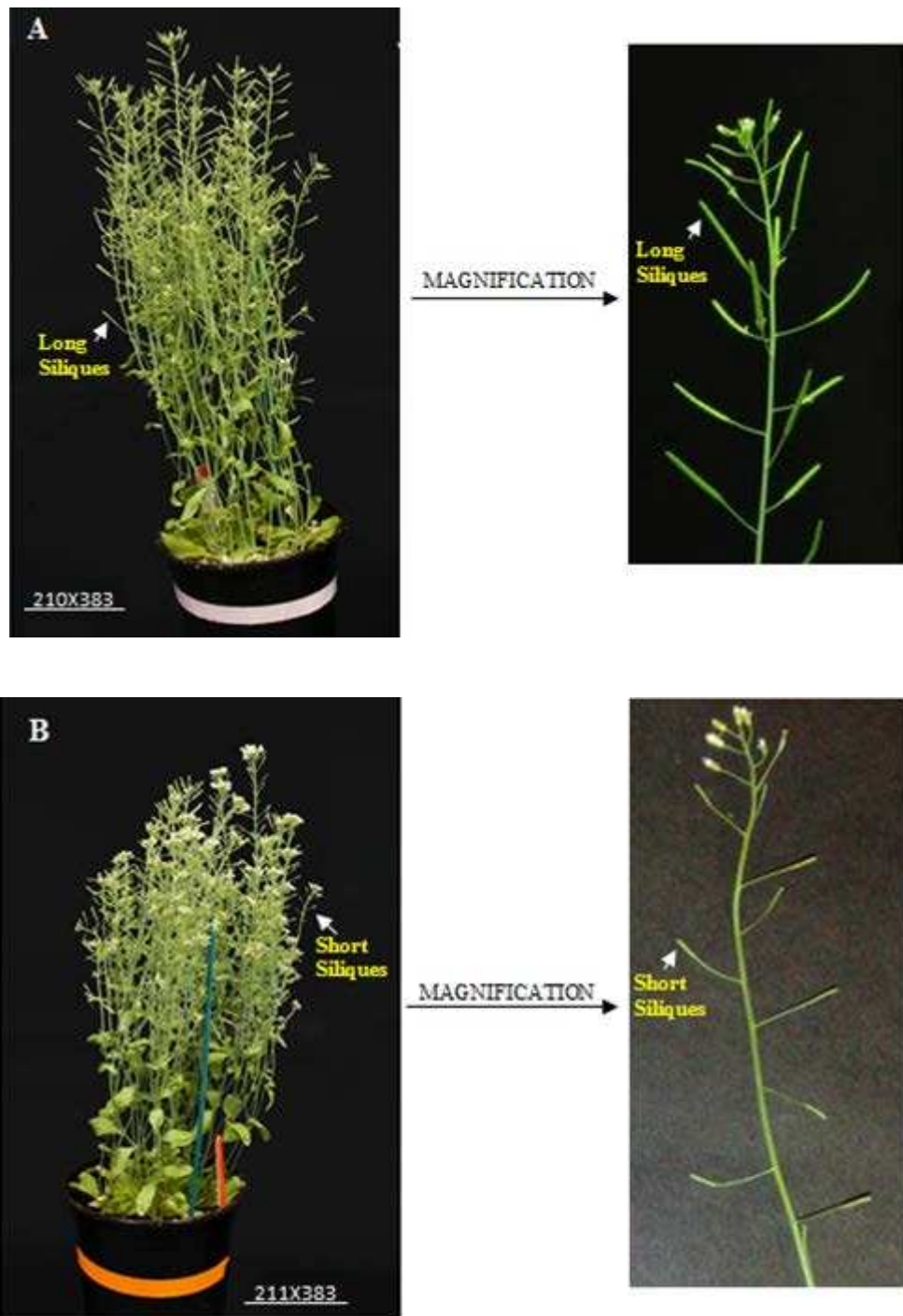
Figure 4.8 Phenotypic analysis of open flowers of (**A**) Wild type *Ler* and (**B**) F2 male sterile mutant 2Y carrying MS1Prom:MS1-GFP functional fusion. Compared to Wild *Ler* MS1-GFP parent plant in (**A**) 2Y in (**B**) displays short siliques.
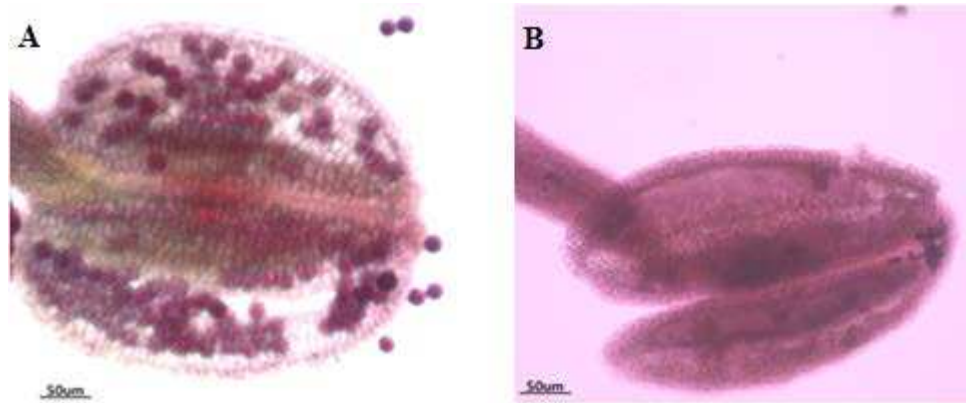
Figure 4.9 Analysis of pollen viability of (**A**) Wild type L*er* and (**B**) F2 male sterile mutant 2Y carrying MS1Prom:MS1-GFP functional fusion. Compared to wild type anthers in (**A**) which showed viable pollen Anthers of 2Y open flowers in (**B**) were devoid of pollen grains after Alexander (1969) staining.

4.2.4   MAPPING ANALYSIS FOR THE NOVEL MUTATION

The primers used as molecular markers were initially tested on genomic DNA from *Arabidopsis* wild type L*er* and Col ecotypes to confirm the product sizes they would amplify and the presence of polymorphisms (Table 4.5) (Figure 4.10). PCR analysis was subsequently performed on the segregating male sterile lines from the mapping population. The male sterile lines were compared to Col and L*er* and scored for their genotype at each marker to identify recombinants around the mutant locus. Linkage was detected between 2Y and two of the PCR markers used, At3g26605 and At5g22545, which lie on chromosomes 3 and 5 respectively (Table 4.6). Analysis of the regions associated with these genes identified a number of previously characterised genes linked to anther and pollen development

(Table 4.7 and 4.8). On chromosome 3 AT3G26550, AT3G26600 and AT3G26618 were identified in the region 26, 000 bp upstream, 5, 000bp upstream and 25, 000 bp downstream, respectively, of the marker At3g26605. Of these three genes, AT3G26600 has been shown to have an involvement in male reproduction and to exhibit anther specific expression (Figure 4.11). AT5G22480 and AT5G22260 were analysed in the region 26, 000 bp and 113, 000 bp upstream of the marker respectively, of these AT5G22260 (MS1) has an involvement in male reproduction and shows anther specific expression (Figure 4.12) The marker on chromosome 5 (At5g22545) that showed linkage to the male sterility locus in this analysis (Table 4.6) was closely associated with the *MS1* locus (Table 4.8). This and the distorted segregation of two regions associated with the 2Y location may be a reflection of contamination and mapping of the *ms1* mutation within this population. This could have been the result of the ProMS1::*MS1-GFP* line that was used to backcross to the 2Ymutant carrying the *ms1* mutation; this would therefore mean that the *ms1* mutation was mapped alongside of the 2Y mutation.

Table 4.5 PCR analysis of genomic DNA from *Arabidopsis* wild type *Ler* and *Col* ecotype to confirm the product sizes they would amplify to test the presence of polymorphisms.

GENE ID	PRIMER	EXPECTED SIZE (BP) COLUMBIA	EXPECTED SIZE (BP) LANDSBERG *ERECTA*	PCR ANNEALI-NG TEMP. (^{0}C)
At1g07810	1	C >300	L<300	55
At1g09940	2	C >200	L<200	55
At1g30930	3	C = 239	L=142	55
At2g39010	4	C<150	L >150	55
At3g11220	5	C<200,	L >200	55
At3g50820	6	C 150/151	L= 130/135	55
At4g29860	7	C= 193	L= 174	55
At4g10360	8	C <200	L>200	55
At5g14320	9	C= 190	L= 215	55
At5g42600	10	C >150,	L <150	55
At5g22545	11	C= 492	L= 404	50
At5g63640	12	C= 268	L= 188	50
At1g49610	13	C= 225	L= 271	50
At2g14890	14	C= 100	L >130	53
At3g26605	15	C= 165	L= 145	53
At4g01710	16	C = 200	L>200	53
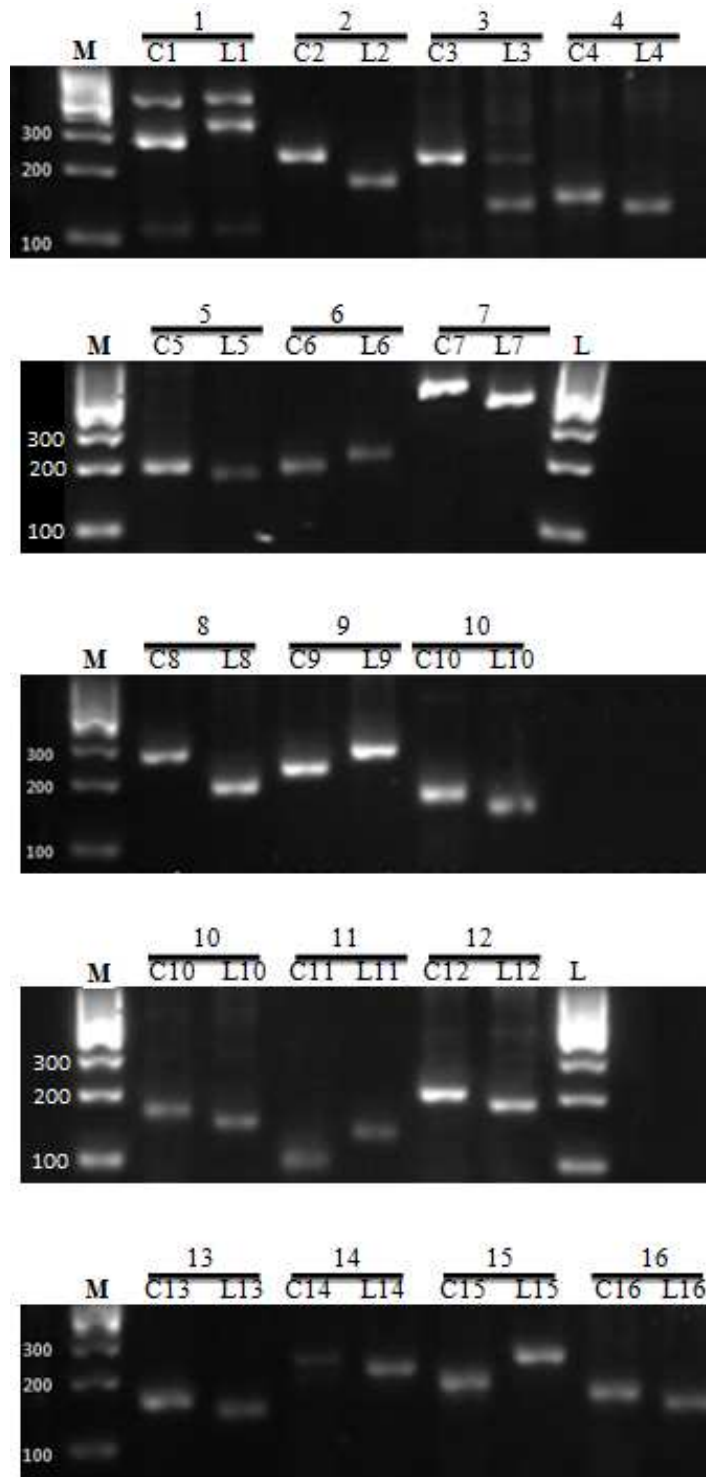
Figure 4.10 PCR polymorphisms between *Ler* and *Col* for the 16 different markers used for mapping the mutations on 5 chromosomal arms. C: Columbia, L: Landsberg *erecta*. Number represents primer number; M: marker.

Table 4.6 Linkage analysis of 2Y mutation with molecular markers on all of the Arabidopsis chromosome arms. Chi-square analysis indicated linkage between the 2Y mutation and two markers, At5g22545 and AT3g26605. A p-value of 0.05 or less is usually regarded as statistically significant, i.e. the observed deviation from the null hypothesis was significant. Highlighted region equates to markers showing linkage.

SUM L*ER*	SUM HET	SUM COL	SUM OF UNKNOWNS	F2 EXP L*ER*	F2 EXP HET	F2 EXP COL	P-VALUE	GENE
29	24	15	16	17	34	17	0.00295715	At1g07810
44	22	11	7	19.25	38.5	19.25	6.12002E-10	At1g09940
31	38	10	5	19.75	39.5	19.75	0.003555568	At1g30930
47	22	6	9	18.75	37.5	18.75	3.04532E-13	At1g49610
62	5	8	7	18.75	37.5	18.75	7.62084E-30	At2g14890
57	2	10	14	17.25	34.5	17.25	6.30881E-28	At2g39010
57	1	11	13	17.25	34.5	17.25	3.5849E-28	At3g11220
66	2	1	14	17.25	34.5	17.25	1.29121E-40	At3g26605
44	8	12	20	16	32	16	1.71391E-15	At3g50820
19	21	12	32	13	26	13	0.14899446	At4g01710
19	24	13	28	14	28	14	0.296922026	At4g29860
26	30	16	12	18	36	18	0.091731551	At4g10360
57	4	7	16	17	34	17	3.44557E-28	At5g14320
68	6	3	7	19.25	38.5	19.25	1.80015E-36	At5g22545
56	8	8	12	18	36	18	4.41172E-24	At5g42600
55	5	12	12	18	36	18	1.79402E-23	At5g63640

Table 4.7 Previously characterised genes linked to anther and pollen development that are associated with AT3G26605 marker.

GENE NAME	DESCRIPTION	REFERENCE
AT3G26618	Eukaryotic release factor 1-3 (ERF1-3) which functions in translation release factor activity. It is involved in translational termination. Located in the cytoplasm and expressed in male gametophyte, leaf, pollen tube. It is also expressed during pollen germination. It contains InterPro domains eRF1 domain 2 (InterPro: IPR005141), eRF1 domain 3 (InterPro: IPR005142), eRF1 domain 1 (InterPro: IPR005140), Peptide chain release factor eRF1/aRF1 (InterPro: IPR004403).	http://www.arabidopsis.org/servlets/TairObject?id=38154&type=locus
AT3G26600	Armadillo repeat protein. One of a family of four in Arabidopsis. Expressed in vegetative tissues, anthers and ovules.	http://www.arabidopsis.org/servlets/TairObject?id=38149&type=locus
AT3G26550	Cysteine/Histidine-rich C1 domain family protein which functions in zinc ion binding and is expressed in hypocotyls, roots and flowers during petal differentiation and expansion stage. It contains InterPro domain/s: Zinc finger, RING-type (InterPro: IPR001841), DC1 (InterPro: IPR004146), Zinc finger, PHD-type (InterPro: IPR001965), C1-like (InterPro: IPR011424).	http://www.arabidopsis.org/servlets/TairObject?id=38150&type=locus

Table 4.8 Previously characterised genes associated with anther and pollen development and show expression in the anther that are linked to the AT5G22545 marker.

GENE NAME	DESCRIPTION	REFERENCES
AT5G22480	ZPR1 zinc-finger domain protein which functions in zinc ion binding and expressed in male gametophyte and pollen tube during: mature pollen stage and germinated pollen stage respectively. It contains InterPro domains: Zinc finger, ZPR1-type (InterPro: IPR004457) and best *Arabidopsis thaliana* protein match is ZPR1 zinc-finger domain protein (TAIR: AT5G37340.3).	http://www.arabidopsis.org/servlets/TairObject?id=134224&type=locus
AT5G22260 *Male sterility 1* (*ms1*) mutant	Sporophytic factor controlling anther and pollen development. Mutants fail to make functional pollen. Pollen degeneration occurs after microspore release and the tapetum also appears abnormally vacuolated. Similar to PHD-finger motif transcription factors.	Li *et al*., 2011 Wang *et al*., 2012 Feng *et al*., 2012 Ito *et al*., 2007 Yang *et al*., 2007 Wilson *et al*., 2001
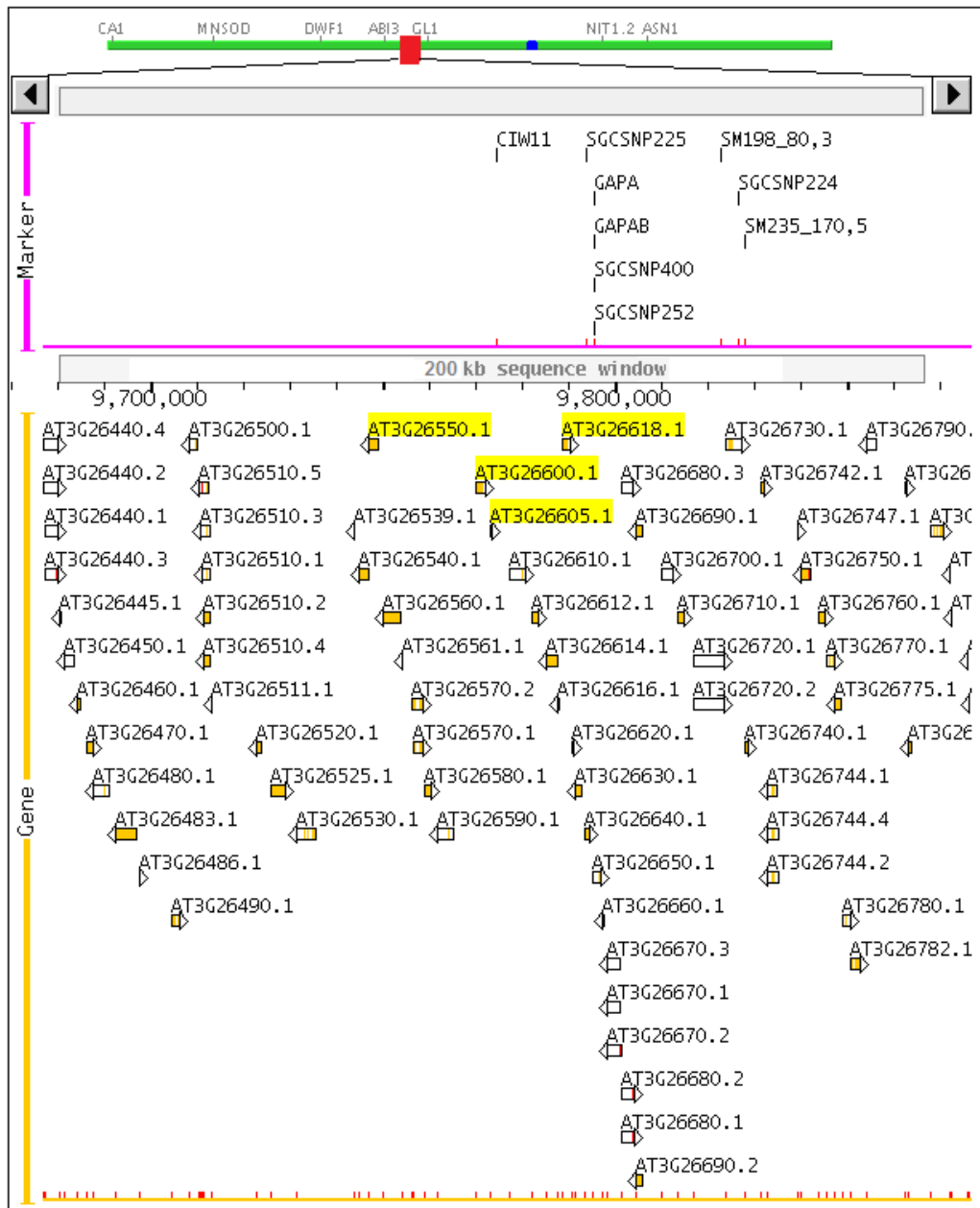
Figure 4.11: A close up map view of markers and gene models on chromosome 3. The Green line displays the map sequence for chromosomes 3. The red box on green line represents the exact position of the genes associated with the marker on AT3G26605. Blue circles on the map indicate the centromeric region. Highlighted in yellow represents AT3G26605 and the three previously characterised genes (AT3G26618.1, AT3G26600.1, and AT3G26550.1) associated with it and expressed in male gametophyte including anthers.
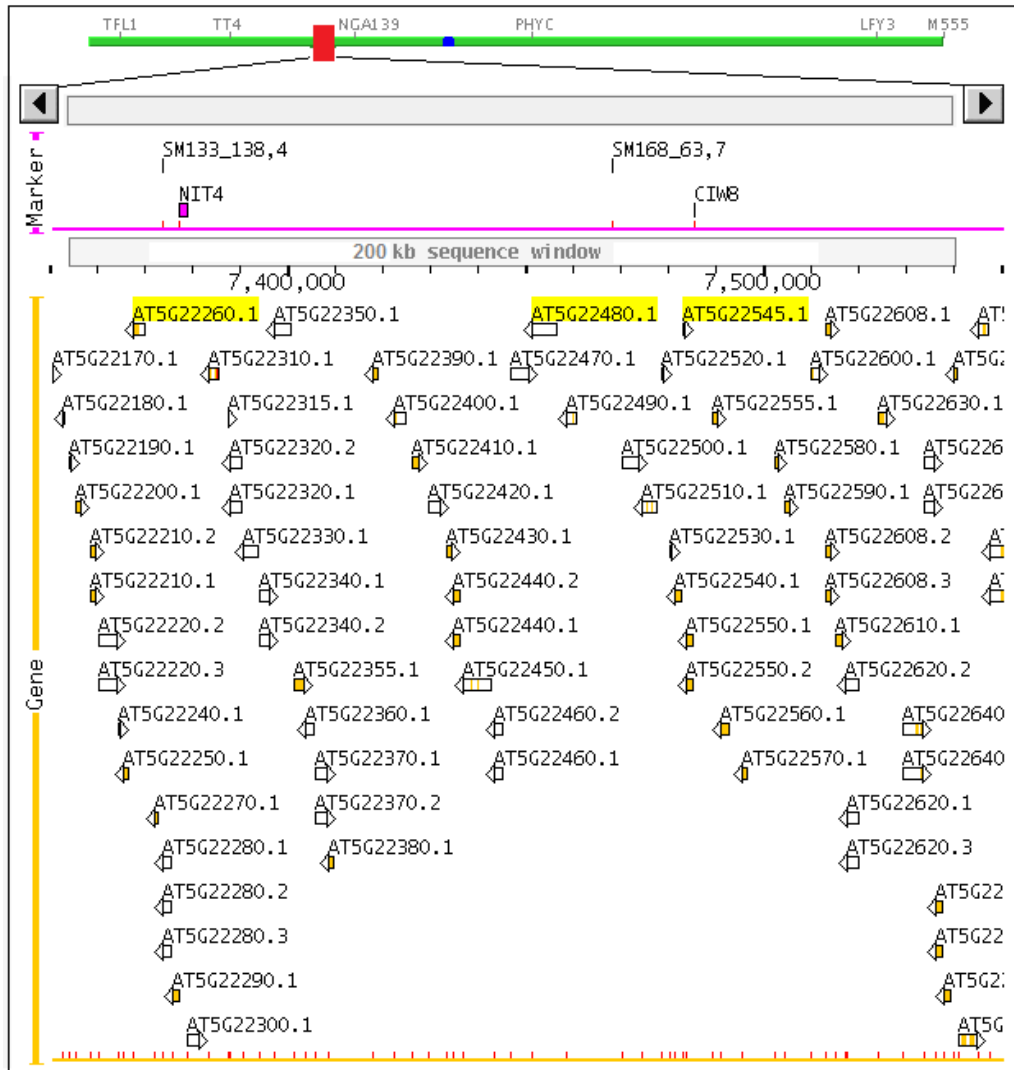
Figure 4.12: A close up map view of markers and gene models on chromosome 5. The Green line displays the map sequence for chromosomes 5. Red box on green line represent the exact position of the three genes associated with marker AT5G22545. Blue circles on the map indicate the centromeric region. Highlighted in yellow represents AT5G22545 and the two previously characterised genes (AT5G22260, AT5G22480) associated with the marker and expressed in male gametophyte during mature pollen stage.

4.3 DISCUSSION AND CONCLUSIONS

During the past decade major advances in genetic and genomic technologies have aided our understanding of pollen development at the molecular level. Dynamic interactions between transcription factors such as MS1 and AMS and their target genes during pollen development have been studied, however there are currently still significant gaps in our understanding of these networks (Sanders *et al*., 1999; Wilson *et al*., 2001, Sorensen *et al*., 2003; Scott *et al*., 2004; Ma, 2005; Vizcay-Barrena and Wilson, 2006; Yang *et al*. 2007; Ito *et al*., 2007; Wilson and Zhang, 2009; Feng and Dickinson, 2010; Xu *et al*. 2010; Wilson *et al*., 2011; Wang *et al*., 2012;Ma *et al*, 2012). The primary challenge to understand and measure the transcriptional regulation network activities has been at the genome-scale. Computational methods have been developed to infer the activities of transcription factor and predict the regulatory relationships between transcription factors and their target genes (Wang *et al*., 2011; Lamesch *et al*., 2012). The achievement also includes the highly annotated *A. thaliana* genome, comprehensive *A. thaliana* transcriptomics datasets, and various gametophytic mutants (www.plantgdb.org/AtGDB/; www.arabidopsis.org/; http://signal.salk.edu/cgi-bin/atta). Although significant progress has been made towards understanding pollen development at each development stage as indicated by the recent Plant Male Reproduction Database (http://www.pmrd.org), the dynamic regulatory network requires further characterization, particularly in identification of the key transcription factors, their regulators and their target genes.

In this work fast neutron ionising radiation was used as a mutagen to create mutations in genes linked to male fertility and in particular those causing failure of pollen development resulting in complete male sterility. Fast neutrons have been previously reported as a highly effective mutagen in *Arabidopsis* resulting in large deletions, whilst maintaining plant viability [Belfield *et al*., 2012; Bolon *et al*., 2011]. This was supported in this work in that a relatively high frequency of male sterile mutants were identified, with an average of 50% observed, whilst overall plant viability at both the M1 and M2 levels were not significantly affected. Analysis of male-sterile mutants is necessary for elucidating the mechanism of pollen development; however the target for this mutation was particularly focussed upon genes that had an effect on the expression of the MS1 transcription factor. This was detected by looking at expression of MS1-GFP fusion protein in the mutant lines by confocal microscopy to identify mutant lines showing altered expression of the MS1-GFP protein. Wildtype expression of MS1 is seen extremely transiently, for approximately a 12-hour period in the anther tapetum (Yang *et al*., 2007). However the male sterile mutants that were identified in this work showed prolonged expression of MS1 and also failure of pollen development. These mutants may therefore involve genes associated with the regulatory control of MS1. Allelism tests identified that three of the mutants were allelic to MS1 and may therefore represent new allelic versions of the MS1 locus. However, the 2Y mutant was found to be novel, with the allelism cross between *ms1* and 2Y showing complementation of both the 2Y and *ms1ttg* mutations. Mapping of this mutation indicated that there was a distortion of segregation associated with two of the markers analysed which were associated with different chromosomes, At3g26605 and At5g22545 on chromosomes 3 and 5, respectively.

It was unexpected to see linkage on two separate chromosomes; however the marker At5g22545 shows very close association with the *MS1* locus (At5g22260). It therefore seems possible that there may have been a level of *ms1* mutant contamination in the line used for backcrossing and mapping of the 2Y line. This line carried the MS1Prom::MS1-GFP fusion protein therefore would be fertile due to the presence of the transgene regardless of whether the line was carrying mutations in the *ms1* background, this had been previously shown for functionality by transformation into and rescue of the *ms1ttg* mutant (Yang et al, 2007). The presence of the transgene confers fertility to the *ms1* mutant therefore presence/segregation for the *ms1* mutation in lines carrying the transgene would not be easily detectable, except by molecular screening. There was also a higher than expected incidence of the L*er* genotype in the segregating population, this could also be explained based on the mapping of an additional mutation (*ms1* in the L*er* background). This suggests that the 2Y mutation may actually reside on chromosome 3 linked to At3g26605, however of course it could be that the mutation is an additional mutation found in the region of the *MS1* locus. Further molecular analysis of the lines showing male sterility and linkage with the At5g22545 marker needs to be conducted to confirm this. Moreover, the use of next generation sequencing technology may provide a faster way to pinpoint the 2Y gene. With next generation sequencing it is possible to resequence the entire plant genome more efficiently and economically and in greater depth than ever before (Varshney *et al*., 2009). Therefore sequencing the entire genome of the 2Y mutant would potentially reveal the identity of the mutant gene. This is particularly the case since a basic map location has already been established that could be used to confirm mutant identify, however this strategy provides

additional complications if there is a level of background mutations still remaining in the 2Y mutant population.

Nevertheless there are a number of previously characterised genes that have been linked to anther and pollen development that are associated with the At5g22545 and At3g26605 markers (Table 4.7 and 4.8). For example AT3G26618.1, AT3G26600.1 and AT3G26550.1 are associated with At3g26605 marker and of these AT3G26600.1 is 5, 000bp upstream of the marker At3g26605 (Figure 4.11) and has an involvement in male reproduction and show anther specific expression. In addition to At5g22260 (*MS1),* AT5G22480.1 is also associated with the AT5G22545 marker and this could potentially play a role in relation to *MS1* expression as it is a ZPR1 zinc-finger domain protein which functions in zinc ion binding and is expressed in the male gametophyte during the mature pollen stage. AT5G22480.1 was analysed in the region of 26, 000 bp upstream of the AT5G22545 marker a closer linkage to the marker compared to At5g22260 which was 113, 000bp upstream of the AT5G22545 markers (Figure 4.12).

Further work to characterise the exact location of 2Y could involve the analysis of more markers in this region to provide a tighter resolution of the exact map position. In addition molecular characterisation could be conducted on the potential targets in the region to see if they are showing normal expression of the genes in the mutant background. Currently the information available about these putative genes is very limited, except that they have been linked to pollen development and shown to be expressed in floral organs during *Arabidopsis* male

gametogenesis (Table 4.7 and 4.8); further complementation analysis of 2Y mutant could also be conducted to confirm the gene identity. This data will then provide greater characterisation of genes involved in MS1 regulatory network and pollen development. Once the putative gene locus has been identified, the gene could be confirmed by complementation of the 2Y mutant phenotype using the corresponding wildtype sequence, or phenotypic and complementation analysis of additional mutant alleles of the putative gene of interest (for example using SALK insertional mutant lines).

CHAPTER 5: MS1 PROTEIN-PROTEIN

INTERACTIONS

5. MS1 PROTEIN-PROTEIN INTERACTIONS

5.1 INTRODUCTION

Protein-protein interactions and interactions between proteins and other macromolecules are necessary for cell function and survival. Identifying interactions between proteins involved in common cellular functions provides us with a broad view of how they work cooperatively in a cell. As protein–protein interactions are so important there are a multitude of methods to detect them and each of the approaches has its own strengths and weaknesses, especially with regard to the sensitivity and specificity of the methods. Biochemical, biophysical and theoretical approaches are well known methods of investigating protein interactions. In this study the Yeast-2-Hybrid system was used to analyse proteins thought to interact with the MS1 protein using an Arabidopsis anther specific Yeast-2-Hybrid library previously generated by Yang and Wilson, (unpublished data). Subsequent validation of these putative interactions was then conducted.

5.1.1   THE YEAST-2-HYBRID SYSTEM

The Yeast-2-Hybrid system is an *in vivo* yeast-based system that identifies the interaction between two proteins (X and Y) by reconstituting an active transcription factor. The active transcription factors are formed as a dimer between two fusion proteins, one of which contains a DNA-Binding Domain (DB) fused to the first protein of interest (DB-X; also known as the "bait") and the other, an Activation Domain (AD) fused to the second protein of interest (AD-Y;

also known as the "prey" or "target protein"). DB-X:AD-Y interaction reconstitutes a functional transcription factor that activates chromosomally-integrated reporter genes driven by promoters containing the relevant DB binding sites (Figure 5.1)**.**
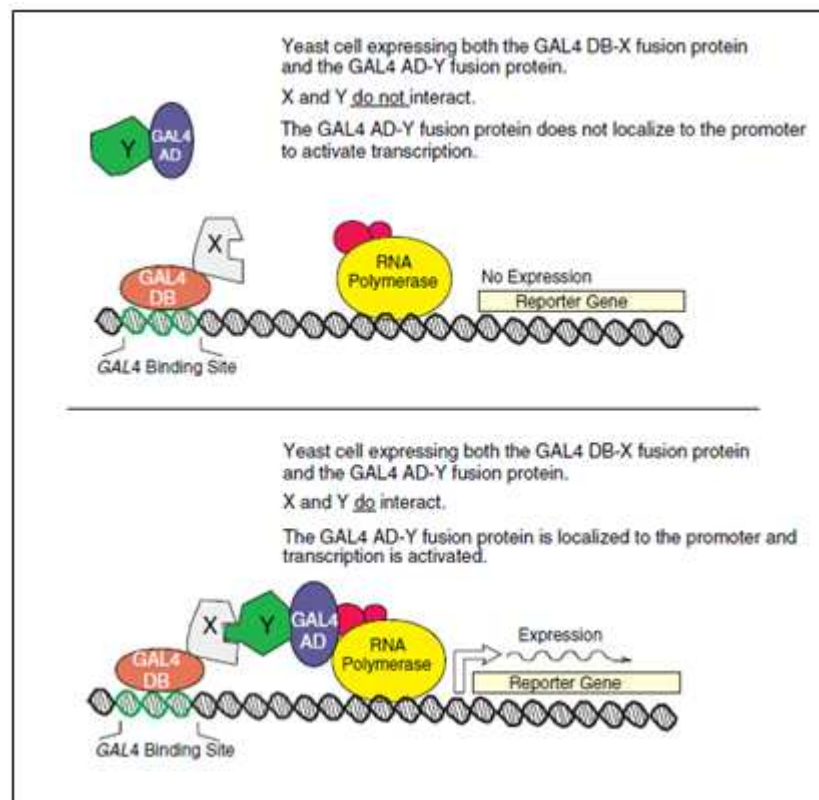
Figure 5.1: Overview of the Yeast-2-Hybrid assay, checking for interactions between two proteins, called here X and Y. *Gal4* transcription factor gene produces two domain proteins (*DB* and *AD*), which is essential for transcription of the reporter gene (*LacZ*). Two fusion proteins are prepared: *Gal4DB+X* and *Gal4AD+Y*. Neither of them is usually sufficient to initiate the transcription (of the reporter gene) alone. When both fusion proteins are produced and the X part of the first interacts with the Y part of the second, transcription of the reporter gene occurs (Image from Invitrogen Life Technologies Instruction Manual).

When a selectable marker such as *HIS3* is used as a reporter gene, Yeast-2-Hybrid dependent transcription activation can be monitored by growth of cells on plates lacking histidine, thereby providing a means to detect protein:protein interactions genetically. The Y2H system has become one of the techniques that has proven extremely valuable for the mapping of protein-protein interactions on a global scale and has thereby provided interesting functional insights. It offers many advantages over biochemical methods, such as cost, convenience and sensitivity (Causier and Davies, 2002). However it still has several associated problems for studying protein-protein interactions *in vivo* (Causier and Davies, 2002). In particular false-positives are often generated in two-hybrid library screens; one approach used to minimise this has been the use of two or more reporter genes to assay for an interaction, however interactions are generally confirmed with alternative subsequent procedures. In this work a Yeast-2-Hybrid screen was used as a starting point for the identification of MS1 protein interactions, whilst other independent approaches, such as transient expressions of gene constructs in tobacco/onion epidermal cells, bifunctional fluorescent complementation and recombinant protein expression in *E.coli.* were used for subsequent confirmation and validation of these interactions.

5.1.2   TRANSIENT EXPRESSION OF PUTATIVELY INTERACTING PROTEINS IN TOBACCO

Agrobacterium tumefaciens-mediated transient expression of proteins in tobacco has been a useful procedure for characterization of proteins and their functions in

plants, including analysis of protein-protein interactions (Tsuda *et al*., 2012). It has been widely used due to its ease and high efficiency (Tsuda *et al*., 2012). In comparison with stable transformation, it has shown the advantages of being simple, quick, economical, and effective (Zheng *et al*., 2012). However, the method is limited in its applications because expression levels are low and transient in nature (typically expression disappears after less than 5 days from inoculation) (Voinnet *et al*., 2003). Post-transcriptional gene silencing (PTGS) is a major cause for this lack of efficiency. A system based on co-expression of a viral-encoded suppressor of gene silencing, the p19 protein of Tomato Bushy Stunt Virus (TBSV) prevents the onset of PTGS in the infiltrated tissues and allows high level of transient expression (Voinnet *et al*., 2003). Recently a study conducted by Zheng *et al*., (2012) presented a new approach which could improve transient gene expression system based on *Agrobacterium*-mediated transformation. They mentioned that hyperosmotic pre-treatment of plants significantly improved transient expression in this system. Furthermore, other factors they mentioned which could significantly influence transient expression efficiency included acetosyringone concentration, co-cultivation time, and *Agrobacterium* cell density. In this work transient expression of MS1 and POB2 gene constructs was performed in *Nicotiana benthamiana* leaves to test for protein-protein interactions (see section 5.2.2 for detailed analysis).

5.1.3 *E.COLI* PROTEIN EXPRESSION USING GATEWAY CLONING TECHNOLOGY

The *E.coli* expression system with Gateway technology uses Gateway-adapted destination vectors designed for cloning and high-level inducible expression of recombinant proteins in *E. coli* using the pET system (Studier *et al.*, 1990). The pET system was originally developed by Studier and colleagues and takes advantage of the high activity and specificity of the bacteriophage T7 RNA polymerase to allow regulated expression of heterologous genes in *E. coli* from the T7 promoter (Studier and Moffatt, 1986; Studier *et al.*, 1990). Depending on the vector chosen, the pDEST™ vectors allow production of native, N-terminal, or C-terminal-tagged recombinant proteins. N- or C-terminal fusion tags for detection and purification of recombinant fusion proteins and choice of tag depends on the particular vector used (http://tools.invitrogen.com). Recombinant protein expression in *E. coli* is simple, fast, inexpensive, and robust, with the expressed protein comprising up to 50 percent of the total cellular protein (Francis and Page, 2010). However, it also has disadvantages, for example, the rapidity of bacterial protein expression often results in unfolded/misfolded proteins, especially for heterologous proteins that require longer times and/or molecular chaperones to fold correctly (Francis and Page, 2010). In addition, the highly reductive environment of the bacterial cytosol and the inability of *E. coli* to perform several eukaryotic post-translational modifications results in the insoluble expression of proteins that require these modifications for folding and activity (Francis and Page, 2010). Fortunately, multiple, novel reagents and techniques

have been developed that allow for the efficient, soluble production of a diverse range of heterologous proteins in *E. coli* (Voulgaridou *et al*., 2013).

In the past concerns have been raised about the Gateway cloning with the presence of the residual *attB* sites in the final expression clone constructs (Perehinec *et al*., 2007). However Perehinec *et al*. (2007) reported that the presence of *attB* sites within the expression cassettes did not cause significant differences in rates of induction, although the amount of resultant protein appeared to be lower where *attB* sites were introduced proximal to the promoter. Another concern often expressed when using reporter genes for *in vivo* studies is that the expression of the reporter normally places a metabolic burden on the bacteria and therefore alters growth behaviour (Perehinec *et al*., 2007). It is clear that problems of vector maintenance may be of more significance than any metabolic burden of the reporter genes (Perehinec *et al*., 2007). In addition antibiotic selection alone may be not sufficient to ensure that all cells in a population retain reporter gene constructs, and therefore the choice of antibiotic marker gene is also an important factor when designing vectors. Furthermore past studies by Belfield *et al*., (2007) also suggested that caution should be exercised when employing pDEST17 as a Gateway (Invitrogen) bacterial expression vector for high-level protein expression as it is intrinsically susceptible to −1 ribosomal frameshifting at the sequence C-AAA-AAA (Belfield *et al*., 2007).

The European Molecular Biology Laboratory (EMBL) have also reported that when the levels of protein expression are low, expression levels of the target protein can be optimized by varying the time and/or temperature of induction and

the concentration of the inducer. Using protease-deficient host strains carrying mutations which eliminate the production of proteases can sometimes enhance accumulation by reducing proteolytic degradation. Despite all the setbacks mentioned, the Gateway cloning system was effectively utilised in this study to generate new constructs in different vectors (refer to section 5.2.3).

5.1.4 BIMOLECULAR FLUORESCENCE COMPLEMENTATION

Bimolecular fluorescence complementation (BiFC) is a technique used to investigate the formation of protein complexes and the localization of protein–protein interactions *in planta*. This technique is based on formation of fluorescent complex through the association of two partial fragments of a fluorescent protein; these are brought together by an interaction between the proteins maintained as fusions with the split fluorescent tags (Figure 5.2) (http://sitemaker.umich.edu/kerppola.lab/_protein_interactions_). BIFC involves making a construct of one protein with the C-terminal part of the YFP protein and then another with the N-terminal part of the YFP protein. They are then transiently expressed in plant cells, either onion epidermis/protoplasts or *N. benthamiana* leaves. YFP fluorescence should be seen if the proteins interact (Bhat *et al*., 2006) (Figure 5.2). For example two fragments (YN and YC) of the yellow fluorescent protein (YFP) are fused to two putative interaction partners (A and B) (Figure 5.3) (Kerppola, 2006). An interaction between the proteins facilitates association between the fragments to produce a bimolecular fluorescent complex (Figure 5.3) (Kerppola, 2006); if YN and YC do not interact on their own no fluorescence is seen (Kerppola, 2006). One advantage of BiFC is that it

provides a physiologically-relevant context from which to draw conclusions about protein interactions by enabling direct visualisation of protein interactions in living cells with limited cell perturbation, rather than relying on secondary effects or staining by exogenous molecules that can fail to distribute evenly (Morell *et al*., 2008). This and the ability to observe the living cells for long periods of time has been made possible by the strong intrinsic fluorescence of the reconstituted reporter protein which reduces the chances of an incorrect readout associated with the protein isolation process (Morell *et al*., (2008); Fan *et al*., (2008).
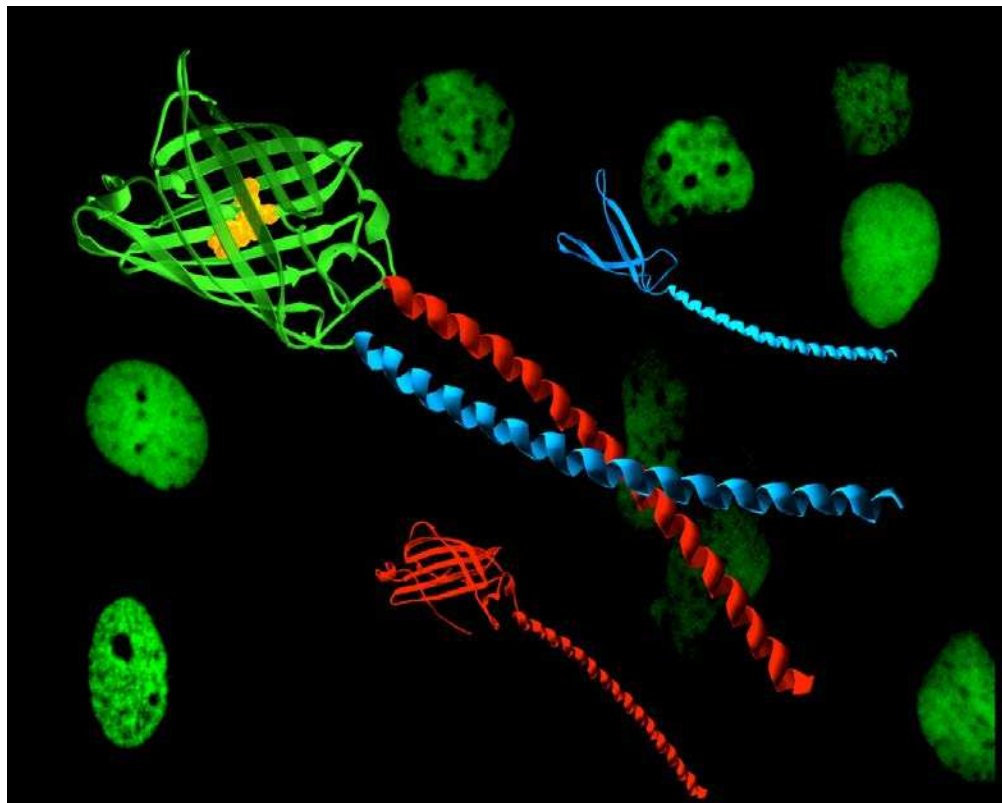
Figure 5.2 A fluorescent protein complex (yellow) is formed by the association of two fragments (red and blue fragment) of a yellow fluorescent protein (YFP) brought together (Bhat *et al*., 2006).
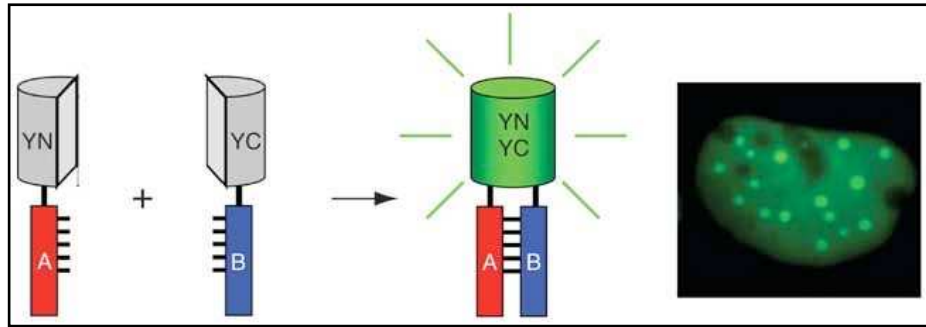
Figure 5.3: Two fragments (YN and YC) of the yellow fluorescent protein (YFP) are fused to two putative interaction partners (A and B). An interaction between the proteins facilitates association between the fragments to produce a bimolecular fluorescent complex (Kerppola, 2006).

Unlike many *in vivo* protein-interaction assays, BiFC does not require protein complexes to be formed by a large proportion of the proteins or at stoichiometric proportions. Instead, BiFC can detect interactions among protein subpopulations, weak interactions, and low expression proteins due to the stable complementation of the fluorescent reporter protein (Hu *et al*., 2003; Rackham *et al*., 2004). Furthermore, the strength of the protein interaction can be quantitatively determined by changes in fluorescent signal strength and does not require specialised equipment, as visualisation is possible with an inverted fluorescence microscope that can detect fluorescence in cells and analysis does not require complex data processing or correction for other sources of fluorescence (Kerppola *et al*., 2008; Tsien, 1998; Richards, 1958). Despite many benefits of the BiFC approach, this technique is unable to provide real-time detection of protein interactions (Morell *et al*., (2008). The delay for chemical reactions to generate fluorophore may also have an effect on the dynamics of complex dissociation and partner exchange (Morell *et al*., 2008; Hu *et al*., 2002; Jach *et al*., 2006; Tsien,

1998). With BiFC the exact interaction relationship is unknown because fluorophore reconstitution can occur at a distance of 7nm or more. However fluorescence complementation may indicate either a direct or indirect (i.e. within the same complex) interaction between the fluorescent fragments' fused proteins (Fan *et al*., 2008). Section 5.2.4 describes the detailed methodology for BiFC analysis to detect MS1 interaction with POB2.

5.2 METHODS

5.2.1   YEAST-TWO-HYBRID SYSTEM

An Arabidopsis stamen Yeast-2-Hybrid library was constructed using Gateway compatible vectors that comprised of all stages of anther and filament tissue (Table 5.1). The fusion proteins were prepared according to manufacturer's instruction manual (ProQuest™ Two-Hybrid System with Gateway® Technology). A previous screen of this library using the MS1 sequence without the PHD motif identified a putative interacting protein POB2. POB2 is a BTB/POZ domain-containing protein that may be involved in ubiquitin-based proteolytic breakdown (Qu *et al*., 2010). POB2 interaction with MS1 was subsequently confirmed by further Yeast-2-Hybrid pair-wise analysis (Figure 5.4). In addition to MS1 interaction with POB2, POB1 interaction with MS1 was also conducted (Figure 5.5).

Table 5.1: Gene constructs of Gateway compatible vectors expressing either the Gal4 Binding Domain or Gal4 activating Domain (Invitrogen).

GENE	GATEWAY DESTINATION VECTOR	GAL4 TRANSCRIPTION FACTOR DOMAIN
MS1 with no PHD motif	pDEST32	Binding domain
POB1	PDEST32	Binding domain
POB2	pDEST22	Activating domain
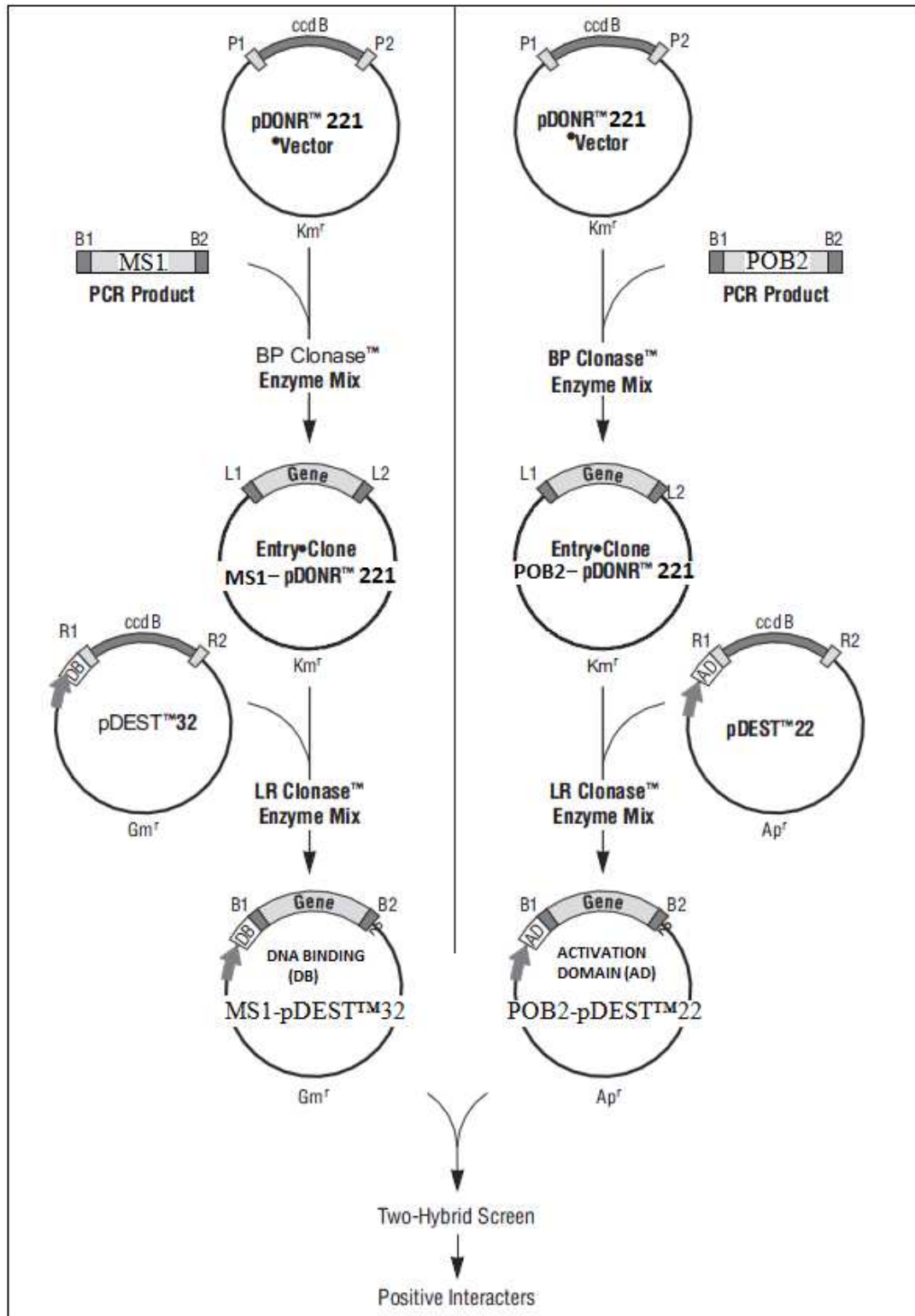
Figure 5.4: An overview of how the Gateway ProQuest Yeast-2-Hybrid analysis (Invitrogen) was used to analyse the interaction between MS1 and POB2. For more information refer to Invitrogen ProQuest™ Two-Hybrid System a manufactures manual.
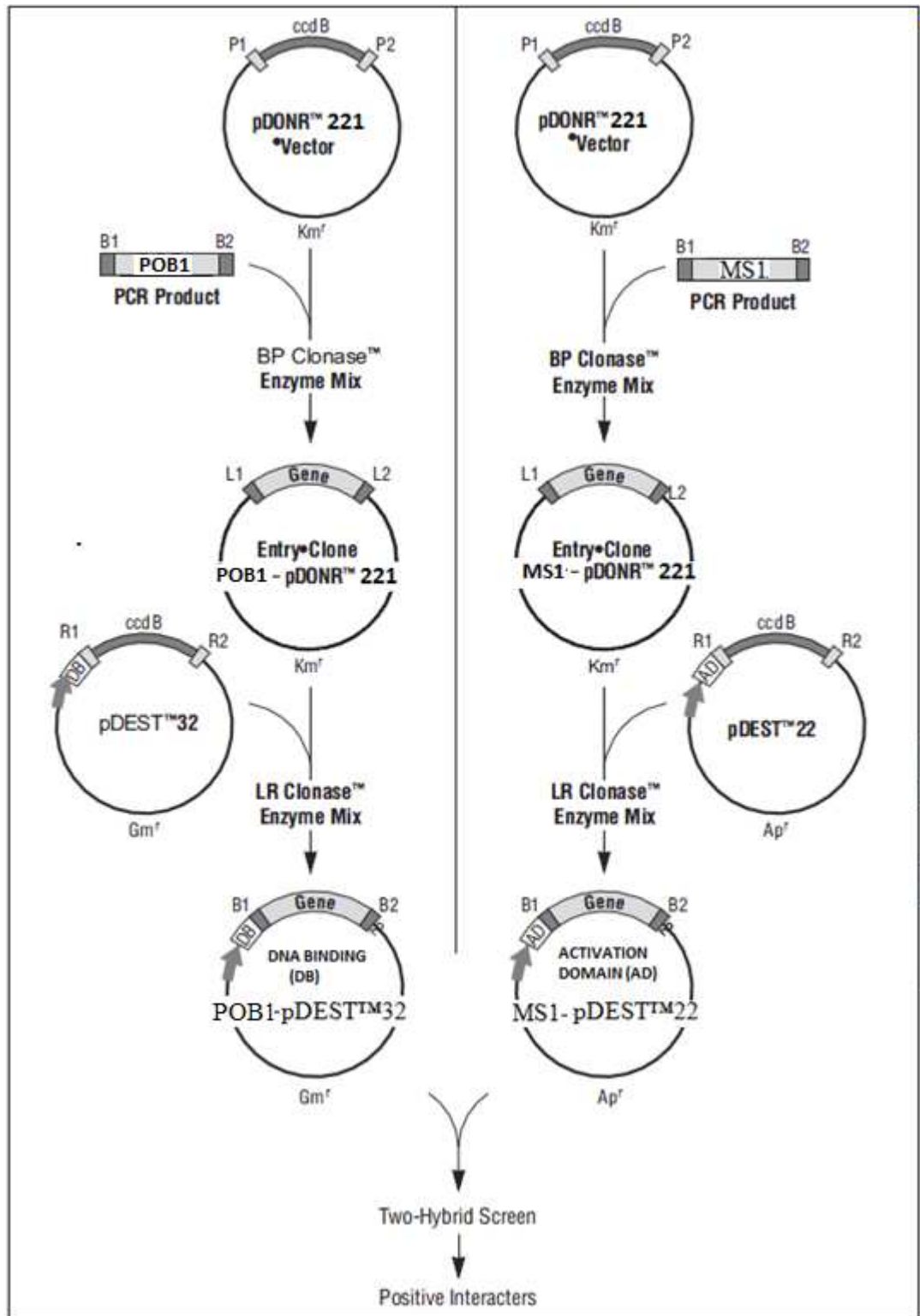
Figure 5.5: An overview of how the Gateway ProQuest Yeast-2-Hybrid analysis (Invitrogen) was used to analyse the interaction between MS1 and POB1. For more information refer to Invitrogen ProQuest™ Two-Hybrid System a manufactures manual.

5.2.2 TRANSIENT EXPRESSION OF MS1 AND POB2 GENE CONSTRUCTS IN TOBACCO.

Nicotiana benthamiana seeds (from University of Nottingham) were grown at 22ºC +/-2^{0}C in pots containing compost immersed in 0.15g.l-1 of Intercept® solution (Scotts, U.K.) (Chapter 2, Section 2.2). Plants used for transient assays were grown for 5-6 weeks (Kapila *et al*., 1997). To transiently express MS1 and POB2 in tobacco leaves the genes were cloned into Gateway® entry vector pDONR™221 respectively to create entry clones (Chapter 2 Section 2.8.9). The LR recombination reaction was carried out to transfer MS1 and POB2 into PGWB5 (T7 Promoter-N-terminal-GFP tagged region) and PGWB8 (T7 Promoter-C-Terminal-6xHIS tagged region) respectively (Table 5.2). The LR reaction mixture was transformed into *E.coli* DH5a competent cells (supplied by Invitrogen) by heat shock transformation (Section 2.8.8). The culture was spread on LB plates containing kanamycin 100µg/ml and hygromycin 50µg/ml and incubated overnight at 37˚C. Positive clones were identified by plasmid DNA-PCR using appropriate vector primer and gene specific primer pairs. Plasmid DNA was extracted from the expression clones by Miniprep (2.8.10) and transformed into competent *Agrobacteria* cells (GV3101) (Section 5.2.2.2) for transient assays.

Table 5.2 Gene constructs in Gateway compatible vectors (Nakagawa *et al*., 2007) utilised for transient protein expression experiments.

GENE	GATEWAY DESTINATION VECTOR	FEATURES
MS1	PGWB5	(--35S promoter–R1- CmR-ccdB-R2–sGFP--)
POB2	PGWB8	(--35S promoter-R1- CmR-ccdB-R2-6xHis--)

5.2.2.1 TRANSFORMATION OF ELECTRO-COMPETENT *AGROBACTERIUM* GV3101 CELLS WITH PGWB5-MS1 AND PGWB8-POB2 GENE CONSTRUCTS

Competent *Agrobacteria* cells (GV3101) were prepared by Doris Klisch (University of Nottingham) and stored in 50μl aliquots at -80^0C were thawed on ice and 2μl (200ng) of plasmid DNA (PGWB5-MS1 and PGWB8-POB2) was added. The cells were then transferred to a pre-chilled electroporation cuvette (BioRad) and pulsed 500 rpm with the MicroPulserTM electroporation (BioRad) for 1min. 1mL of LB medium was added to the cells, which were then transferred to a microfuge tube and incubated at 28ºC with constant shaking 200rpm for 3h to allow expression of the antibiotic resistance genes. Cells were then pelleted, resuspended in 100μl LB medium and plated on agar plates containing the appropriate antibiotics for selection. The plates were incubated at 28ºC for 2 days until colonies developed. Plasmid DNA PCR was conducted on selected colonies to confirm the presence of the transgenes.

5.2.2.2 PREPARATION OF TRANSIENT ASSAYS FOR *N. BENTHAMIANA* TRANFORMATION

Transient assays in *N. benthamiana* were performed as described in Hammond-Kosack *et al.*, (1998) to confirm the interaction between MS1 and POB2. *Agrobacterium* GV3010 glycerol stock containing PGWB5-MS1 and PGWB8-POB2 gene construct were grown on LB plates at 28ºC for 2 days. A few colonies were then selected and grown overnight at 28ºC with constant shaking (200rpm) in 10mL LB culture vials containing the appropriate antibiotics. The culture was pelleted (1957 x g) 10min and washed once with 10mM $MgCl_2$. The cells were then suspended in 10mM $MgCl_2$ containing 200mM acetosyringone and the OD_{600} was adjusted to 0.2. The cultures were left for 2-3 hours at room temperature before infiltration. The *Agrobacteria* cultures were infiltrated into *N. benthamina* leaves by pressure-infiltration on the lower side of the leaves using a 1mL syringe. The infiltrated plants were then put in a controlled growth cabinet at 22ºC for 2-3 days to allow for gene expression. Protein was later extracted from leaf material infiltrated with both PGWB5-MS1 and PGWB8-POB2 (Chapter 2 Section 2.9.1). Non-infiltrated leaves were used as negative control as well as those infiltrated with either PGWB5-MS1 only or PGWB8-POB2.

5.2.2.3 EXTRACTION, SEPARATION AND VISUALIZATION OF RECOMBINANT FUSION PROTEINS

Plant tissue from infiltrated *N. benthamiana* leaves were harvested for total protein extraction according to protocol in Chapter 2 Section 2.9.1. The total

protein fraction was then mixed with 1x SDS loading buffer [25mM Tris-HCl pH6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 5% (v/v) beta-mercaptoethanol, 0.001% bromophenol blue] and boiled for 5min before being separated by SDS-PAGE (Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis). A pre-stained protein marker (New England Biolabs) was run alongside the samples and used as a reference. A 10% precast SDS-PAGE separating gel was used [0.38M Tris-HCl pH8.8, 10% (w/v) acrylamide 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.07% TEMED] to separate proteins according to their molecular weight (Chapter 2 Section 2.9.2). The gel was fitted in a Mini- PROTEAN II cassette (BioRad) filled with SDS running buffer [25mM Tris-HCl pH8.5, 190mM glycine and 1% (w/v) SDS] and electrophoresed at 100V until they reached the end of the stacking gel, after which the voltage was increased to 150V. All proteins were fractionated on an SDS-PAGE precast gel. The bands were visualised by incubating the gel in Coomassie stain solution (50% (v/v) methanol, 13% (v/v) acetic acid, 0.2% (w/v) Coomassie Brilliant Blue R250) for 2 hours at room temperature with gentle agitation.

5.2.2.4 IMMUNODETECTION OF GFP (MS1-PGWB5) AND GST (POB2-PGWB5) TAGGED PROTEINS

After the SDS-PAGE run, proteins were transferred to a PVDF membrane (BioRad) using the Mini-PROTEAN Trans-Blot transfer cassette (BioRad) filled with cold transfer buffer [25mM Tris-HCl, 190mM glycine, 20% (v/v) methanol] (Chapter 2 Section 2.9.5). The membrane was incubated for 1-3 hours in anti-GFP primary antibodies diluted (1:5000) in TBS-T containing 3% non-fat dried milk.

Membranes were then washed several times with TBS-T for a total of 20min before incubation for 45 min with goat anti-mouse secondary antibodies. The exact concentrations of the primary and secondary antibodies used in this study are described in Tables 2.13 and Table 2.14. Following the incubation, the membranes were washed five times with TBS-T for a total of 25min. Immunodetection was carried out using Immobilon (Millipore) for chemiluminescent detection in accordance with the manufacturer's instructions (Chapter 2 Section 2.9.6). The membrane, washed from any unbound HRP-conjugated secondary antibody, was incubated in 1:1 Luminol Reagent and Peroxide Solution for 5min at room temperature. Excess substrate was drained and the membrane was covered in cling film and placed in an X-ray cassette for overnight. The membrane was then exposed to X-ray film (Kodak) under safe red light conditions. The film was developed using the X-OMAT developing system.

5.2.3 *E.COLI* EXPRESSION SYSTEM WITH GATEWAY® TECHNOLOGY

The *E. coli* Expression System with Gateway® Technology comprises a series of Gateway®-adapted destination vectors (Chapter 2 Section 2.8.9 Table 2.9) designed to facilitate high-level, inducible expression of recombinant proteins in *E. coli*. Depending on the vector chosen, the pDEST™ vectors allow production of Native, N-terminal, or C-terminal-tagged recombinant proteins. To express the gene of interest in *E. coli* using the Gateway® Technology MS1 and POB2 full cDNA sequences were cloned using the BP recombination reaction into Gateway® entry vector pDONR™221, respectively to create entry clones. The

entry clones were transformed into *E. coli* by heat shock (Chapter 2 Section 2.8.8) and plated on 50ug/ml of kanamycin for selection purposes and grown at 37°C overnight. Plasmid DNA PCR was conducted using appropriate primers to identify positive clones. To transfer entry vectors MS1: pDONR™221 and POB2: pDONR™221 into pDEST15 and pDEST17 respectively (Table 5.3), chemically competent DH5α *E. coli* cells were transformed with the LR reaction by heat shock (Chapter 2 Section 2.8.8-2.8.9) and plated on 100μg/ml of ampicillin for selection purposes. Plates were incubated at 37°C overnight. Plasmid DNA PCR was conducted using appropriate primer to confirm positive clones. Plasmid DNA was prepared from positive expression clones and was then transformed into *E.coli* BL21 cells for protein induction (Section 5.2.3.1).

Table 5.3: Gene constructs in Gateway compatible vectors (Invitrogen) transiently expressed in *E.coli* cells.

GENE	GATEWAY DESTINATION VECTOR	FEATURES
POB2	pDEST17	(T7 promoter- N-terminal, 6XHis Tag)
MS1	pDEST15	(T7 promoter, C-terminal, GST Tag)

5.2.3.1 TRANSFORMATION OF BL21 (DE3) CELLS WITH PURIFIED PLASMID DNA FOR EXPRESSION ANALYSIS

E.coli BL21 (DE3) cells can be used with protein expression vectors under the control of the T7 promoter, such as pET vectors (http://www.molecularinfo.com/MTM/A/A3/A3-1.pdf). This strain is lysogenic for lambda-DE3, which contains the T7 bacteriophage gene 1, encoding T7 RNA polymerase under the control of the lacUV5 promoter. BL21 (DE3) also contains the pLysS plasmid, which carries the gene encoding T7 lysozyme. T7 lysozyme lowers the background expression level of target genes under the control of the T7 promoter but does not interfere with the level of expression achieved following induction with L-arabinose. Plasmid DNA of the expression clones MS1-pDEST15 and POB2-pDEST17 was transformed into BL21 (DE3) separately by heat shock transformation (Chapter 2 Section 2.9.9). Two different volumes of the transformation reaction were plated onto LB plates containing 34 μg/mL chloramphenicol and 100 μg/mL of ampicillin. The plates were inverted and incubated at 37°C overnight. Transformants from the plates were confirmed by plasmid DNA PCR using appropriate primers. The plasmid DNA was also sequenced to validate the constructs. Pilot expression of the recombinant proteins was them performed by adding L-arabinose to a final concentration of 0.2% to the cultures to induce expression of the recombinant proteins.

5.2.3.2 PILOT EXPRESSION OF MS1-pDEST15 AND POB2-pDEST17 RECOMBINANT PROTEINS

Transformants from BL21 (DE3) confirmed by plasmid DNA PCR were cultured in 5 ml of LB medium containing 100μg/ml ampicillin. The cells were grown overnight in 37°C incubator until the OD_{600} reached ≥2. The overnight cultures were used to inoculate fresh LB medium containing 100 μl/ml ampicillin to an OD_{600} of 0.05–0.1 (~1:50 dilution of the overnight culture). This dilution allowed the cells to quickly return to logarithmic growth and reach the appropriate cell density. The cultures were left to grow until they reached mid-log phase OD_{600}=~0.4. Each culture was split into two cultures, one culture was uninduced and the other culture was induced by adding L-arabinose to a final concentration of 0.2% w/v. The cultures were incubated for an additional 2–3 hours. Time points were taken to analyse expression of the proteins. For each time point, 500μl of culture was removed from the induced and uninduced cultures. The cells were harvested by centrifugation (10 000xg) and the cell pellet stored at –80°C for native protein extraction, separation, visualisation, western blotting and detection (Chapter 2, Section 2.9.1-2.9.6).

5.2.4 TRANSIENT EXPRESSION OF MS1 AND POB2 GENE CONSTRUCTS IN ONION EPIDERMAL CELLS BY PARTICLE BOMBARDMENT FOR BiFC ANALYSIS

In this project BIFC was performed to check the interaction of MS1 and POB2 in which a construct of POB2 protein with the C-terminal part of the YFP protein and MS1 with the N-terminal part of the YFP protein was transiently expressed in the same onion epidermal cells by particle bombardment (Table 5.4). Initially The LR recombination reaction was carried out to transfer MS1 and POB2 into PG005 (YFP-N-terminal region) and PG006 (YFP-C-Terminal region) respectively (PG005 and PG006 vectors were obtained from Dr Rupert Fray University of Nottingham). Plasmid DNA was extracted from the expression clones and stored at -20^0C in readiness for bombardment into onion epidermal cells. As individual constructs MS1-PG005 and POB2-PG006 were bombardment into separate cells as negative controls. All samples and controls are displayed in Table 5.4 and Table 5.5.

Table 5.4 Gene constructs in gateway compatible vectors transiently expressed in onion epidermal cells for BIFC experiments.

GENE	GATEWAY DESTINATION VECTOR (donated by Dr Rupert Fray University of Nottingham)	FEATURES
MS1	PG005	pG005 Spectinomycin, Cm BIFC gateway binary vector (Hygromycin), 35S promoter, N-terminal) YFP fragment, 227ng/ul
POB2	PG006	pG006 spec, Cm BIFC gateway binary vector (Hygromycin) Hygromycin, 35S promoter-, C-terminal YFP fragment, 200ng/ul

Table 5.5 Onion epidermal strips were bombarded with the following MS1-PG005 (N-YFP) + POB2-PG006 (C-YFP) plasmid DNA and negative controls.

PLASMID DNA PREPARATION	EXPECTED OUTCOME
MS1-PG005 (N-YFP) + POB2-PG006 (C-YFP)	Induction
MS1-PG005 (N-YFP)	Negative control
POB2-PG006 (C-YFP)	Negative control
Gold particles only	Negative control

Transient expression was carried out via bombardment of DNA coated gold particles (AlfaAesar; sphere 0.8 1.5 μm) into onion epidermal cells using the Bio-Rad Particle Delivery System (PDS) He 1000. The gold particles (30 mg) were first washed with 1 ml of ethanol and three times with SDW in a 1.5 ml Low-Bind eppendorf centrifuge tube before suspension in 0.5 ml of SDW. The gold solution was then sonicated for 1 min in a water-bath sonicator (Fisher) and dispensed in 30 μl aliquots into the Low-Bind centrifuge tubes. The plasmid DNA constructs (Table 5.5) were coated onto the gold particles using the $CaCl_2$/spermidine method with slight modifications as outlined in the manual (Bio-Rad). The gold solution was sonicated for 10s and vortexed vigorously for 1 min, then 25 μl of plasmid DNA (100 ng/μl) was immediately added to the gold solution and

followed by addition of 50 μl of 2.5M $CaCl^2$ and 20 μl of 0.1 M spermidine (freshly prepared). The tube was vortexed for a further 1 min then placed on ice for 1-3 min. The gold solution was then pelleted by a 1 sec pulse centrifugation and the supernatant was removed gently without disturbing the gold pellet. Ethanol (150 μl) was gently laid on top of the gold pellet and the tube was again pulse centrifuged for 1 sec. The gold pellet was washed with 150 μl of ethanol three times and the tube was again centrifuged for 1 sec and suspended in 30 μl of ethanol. The tube was then vortexed for 2-3 sec and the gold fully dispersed by pipetting. The gold solution was then loaded to three micro-carriers (Bio-Rad) and dried in a Petri dish for about 20 min in a $30^{°}$C oven before bombardment into the onion epidermal cells. The inner epidermal cell layer of the onion (locally purchased) was peeled and placed with the inner side down onto Petri dishes containing solid MS agar medium supplemented with 1% (w/v) sucrose). Water was sprayed onto the Petri dishes to keep the onion peel moist. Any excessive water on the onion peels was removed by gently blotting with a piece of dry tissue. The gold particles coated with plasmid DNA (Table 5.4) were then bombarded into the onion cell under 26-28 cm Hg vacuum using 900-1000 psi rupture disks. Petri dishes were then sealed with parafilm and covered with foil and incubated at 28°C for 24-48 hours before scoring the transformed cells for YFP fluorescence.

5.3 RESULTS

5.3.1 YEAST TWO HYBRID ANALYSIS

A previous yeast 2 hybrid screen had identified POB2 as a possible interacting protein of MS1; the interaction of MS1 with POB2, and any possible interaction with the closely related protein POB1, was subsequently tested by further pair-wise yeast two hybrid analyses using the ProQuest Yeast-2-Hybrid (Invitrogen) system (Figure 5.2 and 5.3). The full length *MS1* cDNA without PHD finger motif and the full-length POB2 cDNA were generated as *att*B-PCR products for BP recombination cloning into the pDONR™ vectors to create Entry clones for yeast 2 hybrid analysis (Table 5.6 and Figure 5.6).

Table 5.6 PCR conditions and primer details for generation of attB-MS1 and attB-POB2 products for BP recombination cloning into the pDONR™ vectors to create entry clones.

CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMPERATURE (˚C)	PRODUCT SIZE (bp)
attB-MS1	MS1-Start	4R-noPHD-Y2H	56	1700
attB-POB2	19-1_1-xhoF	19-1_end-kpn1R	56	2200
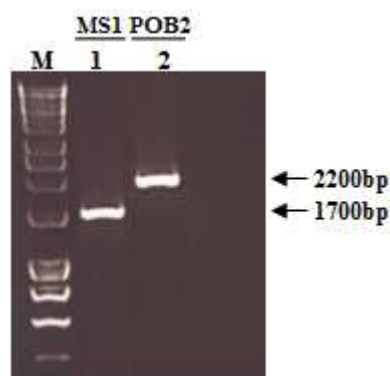
Figure 5.6 attB-MS1 and attB-POB2 PCR products used as substrates for BP recombination cloning into pDONR™ vectors to create Entry clones. The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. Lane 1 is MS1 and lane 2 is POB2. M= molecular weight marker.

The BP recombination reaction *att*B-PCR products were cloned into the pDONR™221 Entry Vector (Section 2.8.9; Figure 5.7). Positive MS1pDONR221 and POB2pDONR221 entry clones were checked by plasmid DNA PCR analysis using vector and gene specific primers (Table 5.7 and Figure 5.8). MS1pDONR221 and POB2pDONR221 plamid DNA constructs (Figure 5.6) were then purified and transferred into the Destination Vectors pDEST 32 and pDEST 22 respectively by LR Gateway recombination reactions (Figure 5.9). DB-MS1-pDEST™32, AD-MS1-pDEST™22, AD-POB2-pDEST™22, DB-POB1pDEST32 purified plasmid DNA is shown in Figure 5.10. DB-POB1pDEST32 was provided by Dr Caiyun Yang (University of Nottingham). Plasmid DNA gene constructs were validated as positive expression clones by PCR analysis using vector and gene specific primers (Table 5.8 and Figure 5.11).

Figure 5.7 Map and Features of pDONR™221 Entry Vector into which the MS1-PHD cDNA (without the PHD finger motif) and POB2 *att*B-PCR products were cloned using BP Clonase™.

Table 5.7 PCR analysis to screen transformants carrying the MS1pDONR221 and POB2pDONR221 gene constructs in *E. coli* DH5a cells by using vector and gene specific primer pairs.

GENE CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMPERATURE ˚C	APPROXIMATE PRODUCT SIZE (bp)
MS1pDONR221	M13F	RT-R	55	630
POB2pDONR22-1	M13F	SALK145_146	55°c	650
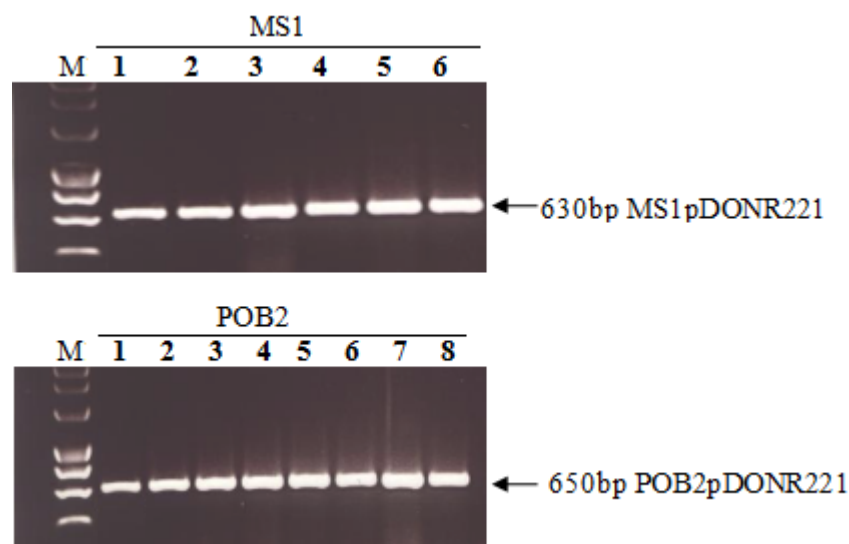
Figure 5.8 Plasmid DNA PCR from colonies of BP Gateway recombination reaction**.** Primers used, anealing temperature and approximate product size are displayed in **Table 5.7**. The Gel pictures show MS1pDONR221 and POB2pDONR221 PCR analysis from purified plamid DNA used for LR Gateway recombination reaction. MS1pDONR221 and POB2pDONR221 transformants were 630bp and 650bp respectively using the primers in Table 5.7. The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.
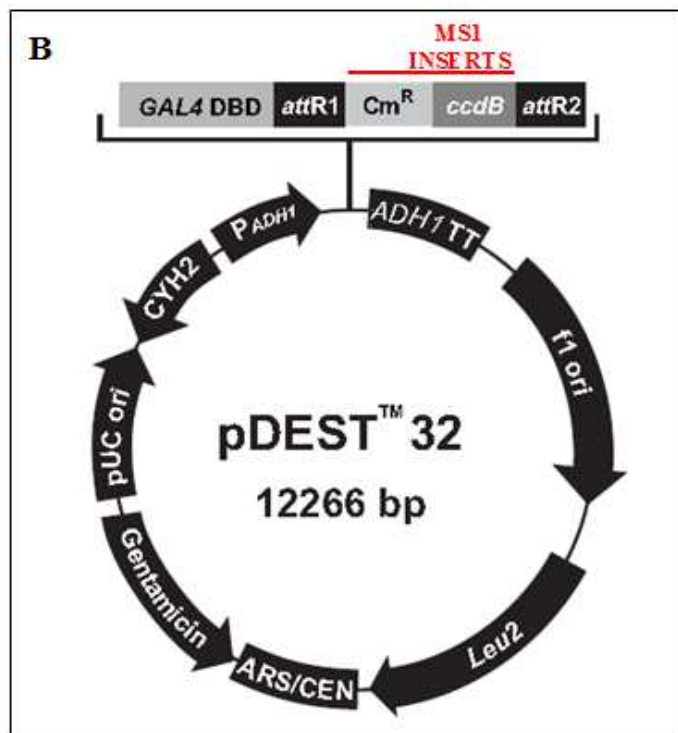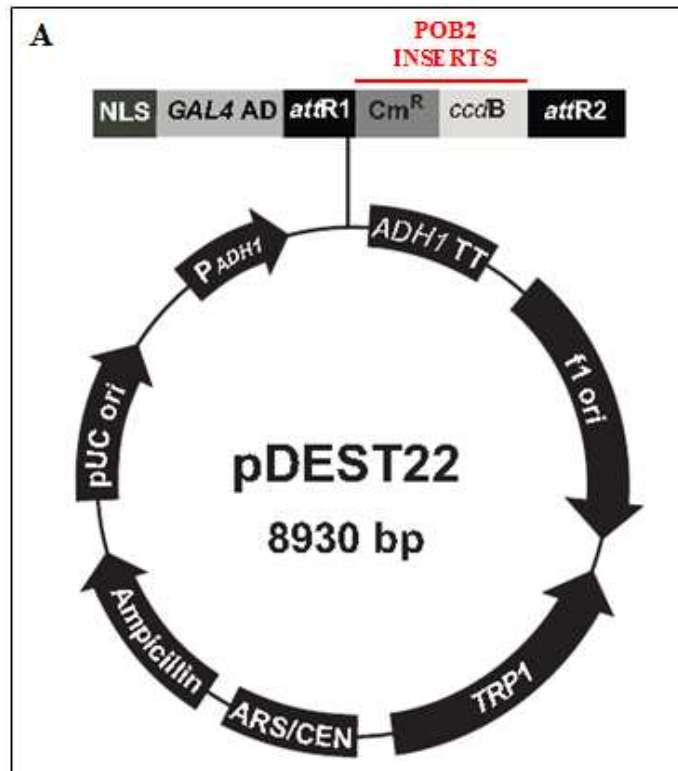
Figure 5.9 (A-B) Map and Features of pDEST™22 and pDEST™32 destination vectors into which the MS1pDONR221 and POB2pDONR221 plamid DNA were cloned using LR Clonase™.
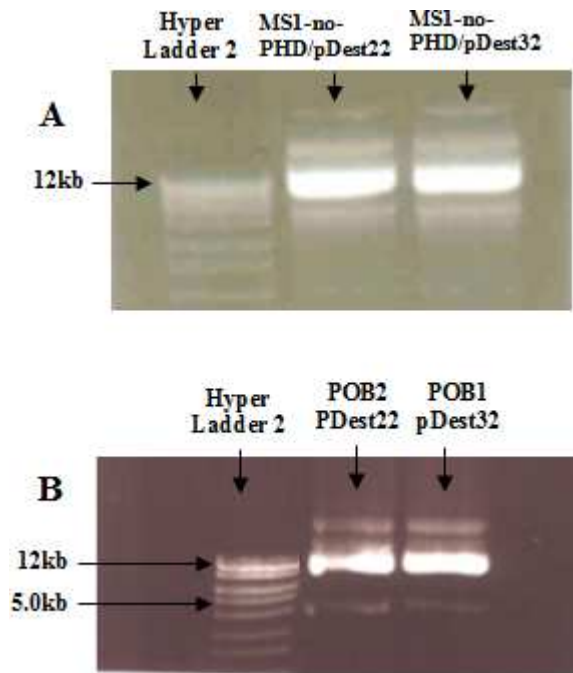
Figure 5.10: Gel pictures showing **(A-B)** MS1pDEST™22, MS1pDEST™32, POB2pDEST™22 and POB1pDEST™32 plamid DNA co-transformed into MaV203 yeast competent cells for checking the induction of reporter genes. The plamid DNA was separated on a 0.8% (w/v) agarose gel with the Hyperladder 2 (Bio-line.com) as molecular weight marker.

Table 5.8: PCR analysis to validate transformants carrying the MS1pDEST™32, and POB2pDEST™22 plamid DNA gene constructs in *E. Coli* DH5a cells by using appropriate vector and gene specific primer pair (Figure 5.11). POB1pDEST™32 and MS1pDEST™22 construct was already confirmed by Caiyun yang (UoN) when provided and therefore no plasmid DNA PCR analysis was performed on them.

PLAMID DNA CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMPERAT-URE	PRODUCT SIZE
MS1pDEST™32	3L-L2-Y2H	DBR	55°C	1500bp
POB2pDEST™22	AT2G46260_725F	ADR	55°C	1100bp
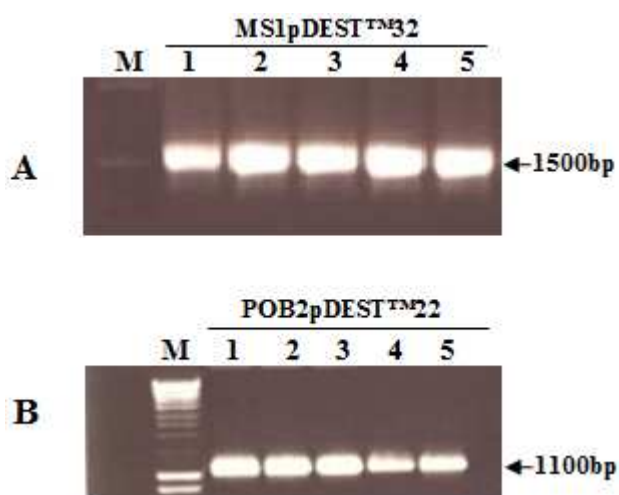
Figure 5.11 Gel picture showing PCR analysis to validate transformants carrying the (**A**) MS1pDEST™32 (**1-5**), and (**B**) POB2pDEST™22 (**1-5**) plamid DNA gene constructs in *E. coli* DH5a cells by using appropriate vector and gene specific primer pair (Table 5.8). The plamid DNA was separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

When a selectable marker such as *HIS3* is used as a reporter gene, Yeast-2-Hybrid dependent transcription activation can be monitored by growth of cells on plates lacking histidine, thereby providing a means to detect protein: protein interactions (Figure 5.12). DB-MS1-pDEST™32, POB1pDEST™32 and AD-POB2-pDEST™22 purified plamid DNA was co-transformed into MaV203 yeast competent cells to check the interactions and induction of reporter genes (Table 5.9; Figure 5.13); they were also screened on SC selection plates to test for self-activation. The MS1 cDNA utilised for yeast 2-hybrid analyses had no PHD finger motif as the PHD finger motif has been previously found to cause self-activation (Yang and Wilson, unpublished data). If POB2 or POB1 self-activates, colonies would grow on Sc-Leu-Trp plates without histidine and activation of the expression of the *HISTIDINE3 AND GALACTOSIDASE1* reporter genes would occur. No interaction between POB1 and MS1 was seen and therefore no expression of the *HISTIDINE3 AND GALACTOSIDASE1* reporter genes was observed. However POB2 interaction with MS1 in yeast MAV203 cells was able to activate the expression of the *HISTIDINE3 AND GALACTOSIDASE1* reporter genes (Figure 5.13). The reporter gene expression levels from the yeast-2-hybrid screen can vary from strong to quite weak (although these levels may not reflect the affinity of the protein: protein interaction observed in the native environment). To help determine the strength of the interactions and which candidate clones represented true interactors, a collection of control strains (derivatives of MaV103) that contain plasmid pairs expressing fusion proteins with a spectrum of interaction strengths was used alongside. The plasmids present in each control strain and the extent of cDNA contained are described in Table 5.10.
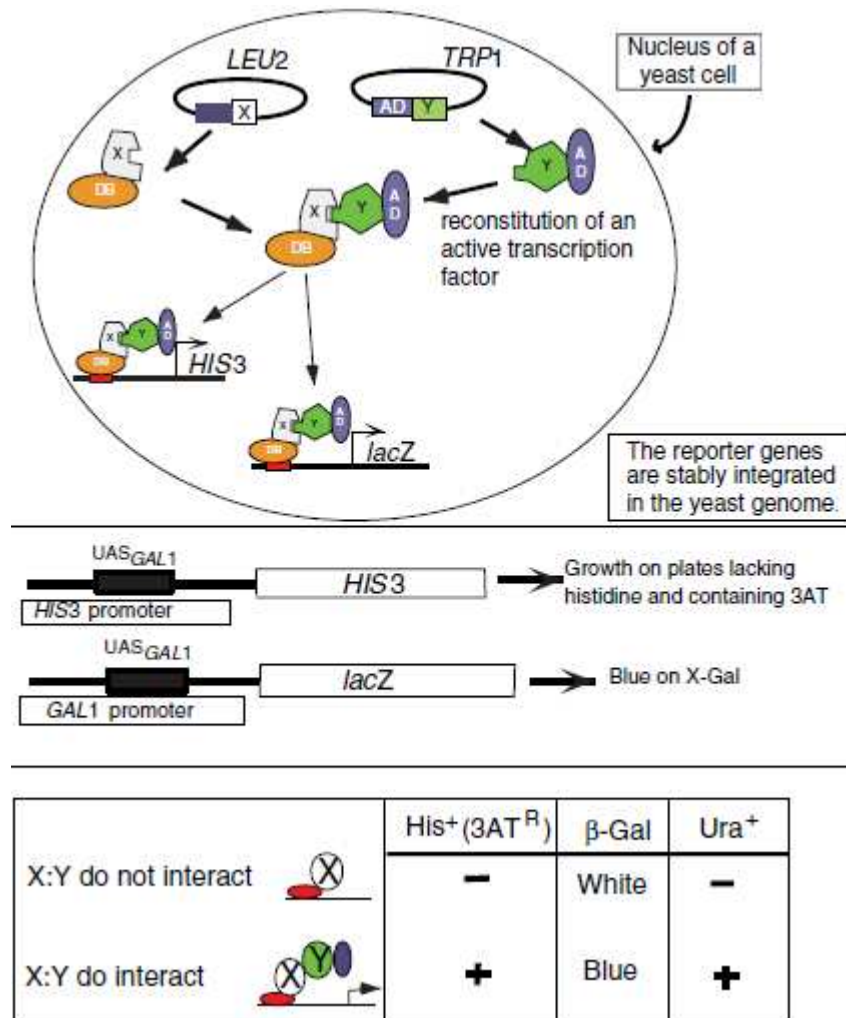
Figure 5.12 An illustration showing how yeast-2-hybrid screening with two reporter genes (*HISTIDINE3 AND GALACTOSIDASE1*) was interpreted. **Top**: MaV203 cells containing DB-S1 and AD-POB2 plasmids encode fusion proteins DB-MS1 and AD-POB2, respectively. The interaction of MS1:POB2 reconstituted an active transcription factor that binds to the *GAL4* DNA binding sequences present in the promoter regions of the reporter genes and activates transcription. **Centre**: Structure of the promoter regions expressing each of the reporter genes and the expected growth/colour results. **Bottom**: Expected growth or colour results when tested for induction of the reporter genes for interacting and non-interacting DBMS1 and AD-POB2.

Table 5.9 Small-Scale Transformation of MaV203 with DB-MS1, POB1, AD-POB2 and control plasmids pDBLeu and pEXP-AD502 (See Table 5.10 for control strains A-D).

TRANSF-ORMATION	PLASMID 1	PLASMID 2	SELECTION MEDIUM	PURPOSE	RESULT SUMMARY
1	Control plasmid (pDBLeu)	None	SC-Leu-Trp-His+3AT	Transformation control	No expression of the Histidine+ and Beta-Galactosidase reporter genes
2	pDBLeu + pEXP-AD502		SC-Leu-Trp-His+3AT	Self-activation control	No expression of the Histidine+ and Beta-Galactosidase reporter genes
3	DB-POB1-pDEST™32	pEXP-AD502	SC-Leu-Trp-His+3AT	Test Self activation of DB-POB1	No expression of the Histidine+ and Beta-Galactosidase reporter genes
4	pDBLeu	AD-POB2-pDEST™22	SC-Leu-Trp-His+3AT	self-activation of AD-POB2	No expression of the Histidine+ and Beta-Galactosidase reporter genes
5	DB-MS1-pDEST™32	AD-POB2-pDEST™22	SC-Leu-Trp-His+3AT	Interaction test for MS1 and POB2	POB2 interaction with MS1 in yeast MAV203 cells was able to activate the expression of the Histidine+ and Beta-Galactosidase reporter genes
6	DB-POB1pDEST™32	AD-POB2-pDEST™22	SC-Leu-Trp-His+3AT	Interaction test for POB1 and POB2	No expression of the Histidine and Beta-Galactosidase reporter genes therefore POB2 does not interact with POB1 in yeast MAV203 cells.

Table 5.10 The plasmids present in each control strain and the extent of cDNA contained. The plasmid vectors used are pPC97 (*GAL4*-DB, *LEU*2), pPC97-*CYH*2S and pPC86 (*GAL4*-AD, *TRP*1) pDBLeu is derived from pPC97. (For more information refer to ProQuest™ Two-Hybrid System with Gateway® Technology).

CONTROL STRAIN	RESIDENT PLASMIDS	cDNA INSERT	INTERACTION STRENGTH	REFERENCE
7. Control A	pPC97-RB	no insert	None	Chevray and Nathans, (1992)
	pPC86-E2F1	no insert		
8. Control B	pPC97	human RB Account M28419 amino acids 302-928	Weak	Vidal, (1997); Vidal *et al*., (1996)
	pPC86	human E2F1 Account M96577 amino acids 342-437		
9. Control C	pPC97-*CYH*2-dDP pPC86-dE2F	*Drosophila* DP Account X79708 amino acids 1-377	Moderately strong	Du *et al*., (1996)
		Drosophila E2F Account U10184 amino acids 225-433		
10. Control D	pPC97-Fos	rat cFos Account X06769 amino acids 132-211	Very strong	Chevray and Nathans, (1992)
	pPC86-Jun	mouse cJun AcccountX12761 amino acids 250-325		
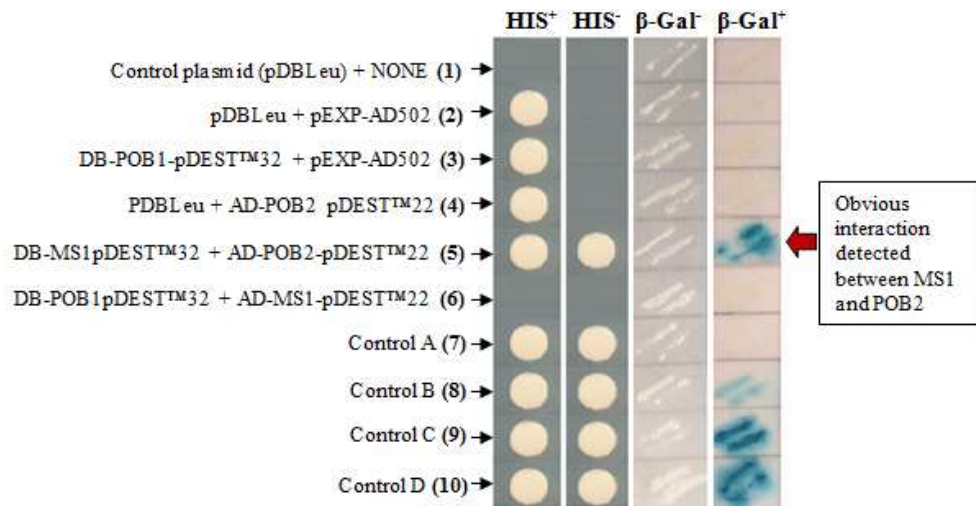
Figure 5.13 Screening candidate interactors by yeast 2 hybrid analysis for the induction of reporter genes. (1) Isolated colonies of Transformation control, (2-4) self-activation tests, (5-6) interacting candidate genes and (7-10) Yeast Control Strains. POB2 interaction with MS1 in yeast MAV203 cells was able to activate the expression of the Histidine (+20mM 3AT) and Beta-Galactosidase reporter genes. However there was no interaction observed between POB1 and MS1 as no expression of the Beta-Galactosidase reporter gene was seen by X-gal assays.

5.3.2 TRANSIENT EXPRESSIONS OF MS1 AND POB2 GENE CONSTRUCTS IN *N. BENTHAMANIA*

The full length cDNA of MS1 and POB2 were generated as *att*B-PCR products for BP recombination cloning into the pDONR™ vectors to create Entry clones (Table 5.11 and Figure 5.14). By BP recombination reactions the *att*B-PCR products were cloned into the pDONR™221 Entry Vector (Section 2.8.9; Figure 5.15).

Table 5.11 attB-MS1 and attB-POB2 PCR products were generated for BP recombination cloning into the pDONR™ vectors to create entry clones.

| CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMPERA- TURE ˚C	PRODUCT SIZE bp
attB-MS1	MS1-Start	Complete gene reverse	56	2194
attB-POB2	19-1_1- xhoF	19-1_end- kpn1R	56	2200
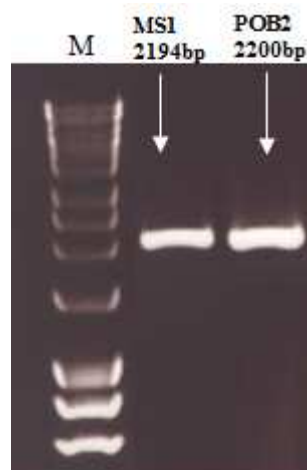
Figure 5.14 attB-MS1 and attB-POB2 PCR products used as substrates for BP recombination cloning into the pDONR™ vectors to create Entry clones. The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

Figure 5.15 Gel picture showing MS1pDONR221 and POB2pDONR221 plamid DNA purified and used for LR Gateway recombination reaction. The PCR products were separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

Positive MS1pDONR221 and POB2pDONR221 entry clones were checked by plasmid DNA PCR analysis using vector and gene specific primers (Table 5.12 and Figure 5.16). MS1pDONR221 and POB2pDONR221 plamid DNA constructs were then purified and transferred into the Destination Vectors PGWB5 and PGWB8 respectively by LR Gateway recombination reactions (Section 2.8.9) (Table 5.13). MS1PGWB5 and POB2PGWB8 plasmid DNA (Figure 5.17) were validated as expression clones by PCR analysis using vector and gene specific primers (Table 5.14 and Figure 5.18).

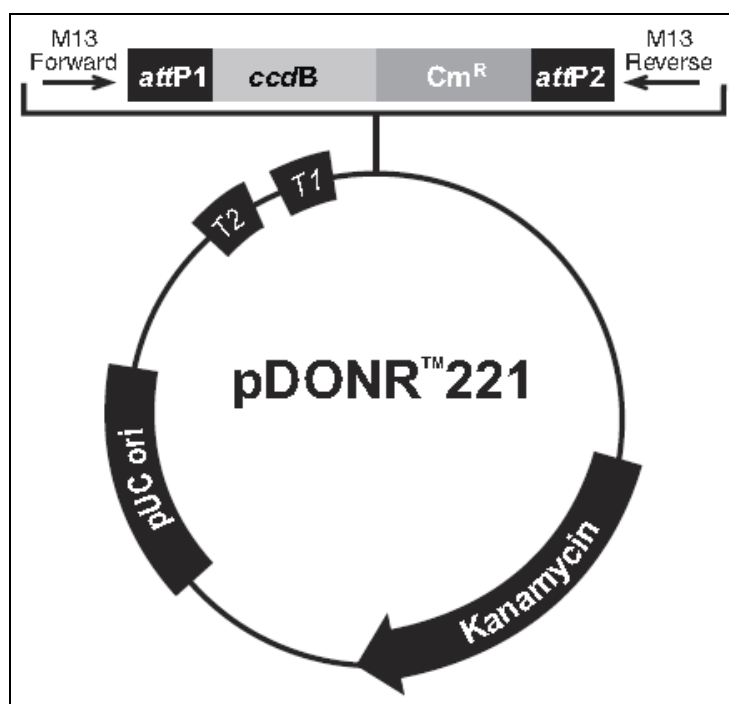**Table 5.12** PCR analysis to screen transformants carrying the MS1pDONR221 and POB2pDONR221 gene constructs in *E. Coli* DH5a cells by using appropriate vector and gene specific primer pair (Figure 5.16 for gel picture).

CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMP. ˚C	APPROXIMATE PRODUCT SIZE bp
MS1pDONR221	M13F	RT-R	55	630
POB2pDONR221	M13F	SALK145_146	55ºc	650
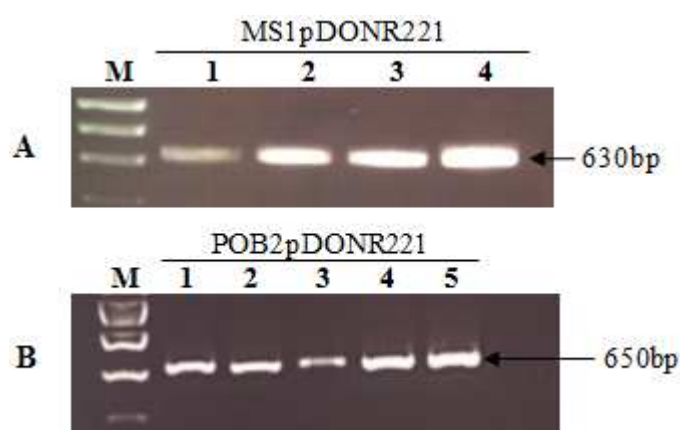
Figure 5.16 Plasmid DNA PCR from colonies of BP Gateway recombination reaction**.** Primers used, anealing temperature and approximate product size are displayed in **Table 5.12**. The gel pictures show (**A**) MS1pDONR221 and (**B**) POB2pDONR221 PCR analysis from plamid DNA purified and used for LR Gateway recombination reaction. MS1pDONR221 and POB2pDONR221 transformants were 630bp and 650bp respectively using the primers in Table 5.12. The PCR products were separated on a 1% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

Table 5.13 Gene constructs in Gateway compatible vectors utilised for transient expression experiments. For more information refer to Gateway instruction manual (Invitrogen).

GENE	GATEWAY DESTINATION VECTOR	FEATURES
MS1	PGWB5	(--35S promoter–R1- CmR-ccdB-R2–sGFP--)
POB2	PGWB8	(--35S promoter-R1- CmR-ccdB-R2-6xHis--)
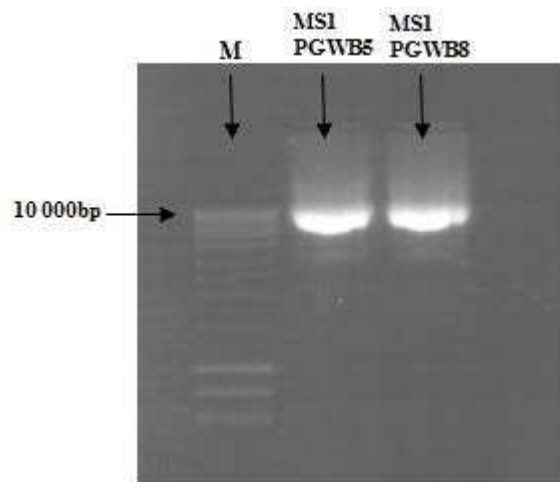
Figure 5.17 Gel pictures showing MS1PGWB5 and POB2PGWB8 plamid DNA purified and co-transformed into GV3101 competent cells for transient expression into *N. benthamania* plants. The plamid DNA were separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

Table 5.14: PCR analysis to validate transformants carrying the MS1PGWB5 and POB2PGWB8 plamid DNA gene constructs using appropriate vector and gene specific primer pair (**Figure 5.19**).

CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMP. ˚C	APPROXIMATE PRODUCT SIZE bp
MS1PGWB5	PGWB8F	R1451	55°C	1500bp
POB2PGWB8	PGWB8F	SALK145_146R	55°C	1100bp
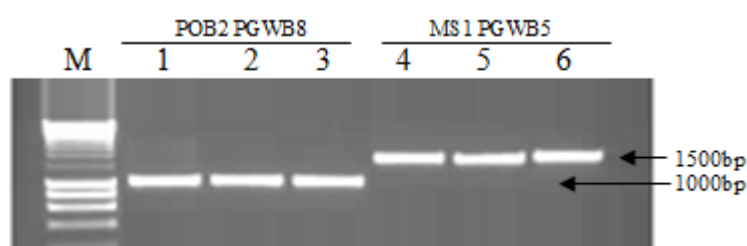
Figure 5.18 Gel picture showing PCR analysis to validate transformants carrying MS1PGWB5 (1-3) and POB2PGWB8 (4-5) plamid DNA gene constructs in *E. coli* DH5a cells using appropriate vector and gene specific primer pairs (Table 5.14) DNA was separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

MS1-PGWB5 and POB2-PGWB8 plamid DNA was purified and transformed into *Agrobacterium* strain GV3101. The *Agrobacterium* carrying expression constructs were checked by PCR using a gene and vector specific primer to make sure they were carrying the correct gene constructs (Table 5.15; Figure 5.19). Validated clones were transformed into *N. benthamania* plants by infiltrating (Section

5.2.2.2) the abaxial side of *N. benthamania* 6 week old leaves with the *Agrobacterium* (GV3101) strains carrying PGWB5-MS1 (GFP tagged) and PGWB8-POB2 (HIS tagged) gene constructs (Figure 5.20).

Table 5.15: PCR analysis to validate transformants carrying the MS1PGWB5 and POB2PGWB8 plamid DNA gene constructs in GV3101 *Agrobacterium* cells by using appropriate vector and gene specific primer pair (Refer to Figure 5.19 for gel picture).

CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMP. ˚C	APPROXIMATE PRODUCT SIZE bp
MS1PGWB5	PGWB8F	R1451	$55^{O}C$	1500bp
POB2PGWB8	PGWB8F	SALK145_146R	$55^{O}C$	1100bp
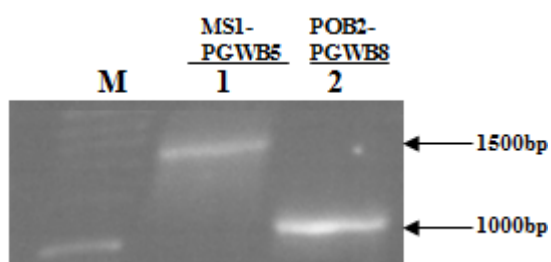
Figure 5.19 Gel picture showing PCR analysis to validate transformants carrying plasmid constructs of MS1PGWB5 (1) and POB2PGWB8 (2) in GV3101 *Agrobacterium* cells by using appropriate vector and gene specific primer pair (Table 5.15). The plamid DNA were separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.
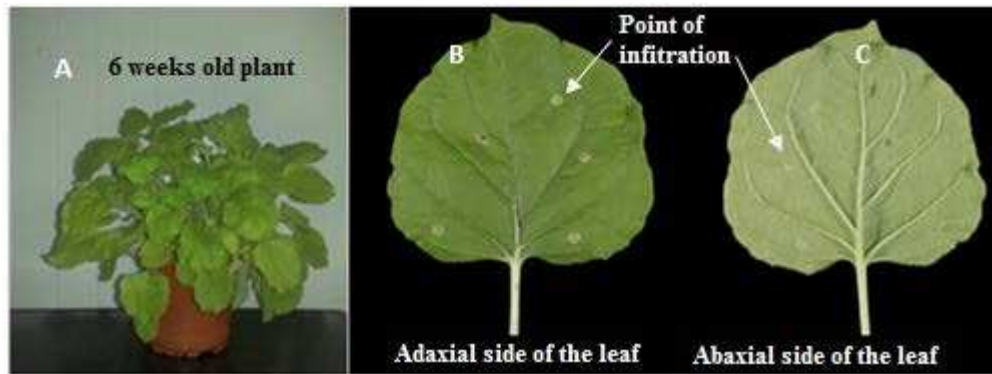
Figure 5.20 Infiltration transformation of *N. benthamiana* leaves. **(A)** Plants grown for 6 weeks prior to infiltration. The abaxial side of the leaf was used for microinjecting PGWB5-MS1 and PGWB8-POB2 gene constructs. **(B)** Adaxial side with points of injection infected with *Agrobacterium* (see arrow) and **(C)** Abaxial side of leaves after infiltration with infected tissue on point of injection (see arrow) (For detailed protocol refer to section 5.2.2).

Total protein was isolated (Chapter 2 Section 2.9.1) from the infiltrated *N. benthamiana* leaf tissues and immunodetection of gene constructs was conducted (Chapter 2 Section 2.9.6). The protein was transferred to a PVDF membrane and hybridised with antibodies to the tags carried on the constructs. The results showed a very faint MS1-GFP band and a GFP positive control band (MAP4) (Figure 5.21). During our study using this procedure proved challenging specifically with immunodetection of GFP and HIS tag fused to MS1 and POB2 respectively. No POB2/MS1 protein complex was detected when the membrane was incubated consequently this approach could not be used to confirm the interaction of MS1 with POB2. However 2 more different approaches were employed to validate MS1 interaction with POB2 (See section 5.3.3 and 5.3.4)
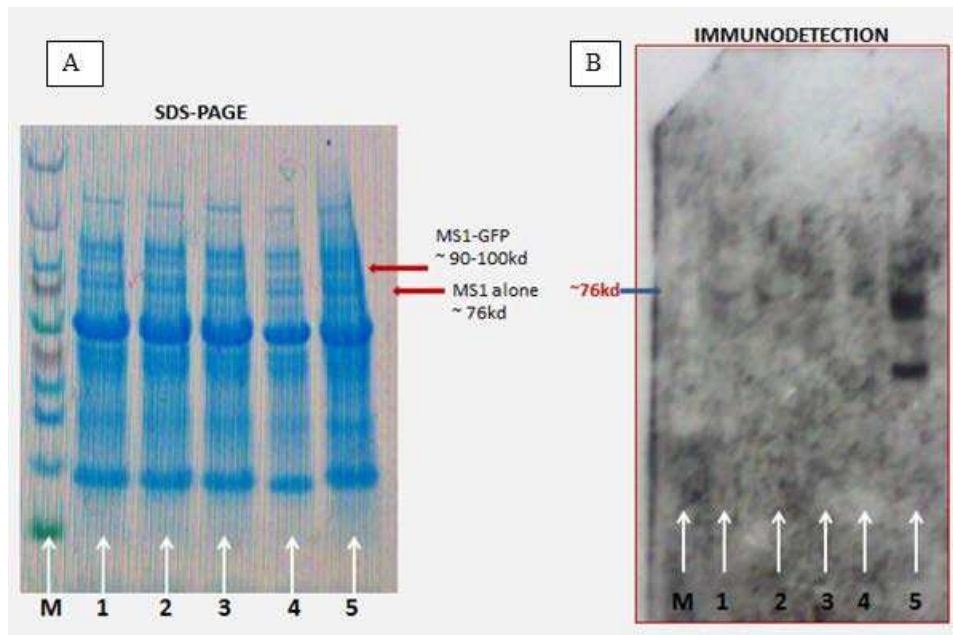
Figure 5.21 (A) SDS-PAGE (4-20% gel) run showing native total protein extraction in lanes (1-4) from infiltrated *N. benthamiana* leaf tissues. Lane 5 represents native total protein extraction from *Arabidopsis* plants carrying the Microtubule-associated protein 4 (MAP4). M represents the SDS-PAGE band profile of the Thermo Scientific Spectra Multicolor Broad Range Protein Ladder and subsequent transfer to the membrane. **(B)** Western blot of total protein from Total protein; Immunodetection was carried out using anti-GFP primary antibody from rabbit polyclonal. The exact concentrations of the primary and secondary antibodies used are described in Tables 2.13 and Table 2.14. Lanes 1-4 showed a very faint MS1 band (76kd) and in lane 5 a GFP positive control band (Microtubule-associated protein 4 (MAP4) = 74.8kd), therefore No POB2/MS1 protein complex was detected.

5.3.3 EXPRESSION OF MS1 AND POB2 USING AN *E. COLI* EXPRESSION SYSTEM WITH GATEWAY® TECHNOLOGY

In addition to the above techniques, the interaction between MS1 and POB2 proteins was also studied using *in vitro* expression in the *E. coli* with Gateway Expression system (Invitrogen®) to create *E. coli* expression clones containing the T7 promoter. As previously described the full length *MS1* and *POB2* cDNA pDONR™221 Entry clones (Figure 5.14-16) were used to perform an LR recombination reaction with the appropriate pDEST15™ and pDEST17™ destination vectors, to generate MS1-pDEST15 and POB2-pDEST17 expression clones (Table 5.17, Figure 5.22 and Figure 5.23). Confirmation of the expression clones was performed by plasmid DNA PCR analysis (Figure 5.22 and Figure 5.23). Plasmid DNA constructs carrying MS1-pDEST15 and POB2-pDEST17 was transformed into BL21 (DE3) competent cells for high-level protein expression (Section 5.2.3.2).

Table 5.16 Genes constructs prepared in Gateway compatible vectors for transient expression in *E.coli*. For more information refer to Section 5.2.4 and *E. coli* Expression System with Gateway Technology Invitrogen.

GENE	GATEWAY DESTINATION VECTOR	FEATURES
MS1	pDEST15	T7 promoter, C-terminal-, GST- Tag
POB2	pDEST17	T7 promoter- N-terminal, 6XHis- Tag
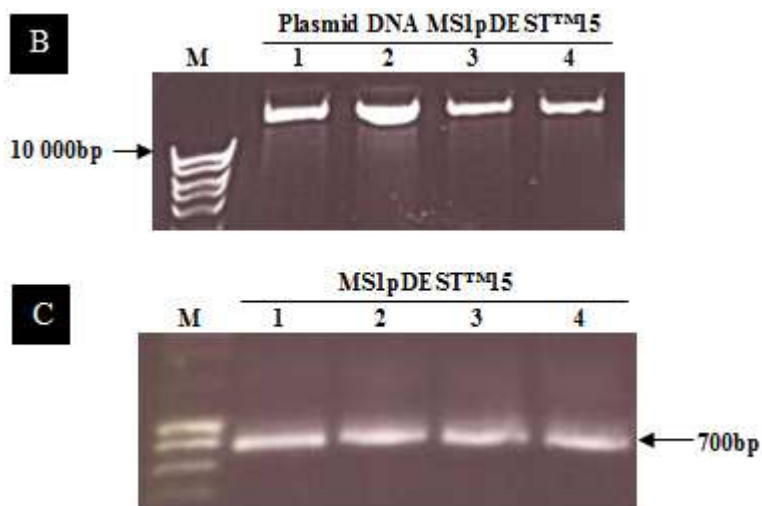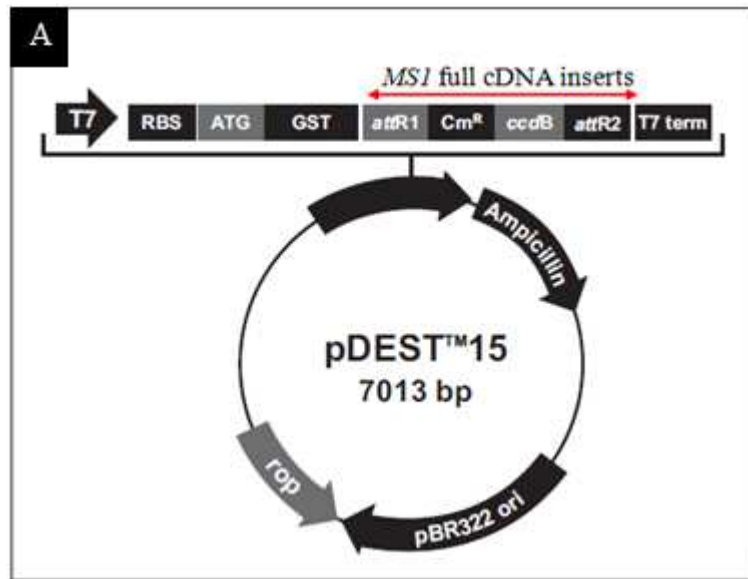
Figure 5.22 (**A**) Map and Features of pDEST™15 destination vector into which the MS1pDONR221 plamid DNA were cloned using LR Clonase™. (**B**) Gel picture showing purified MS1pDEST™15 plasmid DNA (1-4). (**C**) Gel picture showing PCR analysis to validate transformants carrying MS1pDEST™15 (1-4), Forward Primer: T7 primer and Reverse primer: RRT. The plamid DNA and PCR analysis were both separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.
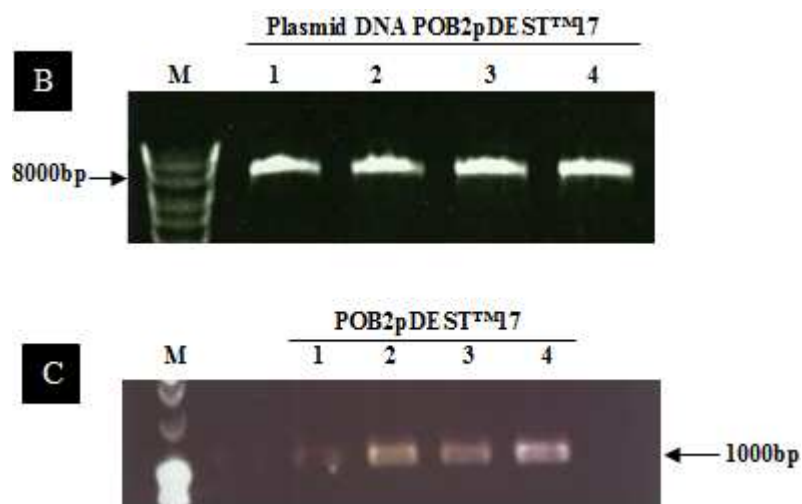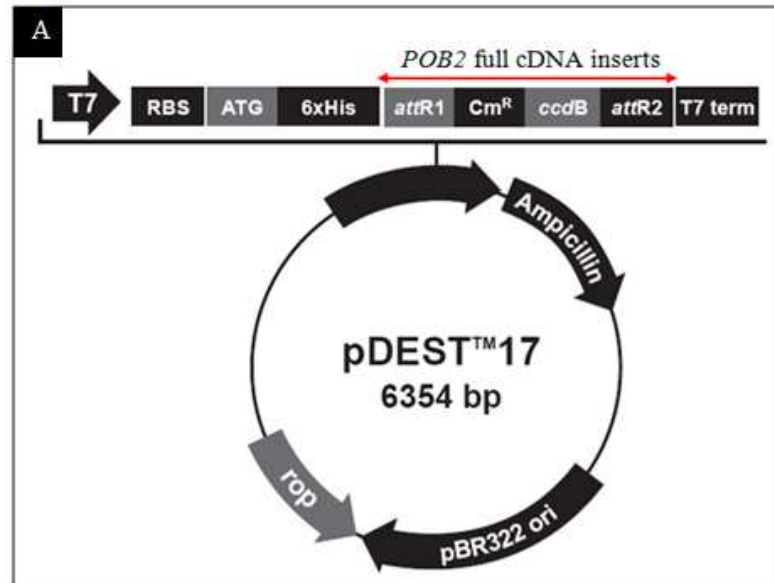
Figure 5.23 (**A**) Map and Features of pDEST™17 destination vector into which the POB2pDONR221 plamid DNA were cloned using LR Clonase™. (**B**) Gel picture showing purified POB2pDEST™17 plasmid DNA (1-4). (C) Gel picture showing PCR analysis to validate transformants carrying POB2pDEST™17 (1-4) Forward Primer: At2g46260-725F Reverse primer: T7_TermR. The plamid DNA and PCR analysis were both separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

The plasmid DNA constructs were also sequenced to confirm that the gene of interest was cloned in frame with the appropriate N-terminal (GST) or C-terminal (6XHIS) tag before induction (Appendices 8.5-8.6). The *MS1* and *POB2* cDNA maps in Figures 5.24 and Figure 5.25 respectively show the primers that were used to sequence the plasmid DNA constructs. The plasmid sequencing results for MS1pDEST15 clone at 3'and 5' end and pDEST17POB2 clone at 5'-3' end aligned with start of POB2 cDNA is shown in Appendices 8.6-8.6. Both constructs were shown to be in frame and correct without any mutations.
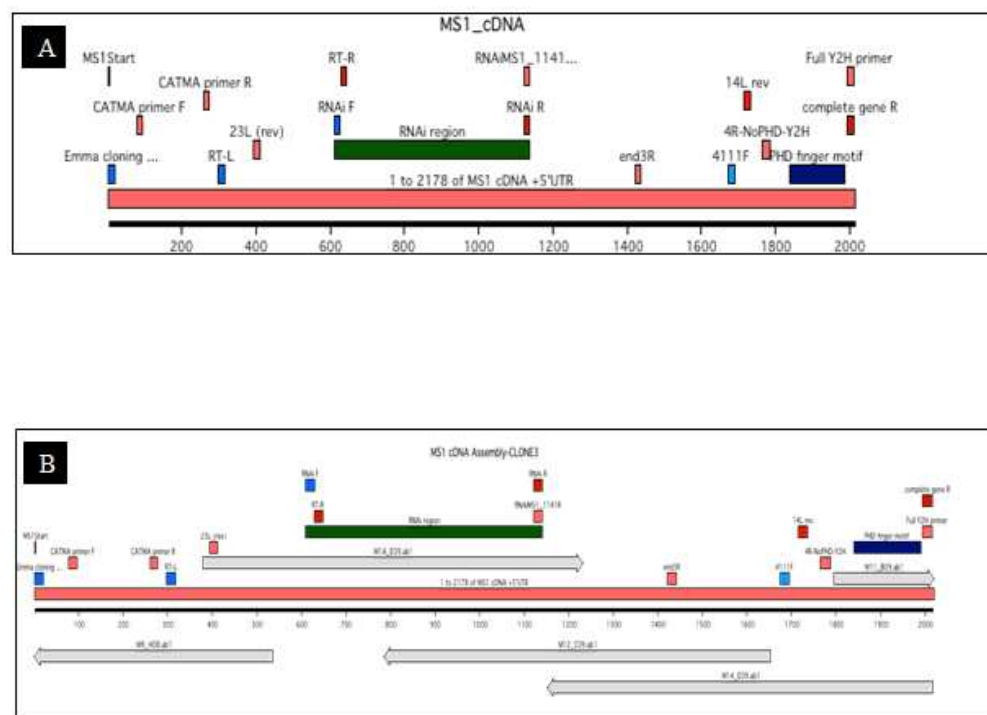
Figure 5.24: (A) MS1 cDNA showing primers utilised for sequencing the pDEST15-MS1 gene constructs. (**B**) The grey arrows show overlaping sequenced regions from 5'to 3' end. Sequencing fomatted alignment of pDEST15-MS1 clone with MS1 cDNA is shown Appendices 8.5.
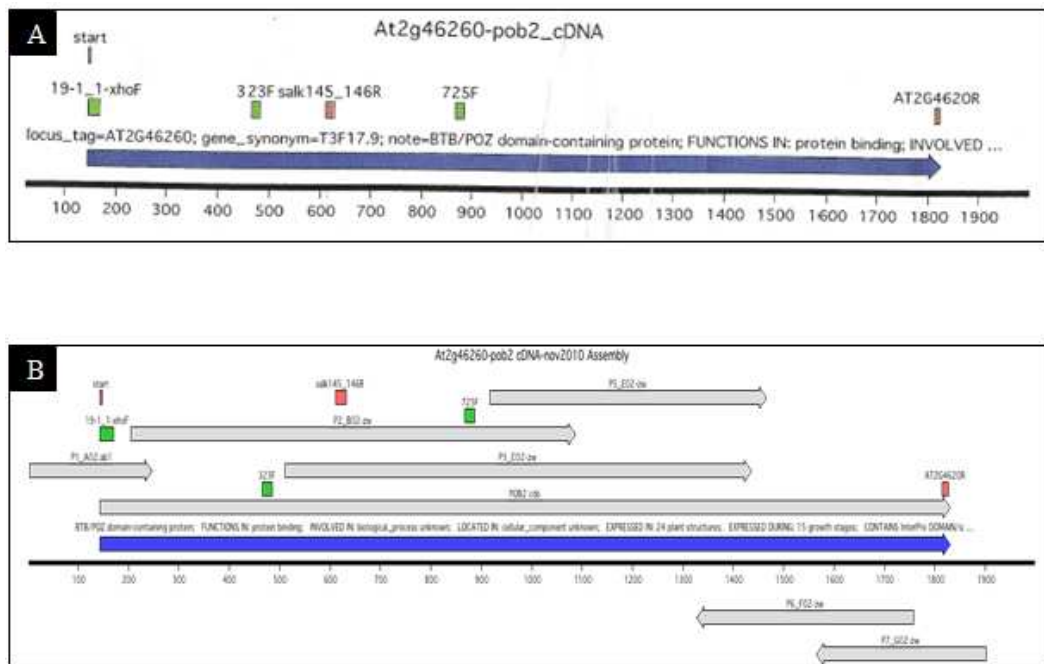
Figure 5.25: POB2 (At2g46260) cDNA showing primers utilised for sequencing the pDEST17-POB2 gene constructs. The grey arrows show overlaping sequenced regions from the 5'to 3' end. Sequencing fomatted alignment of pDEST17-POB2 clone with POB2 cDNA is shown in Appendices 8.6.

To induce MS1 and POB2 protein expression in BL21 (DE3) host strains, L-arabinose was added to the culture to a final concentration of 0.2%. Induction samples of both clones in BL21 (DE3) were collected at 1 hour interval for 6 hours and overnight. Control non-induced culture samples were also collected. The induction of expression of MS1 and POB2 was checked by gel electrophoresis of total protein and staining with Coomassie® Brilliant Blue (Figure 5.26) (for detailed protocols refer to www.invitrogen.com). Overexpression of the recombinant proteins in BL21 (DE3) competent cells with L-Arabinose induction could not be achieved as no induction of a band at the appropriate size for MS1 or POB2 was seen, there were also no difference in other

band intensities between the non-induced and induced cultures (Figure 5.26). Therefore an alternative approach was used to validate the protein-protein interactions (Section 5.3.4).
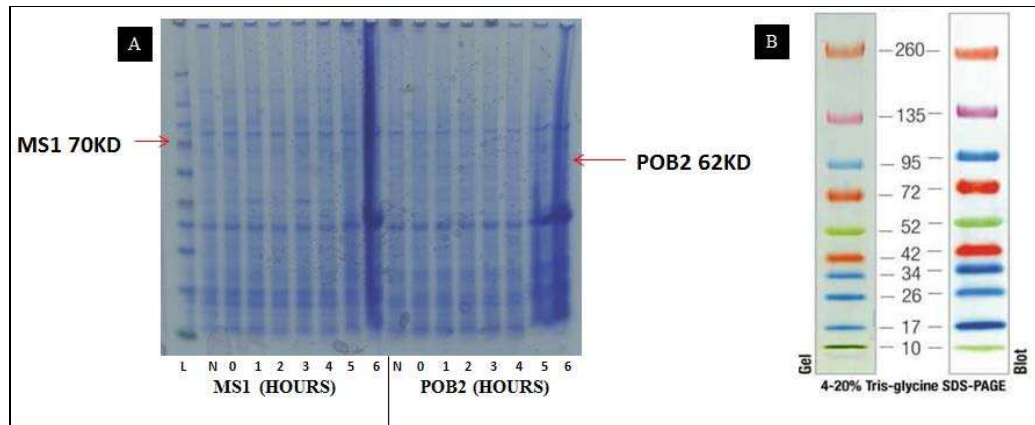
Figure 5.26: SDS-PAGE analysis of MS1 and POB2 protein expression levels from total protein isolated from *E.coli* BL21 (DE3) cultures. (A) No induction of a band at the appropriate size for MS1 or POB2 was seen, there were also no difference in other in band intensities between the non-induced and induced cultures. Arrows indicated expected sizes of transgene proteins (B) Thermo Scientific Spectra Multicolor Broad Range Protein Ladder. L= Protein ladder, N= None induced culture.

5.3.4 BIMOLECULAR FLUORESCENT COMPLEMENTATION (BiFC) ANALYSIS

The previously confirmed MS1pDONR221 and POB2pDONR221 entry clones (Figure 5.14-16) were used for LR Gateway recombination reaction (Section 2.8.9) into the Destination split BiFC vectors pG005 (YFP N-terminal region) and

pG006 (YFP C-terminal region) respectively generating MS1-pG005 and POB2-pG006 expression clones (Figure 5.27). The vectors pG005 and pG006 were donated by Dr Rupert Fray (UoN). MS1pG005 and POB2pG006 plasmid DNA were validated as expression clones by PCR analysis using vector and gene specific primers (Table 5.17 and Figure 5.28). MS1-pG005 and POB2-pG006 plamid DNA was purified and bombarded into living onion epidermal cells (Section 5.2.4.2) to test for interaction and reconstitution of the split YFP fluorescent protein
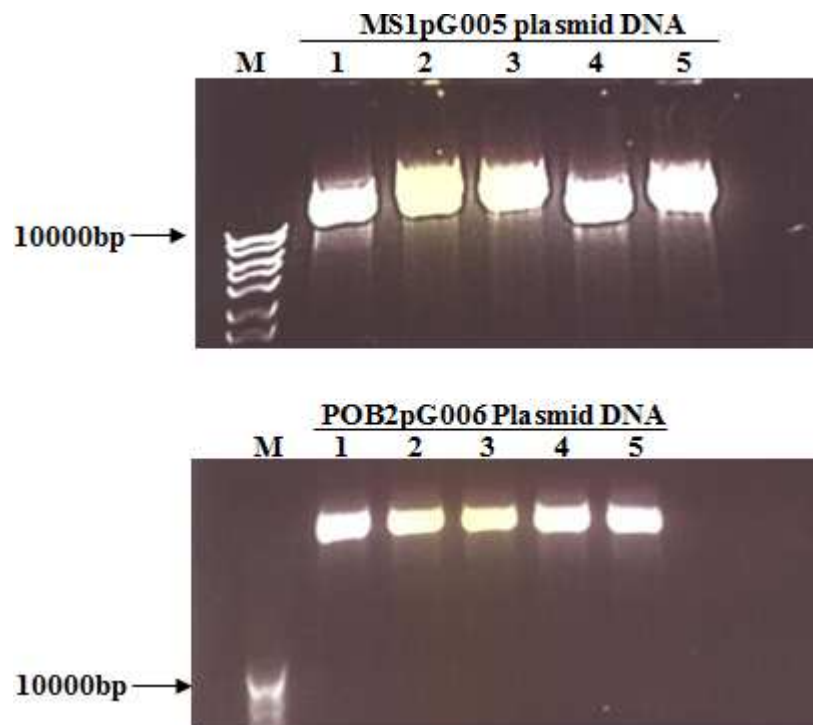
Figure 5.27: Gel pictures showing MS1pG005 and POB2pG006 plamid DNA purified and co-transformed into *E.coli* DH5a competent cells for bombardment into living onion epidermal cells (Table 5.17 and Figure 5.28). The plamid DNA were separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

Table 5.17: PCR analysis to validate transformants carrying the MS1pG005 and POB2pG006 Plasmid DNA gene constructs by using appropriate vector and gene specific primer pair (Figure 5.28).

CONSTRUCT	FOWARD PRIMER	REVERSE PRIMER	ANEALING TEMP. ˚C	APPROXIMATE PRODUCT SIZE bp
MS1pG005	35sF	RRT	52	350
POB2pG006	35sF	SALK_14 5146R	56	500
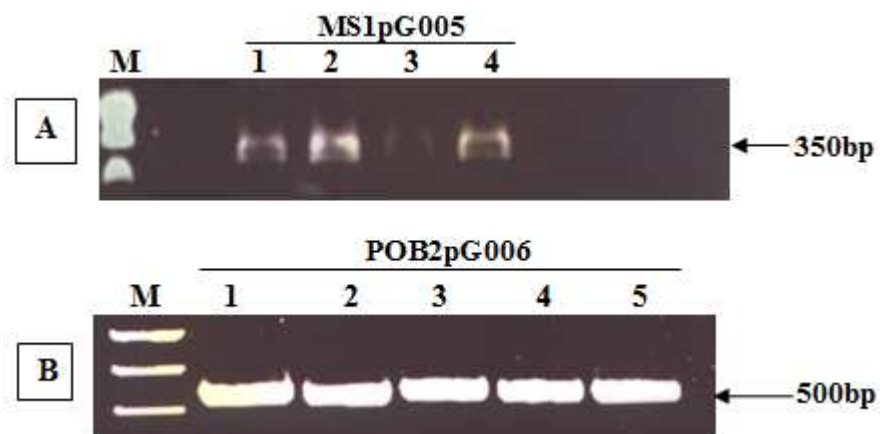
Figure 5.28 Gel picture showing PCR analysis to validate transformants carrying (**A**) MS1pG005 (**B**) POB2pG006 plamid DNA gene constructs by using appropriate vector and gene specific primer pairs (**Table 5.17**). The plasmid DNAs were separated on a 0.8% (w/v) agarose gel with the Hyperladder 1 (Bio-line.com) as molecular weight marker. M= molecular weight marker.

Treatments containing both the MS1-pG005 and POB2-pG006 plasmid gene constructs showed functional nuclear YFP protein expression, indicating a

positive protein interaction between the MS1 and POB2 proteins (Figure 5.29 and Figure 5.30). No fluorescent signal was observed in control samples where only one of the BiFC clones was used for bombardment (Figure 5.29 and Figure 5.30). BiFC therefore enabled direct visualisation of MS1 protein interaction with POB2 protein in living onion epidermal cells. The strong intrinsic fluorescence of the reconstituted reporter protein made it possible to observe the onion epidermal living cells for 48 hours post bombardment.
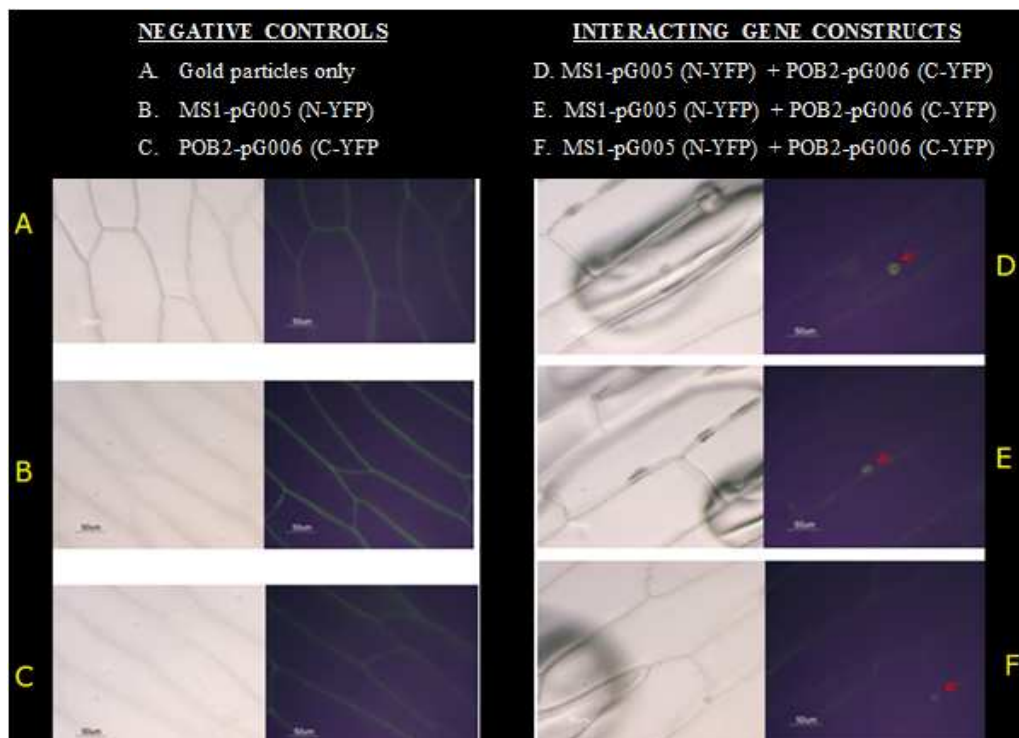
Figure 5.29 Bifluorescent microscopic analyses of onion epidermal cells bombarded with MS1-pG005 and POB2-pG006 plasmid gene constructs. YFP reporter gene induction (red arrow) was observed after reconstitution of a functional YFP protein by interaction of POB2 and MS1 in the nucleus, compared to the control samples containing only the C or N terminal part of the YFP protein that showed no fluorescence.
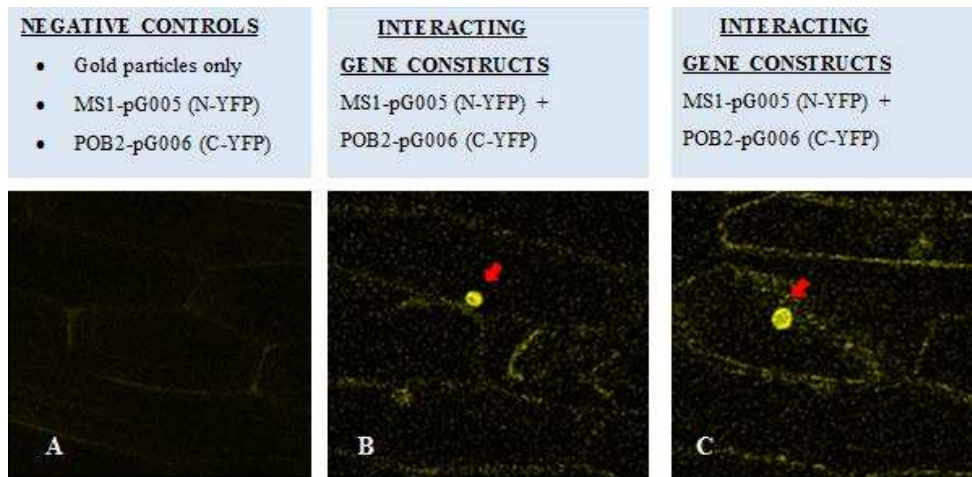
Figure 5.30 Confocal microscopy analyses of onion epidermal cells bombarded with both MS1-pG005 and POB2-pG006 plasmid gene constructs showed YFP reporter gene induction (**B-C** red arrow) by reconstituting a functional YFP protein in the nucleus compared to (**A**) the negative YFP controls.

5.4 DISCUSSION

During the past decade, significant progresses have been made towards understanding *A. thaliana* pollen development at each development stage under the molecular level, yet the regulatory mechanisms that control the dynamic pollen development processes remain largely unknown. To understand MS1 regulatory networks especially the coordinated dynamic regulation between transcription factors (TFs) and their corresponding target genes during pollen development, several approaches have been represented significantly in the genome-wide expression analysis of MS1 regulatory targets. In this work the in-vivo interaction between MS1 and POB2 has been confirmed by Y-2-H and BiFC analyses.

5.4.1 *IN VIVO* INTERACTION OF MS1 WITH POB2

To validate the interaction between MS1 and POB2 4 approaches including Yeast-2-Hybrid System, Transient Expression, Gateway *E. coli* Expression System and BiFC were used. Of these only BiFC and Y2H analysis confirmed the interaction between MS1 and POB2 which could in future provide a fuller understanding of the regulatory role of MS1 during pollen development. This suggests that if POB2 is interacting *in vivo* with MS1, it may be serving to aid in the turnover of the MS1 protein via a proteasome dependant pathway. Therefore POB2 could be playing a fundamental role in the metabolic processes of pollen wall biosynthesis by regulating the break-down of MS1 by ubiquitylation and thus serving to maintain normal tapetal development and the onset of tapetal PCD.

While MS1 is a sporophytic factor controlling anther and pollen development, POB1 and POB2 are both POZ/BTB containing-proteins predicted to be involved in protein ubiquitylation based on physical interaction with CULLIN 3 proteins (Table 5.18). POB1 and POB2 function redundantly as negative regulators of defense against pathogens in *Arabidopsis* (Qu *et al*., 2010). Their BTB domain is a protein–protein interaction motif found throughout eukaryotes. It has been shown also that in Yeast-2-Hybrid assays, POB1 and POB2 dimerize through their BTB domains (Qu *et al*, 2010). In our studies, the Y2H system was performed to check if POB1 and POB2 interact with MS1 by reconstituting an active transcription factor. POB1 contained a DNA-Binding Domain (DB) fused to it known as "bait" and MS1 fused to Activation Domain (AD) also known as the "prey" or "target protein". DB-POB1: AD-MS1 interaction failed to

reconstitute a functional transcription factor that could activate chromosomally-integrated reporter genes driven by *HIS3* and *GAL1* promoters containing the relevant DNA binding sites (Table 5.9, Figure 5.12 and Figure 5.13). Therefore there was no interaction between POB1 and MS1 even though *POB1* and *POB2* (Qu *et al*, 2010) are highly similar genes (Table 5.18). POB1 and POB2 sequences differ within specific regions of the coding sequence and these differences may lead to differing protein conformations between POB1 and POB2, resulting in differences in their interaction partners.   To check the POB2 interaction with MS1, MS1 contained a DNA-Binding Domain (DB) fused to it known as "bait" and POB2 fused to Activation Domain (AD) also known as the "prey" or "target protein". DB-MS1: AD-POB2 interaction reconstituted a functional transcription factor that activated chromosomally-integrated reporter genes driven by *HIS3* and *GAL1* promoters containing the relevant DNA binding sites (Table 5.9, Figure 5.12 and Figure 5.13). POB2 interaction with MS1 in yeast MAV203 cells was able to activate the expression of the Histidine (+20mM 3AT) and Beta-Galactosidase reporter genes. Even though POB1 did not interact with MS1 a number of genes have already been shown to be coexpressed with POB1 (http://atted.jp/data/locus/At3g61600.shtml) (Figure 5.31) and none of these are closely linked with *MS1*.

Bimolecular Fluorescence Complementation (BiFC) technique confirmed MS1/POB2 protein–protein interactions in onion epidermal cells. This approach was based on the formation of a fluorescent complex by two non-fluorescent fragments (N and C terminal ends) of the Yellow Fluorescent Protein (YFP) brought together by the association of interacting proteins fused to these

fragments (MS1 fused to N-YFP terminal end and POB2 fused to C-YFP terminal end). Using BiFC assays, it was possible to confirm that MS1 protein interacts with POB2 by direct visualisation of a reconstituted YFP protein fluorescent under simple and confocal microscope view.

Figure 5.32: Coexpressed gene network around POB1 (At3g61600). (http://atted.jp/data/locus/At3g61600.shtml).

Table 5.18 A summary of MS1, POB1 and POB2 gene description, molecular function, expression pattern during growth and developmental stages.

GENE NAME	DESCRI-PTION	INVOLVED IN	LOCATED IN	MOLECULAR FUNCTION	EXPRESSED DURING (Growth and Developmental Stages	EXPRESS-ED IN (Plant structure)
MS1 At5g22260	Sporophytic factor controlling anther and pollen development. Mutants fail to make functional pollen;pollen degeneration occurs after microspore release and the tapetum also appears abnormally vacuolated. Similar to PHD-finger motif transcription factors.	cellular response to brassinosteroid stimulus, microgametogenesis, pollen germination, pollen wall assembly, positive regulation of transcription, DNA-dependent, regulation of transcription, DNA-dependent, tapetal layer morphogenesis	Mitochondrion, nucleus	DNA binding i.e. sequence-specific DNA binding transcription factor activity, zinc ion binding	Microsporogenesis	Anther Not in microspore
POB1 At3g61600	POZ/BTB containing-protein AtPOB1. Involvement in protein ubiquitylation is predicted based on physical interaction with CULLIN 3 proteins.	floral organ abscission, flower morphogenesis, proximal/distal pattern formation, response to red light	cytoplasm, nucleus	protein binding i.e. protein heterodimerization activity, protein homodimerization activity	anthesis stage, stage, stage, bilateral, expanded, mature embryo stage, mature pollen stage, two leaves visible stage, four leaves visible stage, six leaves visible stage, eight leaves visible stage, ten leaves visible stage, twelve leaves visible stage, petal differentiation and expansion stage	expressed in all tissues of *Arabidopsis*
POB2 At2g46260	POZ/BTB containing-protein involved in protein ubiquitylation is predicted based on physical interaction with CULLIN 3 proteins	floral organ abscission, flower morphogenesis, proximal/distal pattern formation, response to red light	cytoplasm, nucleus	protein binding protein i.e. heterodimerization activity, protein homodimerization activity	carpel, cauline leaf, collective leaf structure, cotyledon, flower, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, pedicel, petal, petiole, plant embryo, plant sperm cell, pollen, root, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf	expressed in all tissues of *Arabidopsis*

5.4.2 OTHER METHODS USED TO ANALYSE PROTEIN-PROTEIN INTERACTIONS

In addition to the approaches utilised to confirm the protein-protein interactions with MS1, there are a number of other methods that can be utilised.

5.4.2.1 CO-IMMUNOPRECIPITATION (CO-IP)

Co-immunoprecipitation (Co-IP) is a popular technique to identify physiologically relevant protein-protein interactions by using target protein-specific antibodies to indirectly capture proteins that are bound to a specific target protein (http://www.piercenet.com). These protein complexes can then be analyzed to identify new binding partners, binding affinities, the kinetics of binding and the function of the target protein. Co-IP works by selecting an antibody that targets a known protein that is believed to be a member of a larger complex of proteins. By targeting this *known* member with an antibody it may become possible to pull the entire protein complex out of solution and thereby identify *unknown* members of the complex (http://www.piercenet.com). This concept of pulling protein complexes out of solution is sometimes referred to as a "pull-down". Proteins that interact in a typical Co-IP are post-translationally modified and conformationally natural and proteins interact in a non-denaturing condition which is almost physiological. However, with Co-IP the signals of low-affinity protein interactions may not be detected and the choice of the appropriate antibody is crucial as the target protein needs to be properly predicted, or there would not be a positive result in Co-IP (http://www.assay-protocol.com/Immunology/Co-IP). An

advantage of this approach is that one can be relatively confident that the antibody directed against the tag is specific and does not cross react with other proteins (Wuchty, 2002). Furthermore, epitope-tagged proteins can often be eluted by incubation with competing peptides, or other small molecules, instead of boiling. Such specific elution often reduces the amount of contaminating proteins in the eluate (Wuchty, 2002). Alternatively, one can perform Co-IP experiments using cells transfected with tagged versions of two putative interaction partners.

5.4.2.2 FLUORESCENCE RESONANCE ENERGY (FRET)

Fluorescence Resonance Energy Transfer (FRET) is a useful technique to measure direct protein-protein interactions and conformational changes (Masi *et al*., 2010). FRET occurs when an excited donor fluorophore transfers energy to an acceptor fluorophore, resulting in reduced donor fluorescence and enhanced acceptor fluorescence (Masi *et al*., 2010). FRET occurs only when the donor/ acceptor pair is in close proximity (1-10 nm). Unlike biochemical methods, FRET allows the quantitative determination of noncovalent molecular associations in living cells (Chan *et al*., 2001). Using conventional flow cytometry, FRET can be quantified on a cell-by-cell basis on large number of cells (Chan *et al*., 2001). While FRET can be quantified by flow cytometry, microscopy is often required to confirm interpretation of results (Chan *et al*., 2001). FRET is relatively cheap, very efficient and rapidly measures changes in distances in a complex of molecules (Moens and dos Remedios, 2001). One of the major obstacles to the widespread implementation of FRET investigations in living cells has been the lack of suitable methods for labelling specific intracellular proteins with the appropriate

fluorophores (Rizzo *et al*., 2009). The development of fluorescent proteins possessing a wide array of spectral profiles and the increasing sophistication of protein chimeras (fusions as well as biosensors) has resulted in a number of potential fluorescent protein pairs that are useful in FRET experiments (Rizzo *et al*., 2009). In addition the precision of the measure is impaired by the uncertainty of the orientation factor and by the size of the probes (Moens and dos Remedios, 2001). Moreover when measuring a change in distance between two probes, the result is a scalar and gives no indications of which probe (donor and/or acceptor) moves and the presence of free labels in solution could mask a change in energy transfer (Moens and dos Remedios, 2001).

CHAPTER 6

FINAL DISCUSSION AND CONCLUSIONS

6. FINAL DISCUSSION AND CONCLUSIONS

6.1 AMS REGULATORY GENE NETWORK FOR SPOROPOLLENIN BIOSYNTHESIS IN THE TAPETUM

The value of *A. thaliana* as a model for plant biology and the analysis of complex organisms have been enhanced by the availability of genome sequence and the application of bioinformatics. In this research, the use of bioinformatics assisted with identification of putative targets of key transcription factors associated with pollen development, namely some of the direct targets of the *ABORTED MICROSPORE (AMS)* transcription factor. Transcriptional analysis of these putative target genes, alongside genotypic analysis of knockout mutants gave clear understanding of the gene function at the molecular level. Phenotypic analysis of the insertional mutants helped predict gene functions as a result of the mutations. Analysis of the putative AMS targets identified defective pollen development in the ABC transporter mutant (WBC27). The ABC transporter protein belongs to the WBC subfamily of *Arabidopsis* ABC transporters, which function in the transport of a wide variety of substrates (Ariizumi and Toriyama, 2011). ChIP analysis by Xu *et al*., (2010) confirmed that this was a direct target of AMS involved in pollen wall development. The function of WBC27 has been further clarified by Quilichini *et al*., (2010) and Dou *et al*., (2011) who confirmed that the ATP-Binding Cassette Transporter G26 was expressed in the tapetum and that it is required for male fertility and pollen exine formation in *Arabidopsis*. The ABC transporter proteins have been shown to be closely associated with secretion of sporopollenin precursors from the tapetum (Ariizumi and Toriyama, 2011),

however the specific materials being transported have yet to be fully confirmed. In Figure 6.1 it is suggested that sporopollenin precursors are processed from fatty acids produced in plastids by a series of enzymes (Choi *et al*., 2011). After fatty acids have been esterified by ACOS5 in the cytosol and hydroxylated by CYP703A2 and CYP704B1 in the ER, they are then modified further by MS2 and/or other unknown enzymes (Choi *et al*., 2011). These hydroxylated fatty acyl CoA, or further processed fatty acids, are probably transported into the locules as sporopollenin precursors (Fig 6.1).
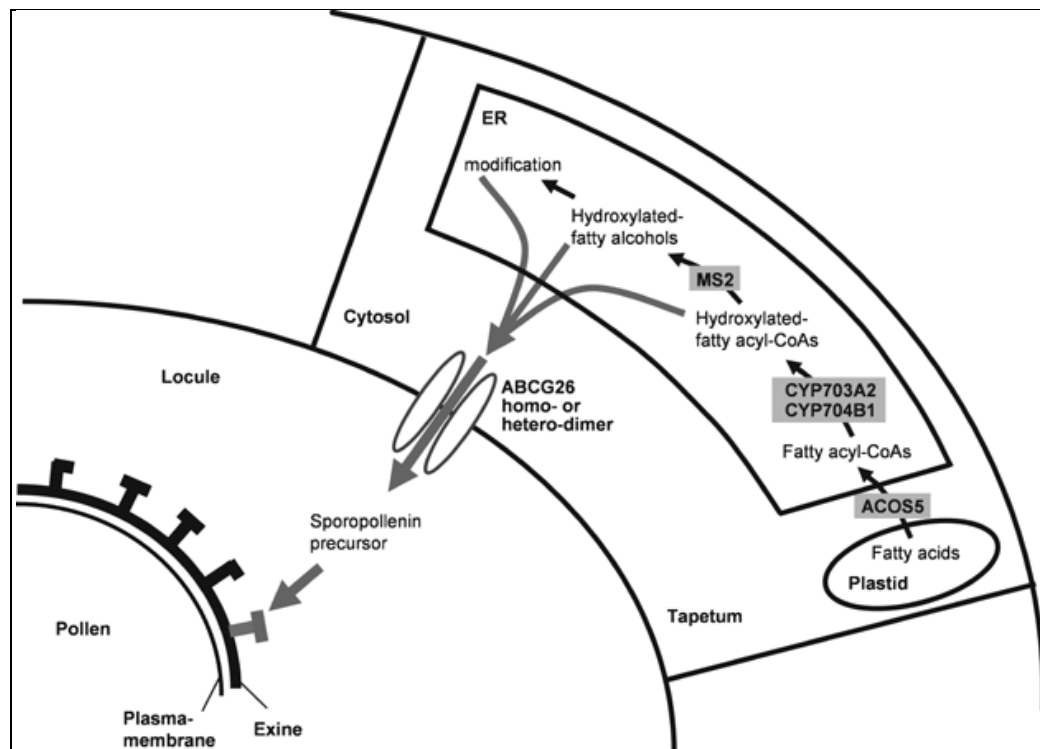
Figure 6.1 Model of ABCG26/WBC27 function in pollen-wall development (Choi *et al*., 2011). During the synthesis of sporopollenin precursors, fatty acids are modified by a series of enzymes. Fatty acids produced in the plastid are esterified with CoA by ACOS5 during transfer into the ER. In the ER, medium- and long-chain fatty acids are hydroxylated by CYP703A2 and CYP704B1, respectively, and then modified further by MS2 and other unknown reactions.

It seems likely that an ABCG/WBC-type ABC transporter may serve as an essential transporter of sporopollenin precursors for exine formation in developing pollen. However, the mechanism underlying this need to be studied further in order to establish the chemical pathway of how transport of sporopollenin precursors occurs and subsequent exine wall components happens. Future work could focus on analysing the expression profile of more putative *AMS* targets in the *ams* mutant background to establishing whether these genes are directly or indirectly regulated by AMS. Additionally it may also be of interest to analyse the expression pattern of *AMS* in the insertional mutants of these targets AMS targets. More work could involve finding out which processes these putative *AMS* targets affect in terms of metabolism and transport, during anther and pollen development. This could be achieved by analysing the chemical composition of the anther tissue and pollen wall. Further work could involve comparing the sequences of putative *AMS* targets and identifying any conserved regions with specialised functions and then linking this to possible chemical events that occur during anther and pollen wall formation.

6.2 FAST NEUTRON SEED MUTAGENESIS

A forward genetics approach was utilised to identify possible factors regulating *MALE STERILITY1* gene expression. This was conducted by Fast Neutron mutagenesis of seed containing the MS1prom:MS1-GFP fusion protein. With this analysis it has been shown that fast neutron mutagenesis is highly efficient, since a mutation frequency of 40% has achieved using Fast Neutron bombardment. From deletion mutant libraries four lines that showed altered expression of MS1

were identified, these were backcrossed over a further four generations to identify homozygous male sterile lines with altered MS1-GFP expression pattern whilst minimising the presence of additional mutations. These lines were crossed to the *ms1* recessive mutant to test for gene complementation. One line (2Y) was identified as a novel mutant since it was able to complement the *ms1* mutant, and it was therefore selected for further analysis. This was crossed with wildtype Columbia ecotype to generate 600 F2 lines for generation of recombinants and gene mapping. Marker analysis showed linkage of the 2Y mutation to a marker on chromosome 5 and 3. The chromosome 5 marker is closely linked to the *MS1* locus and is likely to reflect contamination of the line with the *ms1* mutation during backcrossing. Whereas the linkage on chromosome 3 seems more likely to reflect the location of the novel 2Y mutation, however further work needs to be conducted to confirm the mutant gene ID. Further characterisation of the 2Y line is needed to complete the mutant identification; this could involve sequencing, or restriction digestion of the *MS1* locus in lines showing linkage to the chromosome 5 marker to determine the whether the observed linkage is due to the *ms1* mutation or due to linkage with this genomic region. This could be done by looking for the loss of the *Rsa*1 restriction site that is seen in the *ms1* mutant (a *Rsa*1 restriction site is lost in the *ms1-1* mutant when compared with L*er* wildtype due to a G to A transition, which results in the loss of the Rsa1 restriction site (Wilson *et al.,* 2001). Further work could involve analysing the linked loci that may be the gene of interest by sequencing in the 2Y mutant. The identity of the 2Y loci could also be analysed using KO insertional mutants in the possible candidate genes to see if the same phenotype is seen in the genes showing linkage to the marker. RT-PCR analysis of the potential gene of interest could also be

employed in the 2Y mutant to determine if expression is altered in the 2Y mutant. Further sequencing of possible gene targets and complementation analysis could be conducted to confirm the gene identity. Depending on the nature of the gene its role in the regulation of MS1 expression could subsequently be investigated, for example determining whether the gene is a direct target of the MS1 protein and establishing the effect on MS1 gene expression. This could involve checking if the protein coded by the gene of interest interacts with MS1 protein or promoter sequence and identifying any possible protein structural similarity and regions of interest.

6.3 MS1 PROTEIN- PROTEIN INTERACTIONS

The study of protein-protein interactions (PPIs) is essential to uncover unknown functions of proteins at the molecular level and to gain insight into complex cellular networks, however false-positive protein present problems in such analyses. Therefore, despite the identification of possible interacting proteins by yeast two hybrid analysis a key goal was to confirm the PPI data from MS1 using one or more independent approaches. A number of approaches can be utilised for such work, for example *in vitro* pull-down experiments, transient expression, split YFP and FRET analyses (Steiner *et al*., 2012; Sparkes and Brandizzi, 2012; Kuroda *et al*., 2012; Piotrzkowski *et al*., 2012). During this study the MS1 protein interaction with POB2, previously suggested from screening our stamen Yeast-2-Hybrid library, was confirmed by repeating the Yeast-2-Hybrid analysis on a 1:1 basis and subsequently by Bimolecular Fluorescence Complementation (Chapter 5 Section 5.3.1 and 5.3.4). POB2 expression pattern was determined in

the wildtype background and shown to be expressed throughout plant growth in a similar pattern as reported by Genevestigator (Figure 6.2).
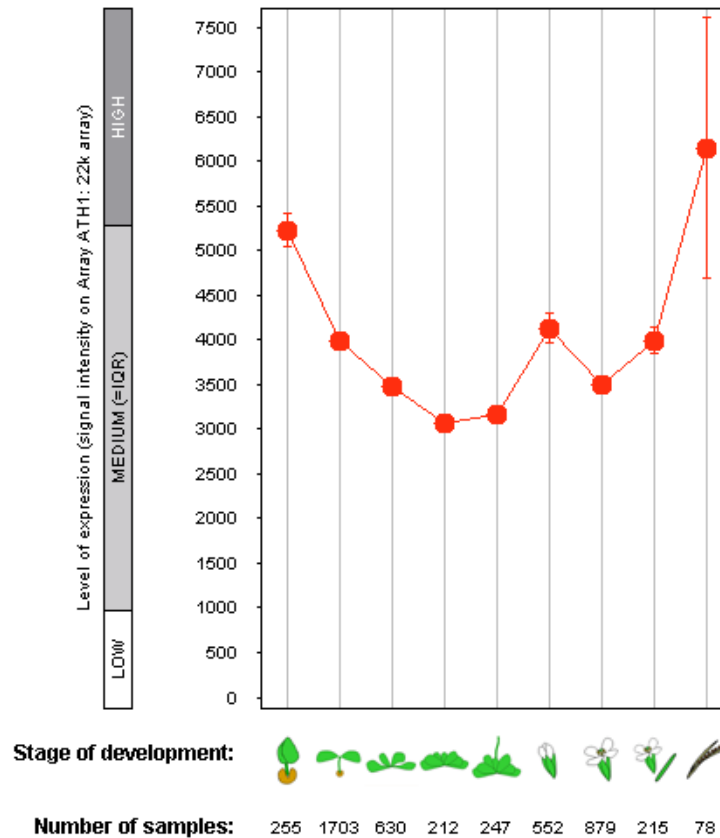
Figure 6.2 POB2 (At2g46260) Genevestigator expression profile. At2g46260 is expressed throughout growth but highly expressed during hypocotyl and cotyledon emergence from seed coat as well as old silique stage. It is also medially expressed in floral organs in unopened buds and during anthesis of Arabidopsis flowers.

MS1 interaction with POB1 was not observed, however POB1 and POB2 have been reported to dimerize through their BTB domains and function redundantly as

negative regulators of JA-mediated pathogen resistance responses (Qu *et al*., 2010). The POB2 BTB domain and some of its surrounding regions have recently been presumed to be involved in target recognition (Christians *et al*., 2012). POB2 now designated as LRB1 has been shown to have a putative nuclear localization sequence (NLS) near the N terminus downstream of the BTB domain (Christians *et al*., 2012). Interestingly Christians *et al*., (2012) also found a higher level of At2g46260 (LRB1) mRNA in pollen. Even though the basis of MS1 interaction with POB2 has not yet been fully characterised, we can hypothesise that the interaction of MS1 with POB2 occurs in the nucleus and while MS1 binds to DNA with its PHD finger motif, POB2 may bind MS1 with its protein binding motif (BTB). This could potentially form a protein complex by either homo, or heterodimerization, resulting in POB2 tagging the MS1 protein with ubiquin molecules, potentially for degradation in the tapetum. If this is occurring, it is likely that once the MS1 protein is tagged with ubiquitin molecules, it could signal to other ligases to attach additional ubiquitin molecules thereby targeting MS1 for degradation by the proteasome. This could thereby serve to regulate the turnover of the MS1 protein and therefore limit anther and pollen gene regulation.

Further work is needed to confirm if MS1 is up or down regulated by POB2. To confirm this genotypic and phenotypic analysis of single homozygous mutants and double mutants could be performed with respect to male sterility. In addition more work is needed on characterisation of POB2 during pollen development by analysing its cell-specific expression pattern in the wildtype and mutant background. It may also be of interest to find out if MS1 interacts with both isoforms of Cullin 3 proteins in yeast and *in planta*, to relate the mode of

interaction POB2 has with Cullin 3 proteins to its interaction with MS1. The BiFC confirmed the Y2H result suggesting that POB2 is a definite interactor of MS1; this information will contribute to currently existing MS1 regulatory network thereby giving more insight into gene regulation during pollen development in *Arabidopsis*. These data could also be extended to research in commercially important crops

Besides the approaches utilised in this PPI work, researchers across the globe are establishing other PPI methodologies which are now seeing the development of new methods and/or improvements in existing methods (Fukao 2012). Some of these techniques involve identifying proteins by mass spectrometry (MS), multidimensional protein identification technology or OFFGEL electrophoresis analyses, one-shot analysis with a long column or filter-aided sample preparation methods (Ros *et al*., 2002; Zhang *et al*., 2012; Fukao 2012). These advanced techniques potentially allow thousands of proteins to be identified, and such detailed proteomic methods should also permit the identification of transient binding or PPIs with weak affinity.

6.4 FUTURE PERSPECTIVES

6.4.1   THE ROLE OF MALE STERILITY IN PLANT BREEDING

Genetically modified (GM) crops have great potential to enhance the productivity and sustainability of agriculture globally, but their adoption raises several important issues, highlighted by the strong consumer aversion to GM food

products in some countries. One of these issues is the potential impact of gene flow. Gene flow is the natural process of movement of genes between individual organisms. Among plants this occurs mainly by pollen from one plant successfully cross-pollinating a flower from another plant and producing viable seed, a process known as outcrossing. Gene flow already occurs from conventionally bred crops but there is heightened interest in gene flow issues because of the release of genetically modified (GM) crops and the potential for the introduced genes (transgenes) to flow into other plants or organisms. Consequently being able to limit gene flow and its implications is an important target for the successful introduction of GM crops. Among many technologies being developed preventing pollen production/release (male sterility) has the potential to reduce gene flow, but this can have significant limitations in some crops, since male sterility may adversely affect crop yield, since unsatisfactory restoration of fertility may occur.

Increasing crop productivity has contributed to global food, feed and fibre security. Production of hybrids is one of the techniques that enhance worldwide agricultural production. However managing the reproduction of crop plants can be a major challenge. Male sterility is one system of pollination control that growers use to produce hybrids with greater ease (Gardner *et al*., 2009). Because hybrids are so desirable and the techniques for making them can be so difficult, researchers have long sought an easier way to breed hybrids by gaining a better understanding of the role of the male sterility in selective plant breeding (breeding plants for particular traits). Examples of male sterility have typically involved cytoplasmic male sterility (governed by cytoplasmic genes) (Pelletier and Budar

2007), genetic male sterility (governed by nuclear genes) (Acquaah, 2007), transgenic male sterility (induced by the technique of genetic engineering) (Rosellini *et al*., 2001) and chemical induced male sterility (induced by the use of chemical) (Acquaah, (2012). In all of these strategies the aim is to generate plants with controllable fertility so that cross-fertilisation can be achieved to generate hybrids and then subsequently fertility can be restored to allow full seed-set to occur.

6.4.2   THE MAIN CHALLENGES FOR UNDERSTANDING POLLEN DEVELOPMENT

During these studies challenges were faced in the analysis of stage specific expression of genes in mutants carrying the MS1-GFP gene constructs. However this problem was improved by using the confocal microscope. Apart from chemical or radiation mutagenesis, male sterility can be caused by many factors, including diseases or unfavourable environmental and growth conditions which could be misleading during phenotypic screening of male sterile lines. This problem is normally solved by genotypic analysis of the plants exhibiting failure to produce pollen-forming tissues in anthers or viable pollen produced in non-dehiscent anthers, deformed or non-viable pollen grains, and abnormal pollen. The study of pollen development by generating genetically induced male sterile mutants can sometimes be limiting because it cannot be used to maintain a pure line of male-sterile crops (Acquaah, 2012). When homozygous recessive male-sterile plants are generated they must be crossed with a heterozygous or wildtype source of pollen so that the seeds that result from crossing the male-sterile with

the male-fertile plants can be used for future breeding (Acquaah, 2007). Despite the many challenges scientists face in elucidating the genetic regulation of anther and pollen development, more discoveries made in the model system of tapetum function and pollen formation will have direct relevance to other areas of plant science.

CHAPTER 7

REFERENCES

7. REFERENCES

Aarts, M. G. M., Hodge, R, Kalantidis, K., Florack, D. and Wilson, Z. A. (1997) The *Arabidopsis MALE STERILITY2* protein shares similarity with reductases in elongation/condensation complexes. Plant Journal **12**: 615-23.

Acquaah, G. (2007) Principles of plant genetics and breeding. Blackwell Publishing, Malden, MA, Oxford, UK.

Acquaah, G. (2012) Mutagenesis in plant breeding, in principles of plant genetics and breeding, 2nd Edition: John Wiley and Sons, Ltd, Chichester, UK.

Ahlers, F., Lambert, J. and Wiermann, R. (1999) Structural elements of sporopollenin from the pollen of *Torreya californica torr*. (Gymnospermae): using the H-1-NMR technique. Zeitschrift für Naturforschung C **54**: 492-95.

Alexander, M. P., (1980) A versatile stain for pollen, fungi, yeast and bacteria. Stain Technology **55**: 13-18.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D.J. (1990) Basic local alignment search tool. Journal of Molecular Biology **215**: 403-410.

Alves-Ferreira, M., Wellmer, F., Banhara, A., Kumar, V., Riechmann, J. and Meyerowitz, E. M. (2007) Global expression profiling applied to the analysis of arabidopsis stamen development. Plant Physiology **145**: 747-762.

Ariizumi, T. and Toriyama, K. (2011) Genetic regulation of sporopollenin synthesis and pollen exine development. Annual Review of Plant Biology **62**: 437-460.

Ariizumi, T., Hatakeyama, K., Hinata, K., Inatsugi, R. and Nishida, I. (2004) Disruption of the novel plant protein NEF1 affects lipid accumulation in the plastids of the tapetum and exine formation of pollen, resulting in male sterility in *Arabidopsis thaliana*. The Plant Journal **39**: 170-81.

Ariizumi, T., Hatakeyama, K., Hinata, K., Sato, S and Kato, T. (2003) A novel male-sterile mutant of *Arabidopsis thaliana*, *faceless pollen-1*, produces pollen with a smooth surface and an acetolysis-sensitive exine. Plant Molecular Biology **53**: 107-16.

Arondel, V., Vergnolle, C., Cantrel, C. and Kader, J. C. (2000) Lipid transfer proteins are encoded by a small multigene family in Arabidopsis thaliana. Plant Science **2000** 157: 1-12.

Ascencio-Ibáñez, J. T., Sozzani, R., Lee, T., Chu, T., Wolfinger, R. D., Cella, R. and Hanley-Bowdoin, L. (2008) Global analysis of arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. Plant Physiology **148**: 436-454.

Asero, R., Mistrello, G., Roncarolo, D., de Vries, S. C., Gautier, M. F., Ciurana, C. L. F., Verbeek, E., Mohammadi, T., Knul-Brettlova, V., Akkerdaas, J. H., Bulder, I., Aalberse, R. C. and Van Ree R. (2001) Lipid transfer protein: A pan-allergen in plant-derived foods that is highly resistant to pepsin digestion. International Archives of Allergy and Immunology **124**: 67-69.

Aya, K., Ueguchi-Tanaka, M., Kondo, M., Hamada, K. and Yano, K. (2009) Gibberellin modulates anther development in rice via the transcriptional regulation of GAMYB. The Plant Cell **21**: 1453-72.

Barry Causier, B. and Davies, B. (2002) Analysing protein-protein interactions with the yeast two-hybrid system. Plant Molecular Biology **50**: 855-870.

Bastola, D.R., Pethe, V.V., and Winicov, I. (1998) Alfin1, a novel zinc-finger protein in alfalfa roots that binds to promoter elements in the salt-inducible *MsPRP2* gene. Plant Molecular Biology **38**: 1123–1135.

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**: 713-25

Bedinger, P. (1992) The remarkable biology of pollen. Plant Cell **4**: 879-887.

Belfield, E. J., Gan, X., Mithani, A., Brown, C., Jiang, C., Franklin, K., Alvey, E., Wibowo, A., Jung, M., Bailey, K., Kalwani, S., Ragoussis, J., Mott, R. and Harberd, N. P. (2012) Genome-wide analysis of mutations in mutant lineages selected following fast-neutron irradiation mutagenesis of Arabidopsis thaliana. Genome Research **22**: 1306-15.

Belfield, E. J., Hughes, R. K., Tsesmetzis, N., Naldrett, M. J. and Casey, R. (2007) The gateway pDEST17 expression vector encodes a −1 ribosomal frameshifting sequence. Nucleic Acids Research **35**: 1322-1332.

Bertone, P. (2005) Global identification of human transcribed sequences with genome tiling arrays. Science **308**: 1149-1154.

Bhat, R. A., Lahaye, T. and Panstruga, R. (2006) The visible touch: *in planta* visualization of protein-protein interactions by fluorophore-based methods. Plant Methods **2**: 12.

Bienz, M. (2006) The PHD finger, a nuclear protein-interaction domain. Trends Biochemical Science **31**: 35-40.

Biswas, K. K., Ooura, C., Higuchi, K., Miyazaki ,Y., Van Nguyen, V., Rahman, A., Uchimiya, H., Kiyosue, T., Koshiba, T., Tanaka, A., Narumi, I. and Oono, Y. (2007) Genetic characterization of mutants resistant to the antiauxin p-chlorophenoxyisobutyric acid reveals that AAR3, a gene encoding a DCN1 like protein, regulates reponses to the synthetic auxin 2'4-dichlorophenoxyacetic acid in Arabidopsis roots. Plant Physiology **145**: 773-785.

Boavida, L. C., Becker, J. D. and Feijó, J. A. (2005) The making of gametes in higher plants. International Journal of Developmental Biology **49**: 595-614.

Boguski, M. S., Tolstoshev, C. M. A. and Bassett, D. E. (1994) The turning point in genome research. Science **265**:1993-1994.

Bolon, Y-T., Haun, W. J., Xu, W. W., Grant, D., Stacey, M. G., Nelson, R. T., Gerhardt, D. J., Jeddeloh, J. A., Stacey, G., Muehlbauer, G. J., Orf, J. H., Naeve, S.L. Stupar R. M. and Vance C. P. (2011) Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. Plant Physiology **156**: 240-253.

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

Bouchez, D. and Höfte, H (1998) Functional genomics in plants. Plant Physiology **118**: 725-732.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D. H., Johnson, D., Luo, S., McCurdy S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S. R., Moon, K., Burcham, T., Pallas, M., DuBridge, R. B., Kirchner, J., Fearon, K., Mao, J. and Corcoran (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nature Biotechnology **18**: 597-8.

Budar, F. and Pelletier, G. (2001) Male sterility in plants: occurrence, determinism, significance and use. Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie **324**: 543-550.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T. L. (2009) BLAST+: architecture and applications. BMC Bioinformatics: **10**: 421.

Chan, F. K. M., Siegel, R. M., Zacharias, D., Swofford, R., Holmes, K. L., Tsien, R. Y. and Lenardo, M. J. (2001) Fluorescence resonance energy transfer analysis of cell surface receptor interactions and signaling using spectral variants of the green fluorescent protein. Cytometry **44**: 361-368.

Chang, H. S., Zhang, C., Chang, Y. H., Zhu, J., Xu, X-F., Shi, Z-H., Zhang, X-L., Xu, L., Huang, H., Zhang, S. and Yang Z-N, (2012) *NO PRIMEXINE AND PLASMA MEMBRANE UNDULATION* is essential for primexine deposition and plasma membrane undulation during microsporogenesis in Arabidopsis. Plant Physiology **158**: 264-272.

Chen, W., Yu, X. H., Zhang, K., Shi, J., Schreiber, L., Shanklin, J. and Zhang, D. (2011) Male Sterile 2 Encodes a Plastid-localized Fatty Acyl ACP reductase required for pollen exine development in *Arabidopsis thaliana*. Plant Physiology **157**: 842-853.

Chen, X., Goodwin, S. M., Boroff, V. L., Liu, X. and Jenks, M. A. (2003) Cloning and characterization of the *WAX2* gene of *Arabidopsis* involved in cuticle membrane and wax production. The Plant Cell **15**: 1170-85.

Chen, Z. J. (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annual Review of Plant Biology **58**: 377-406.

Chevray, P. and Nathans, D. (1992) Protein interaction cloning in yeast: Identification of mammalian proteins that react with the leucine zipper of Jun. Proceedings of the National Academy of Sciences **89**: 5789-5793.

Choi, H, Jin, J. Y, Choi S, Hwang JU, Kim YY, Suh MC, Lee Y. (2011) An ABCG/WBC-type ABC transporter is essential for transport of sporopollenin precursors for exine formation in developing pollen. Plant Journal **65**: 181-93.

Choi, H., Jin, J-Y., Choi, S., Hwang, J.-U., Kim, Y.-Y., Suh, M. C. and Lee, Y. (2011) An ABCG/WBC-type ABC transporter is essential for transport of sporopollenin precursors for exine formation in developing pollen. The Plant Journal **65**: 181-193.

Christians, M. J., Gingerich, D. J., Hua, Z., Lauer, T.D. and Vierstra, R. D. (2012) The Light-Response BTB1 and BTB2 proteins assemble nuclear ubiquitin ligases that modify phytochrome B and D signalling in *Arabidopsis*. Journal of Plant Physiology **160**: 118-34.

Clark, J., Beckage, B., Lambers, H. R. J., Ibanez, I., LaDeau, S., MacLachlan, J., Mohan, J. and Rocca, M. (2002) Dispersal and plant migration. Encyclopedia of Global Environmental Change **3**: 81-93.

Colhoun, C. W., Steer, M. W., (1981) Microsporogenesis and the mechanism of cytoplasmic male sterility in maize. Annals of Botany **48**: 417-24.

Cui, X., Wang, Q., Yin, W., Xu, H. and Wilson, Z. A., Wei, C., Pan, S. and Zhang, D. (2012) PMRD: a curated database for genes and mutants involved in plant male reproduction. BMC Plant Biology **12**: 215.

Da Costa-Nunes J. A and Grossniklaus, U. (2003) Unveiling the gene-expression profile of pollen. Genome Biology **5**: 205.

Dai, Y., Wang, H., Li, B., Huang, J., Liu, X., Zhou, Y., Mou, Z. and Li, J. (2006) Increased expression of MAP KINASE KINASE7 causes deficiency in polar auxin transport and leads to plant architectural abnormality in *Arabidopsis.* Plant Cell **2**: 308-20.

Dassa, E. and Bouige, P. (2001) The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. Research in Microbiology **152**: 211-229.

David, L., Huber, W., Granovskaia, M., Toedling J., Palm, C. J., Bofkin, L., Jones, T., Davis, R. W. and Steinmetz, L. M. (2006) A high-resolution map of transcription in the yeast genome. Proceedings for National Academy Science **103**: 5320-5325.

Davis, R. H., Tanaka, C. K., Vensel, W. H., Hurkman, W. J. and McCormick, S. (2005) Arabidopsis leaves undergoing cold stress. Plant Molecular Biology **52**: 553-567.

Dawson, R. J. P. and Locher, K. P. (2007) Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. FEBS Letters **581**: 935-938.

De Azevedo Souza, C., Kim, S. S., Koch, S., Kienow, L., Schneider, K. (2009) A novel fatty Acyl-CoA synthetase is required for pollen development and sporopollenin biosynthesis in *Arabidopsis*. The Plant Cell **21**:507-25.

Diederichs, K., Diez, J., Greller, G., Muller, C, Breed, J., Schnell, C., Vonrhein, C., Boos, W. and Welte, W., (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon Thermococcuslitoralis. The EMBO Journal **19**: 5951-5961.

Dobritsa, A. A., Shrestha, J., Morant, M., Pinot, F. and Matsuno, M. (2009) CYP704B1 is a long-chain fatty acid omega-hydroxylase essential for sporopollenin synthesis in pollen of *Arabidopsis*. Plant Physiology **151**: 574-89.

Dong, X., Hong, Z., Sivaramakrishnan, M., Mahfouz, M., Verma, D.P (2005) Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in Arabidopsis. Plant Journal **42**: 315-28.

Du, W., Vidal, M., Xie, J. E., Dyson, N. (1996). RBF, a novel RB-related gene that regulates E2F activity and interacts with cyclin E in Drosophila. Genes and Development. **10**: 1206-1218.

Ecker, J. R. (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell **133**: 523-536.

Fan, J. Y., Cui, Z. Q., Wei, H. P., Zhang, Z. P., Zhou, Y. F., Wang, Y. P. and Zhang, X. E. (2008) Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein–protein interactions in living cells. Biochemical and Biophysical Research Communications **367**: 47-53.

Fan, L. M., Wang, Y. F. and Wang, W. H. (2001) In vitro Arabidopsis pollen germination and characterization of the inward potassium currents in *Arabidopsis* pollen grain protoplasts. Journal of Experimental Botany **52**: 1603-1614.

FAO (2008). The state of food insecurity in the world. Rome: Food and Agriculture Organization of the United Nations.

Feldmann, K A. and Marks, M. D. (1987) Agrobacterium-mediated transformation of germinating seeds of *Arabidopsis thaliana*: A non-tissue culture approach. Molecular and General Genetics **208**: 1-2.

Feldmann, K. A. and Marks, M. D. (1987) *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*. Molecular and General Genetics **3**: 1-9.

Foresight. The Future of Food and Farming (2011) Final Project Report. The Government Office for Science, London.

Francis, D. M. and Page, R. (2010) Strategies to optimize protein expression in *E. coli*. Current Protocols in Protein Science **61**: 5.24.1-5.24.29.

Fukao Y. (2012) Protein–Protein interactions in Plants. Plant and Cell Physiology **53**: 617-625.

Fukao, Y. (2012) Protein-Protein interactions in plants. Plant Cell Physiology **53**: 617-25.

Furness, A. (2002) Evolution of microsporogenesis in angiosperms. International Journal of Plant Science **163**: 235-260.

Galperin, M. Y. (2005) The molecular biology database collection: 2005 update. Nucleic Acids Research **1**: D5-D24.

Ganesh, M., Chung-Jui, T., Podila G. K. and Chiang (1994) Agrobacterium-mediated transformation of quaking aspen (*Populus tremuloides*) and regeneration of transgenic plants. Plant Cell Reports **14**: 2-3.

Gardner, N., Felsheim, R. and Smith. A. G. (2009) Production of male- and female-sterile plants through reproductive tissue ablation. Journal of Plant Physiology **166**: 871-881.

Gaudet, R. and Wiley, D. C. (2001) Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. The EMBO Journal **20**: 4964-72.

Gerhard, D. S., Wagner, L., Feingold, E. A. and Shenmen, C. M. (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome research **14**: 2121-7.

Ghosh, S., Bell, I., Gerhard, D. S. and Gingeras, T. R. (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science **20**: 1149-54.

Gibalova, A., Renak, D., Matczuk, K., Dupl'akova, N. and Chab, D. (2009) *AtbZIP34* is required for *Arabidopsis* pollen wall patterning and the control of several metabolic pathways in developing pollen. Plant Molecular Biology **70**: 581-601.

Gibson, G. and Muse, S. V. (2009). A Primer of Genome Science, 3rd Edition: Sinauer Press.

Gietl, C. and Schmid, M. (2001) Ricinosomes: an organelle for developmentally regulated cell death in senescing plant tissues. Naturwissenschaften **88**: 49-58.

Goldberg, R. B., Beals T. P. and Sanders P. M. (1993) Anther development: Basic principles and practical applications. The Plant Cell **5**: 1217-1229.

Grienenberger, E., Besseau, S., Geoffroy, P., Debayle, D. and Legrand, M. (2009) 20th International Conference on Arabidopsis Research. Tair accession: 501734164.

Grini, P. E., Thorstensen, T., Alm, V., Vizcay-Barrena, G., Windju, S. S., Jorstad, T. S., Wilson, Z. A. and Aalen, R.B. (2009) The ASH1 HOMOLOG 2 (ASHH2) histone H3 methyltransferase is required for ovule and anther development in Arabidopsis. PLoS ONE **4**: e7817.

Guan, Y. F., Huang, X. Y., Zhu, J., Gao, J. F., Zhang, H. X. and Yang, Z. N. (2008) *RUPTURED POLLEN GRAIN1*, a member of the MtN3/saliva gene family, is crucial for exine pattern formation and cell integrity of microspores in *Arabidopsis*. Plant Physiology **147**:852-63

Guo, L., Devaiah, S. P., Narasimhan, R., Pan, X., Zhang Y., Zhang, W. and Wang, X. (2012) Cytosolic glyceraldehyde-3-phosphate dehydrogenases interact with phospholipase to transduce hydrogen peroxide signals in the *Arabidopsis* response to stress. The Plant Cell **24**: 2200-2212.

Hammond-Kosack, K.E., Tang, S.J., Harrison, K. and Jones, J. (1998) The tomato *Cf-9* disease resistance gene functions in tobacco and potato to confer responsiveness to the fungal avirulence gene product Avr9. The Plant Cell **10**: 1251-1266.

Hannoufa, A., Negruk, V., Eisner, G. and Lemieux, B. (I1996) The *CER3* gene of *Arabidopsis thaliana* is expressed in leaves, stems, roots, flowers and apical meristems. The Plant Journal **10**: 459-67.

Harbers, M. and Carninci P. (2005) Tag-based approaches for transcriptome research and genome annotation. Nature Methods **2**: 495-502.

Hemsley, A. R. and Gabarayeva, N. I. (2007) Exine development: the importance of looking through a colloid chemistry "window." Plant Systems and Evolution **263**: 25-49.

Higgins, C. F. (1992) ABC transporters: from microorganisms to man. Annual Review of Cell and Developmental Biology **8**: 67-113.

Higgins, C. F. (2001) ABC transporters: physiology, structure and mechanism-an overview. Research in Microbiology **152**: 205-10.

Hiratsu, K., Matsui, K., Koyama, T. and Ohme-Takagi, M. (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. Plant Journal **34**: 733-739.

Holland, I. B. and Blight, M. A. (1999) ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. Journal of Molecular Biology **293**: 381-99.

Holmes-Davis, R., Tanaka, CK., Vensel, W. H., Hurkman, W. J. and McCormick, S. (2005) Proteome mapping of mature pollen of *Arabidopsis thaliana.* Proteomics **5**: 4864-4884.

Honys, D. and Twell, D. (2003) Comparative analysis of the *Arabidopsis* pollen transcriptome. Plant Physiology **132**: 640-652.

Honys, D. and Twell, D. (2004) Transcriptome analysis of haploid male gametophyte development in *Arabidopsis* genome biology. Research **5**: 1-11.

Houman, F. and Sawhney, V. K. (1999) MS32-regulated timing of callose degradation during microsporogenesis in *Arabidopsis* is associated with the accumulation of stacked rough ER in tapetal cells. Sex Plant Reproduction **12**: 188-193.

Hu, C. D. and Kerppola, T. K. (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolour fluorescence complementation analysis. Nature Biotechnology **21**: 539-545.

Huang, L., Berkelman, T., Franklin, A. E. and Hoffman, N. E. (1993) Characterization of a gene encoding a Ca^{2+}-ATPase-like protein in the plastid envelope. Proceedings of National Academy of Sciences **90**: 10066-10070

Hung, L. W., Wang, I. X., Nikaido, K., Liu, P. Q., Ames, G. F. and Kim, S. H. (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. Nature **396**: 703-7

Hvorup, R. N., Goetz, B.A., Niederer, M., Hollenstein, K., Perozo, E. and Locher, K.P. (2007) Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. Science **317**: 1387-1390.

Ikeda, M. and Ohme-Takagi, M. A. (2009) Novel Group of Transcriptional Repressors in *Arabidopsis* Plant and Cell Physiology **50**: 970-975.

Ito, T. and Shinozaki, K. (2002) Protein with a PHD-finger motif, is expressed in tapetal cells and is required for pollen maturation. Plant Cell Physiology **43**: 1285-1292.

Ito, T., and Shinozaki, K. (2002) The MALE STERILITY1 gene of Arabidopsis, encoding a nuclear protein with a PHD-finger motif, is expressed in tapetal cells and is required for pollen maturation. Plant Cell Physiol. **43**: 1285-1292.

Ito, T., Nagata, N., Yoshiba, Y., Ohme-Takagi, M., Ma, H. and Shinozaki K. (2007) *Arabidopsis MALE STERILITY1* encodes a PHD-Type transcription factor and regulates pollen and tapetum development. The Plant Cell **11**: 3549-3562.

Jamet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H. and Pont-Lezica, R. (2008) Recent advances in plant cell wall proteomics. Proteomics **8**: 893-908.

Jander, G., Norris, S. R., Rounsley, S. D., Bush, D. F., Levin L. M. and Last, R. L. (2002) Arabidopsis map-based cloning in the post-genome era. Plant Physiology **129**: 440-450.

Jean-Claude, K. (1996) Lipid-transfer proteins in plants. Annual Review of Plant Physiology and Plant Molecular Biology **47**: 627-654.

Jenks, M. A., Tuttle, H. A., Eigenbrode, S. D, Feldmann, K. A. (1995) Leaf epicuticular waxes of the eceriferum mutants in *Arabidopsis*. Plant Physiology **108**: 369-377.

Jung, K. H., Han, M. J., Lee, D. Y., Lee, Y. S., Schreiber, L. and Franke, R. (2006) Wax-deficient anther1 is involved in cuticle and wax production in rice anther walls and is required for pollen development. The Plant Cell **18**: 3015-3032.

Kapila, J., Rycke, R. D., Montagu, M. V. and Angenon, G. (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. Plant Science **122**: 101-108.

Karpowich N, Martsinkevich O, Millen L, Yuan, Y. R., Dai, P. L., MacVey, K., Thomas, P. J. and Hunt, J. F. (2001) Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. Structure **9**: 571-86.

Kerppola, T. K. (2006) Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. Nature Protocols **1**: 1278-1286.

Kerppola, T. K. (2008) Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. Annual Review of Biophysics **37**: 465-487.

Kerr, I. D., (2002) Structure and association of ATP-binding cassette transporter nucleotide-binding domains. Biochimica et Biophysica Acta **1561**: 47-64.

Kim, O. K., Jung, J. H. and Park, C. M. (2010) An Arabidopsis F-box protein regulates tapetum degeneration and pollen maturation during anther development. Planta **232**: 353-366.

Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., Sasaki, D., Imamura, K., Kai, C., Harbers, M., Hayashizaki, Y. and Carninci, P. (2006) CAGE: cap analysis of gene expression. Nature Methods **3**: 211-222.

Köhl, K. I. and Gremmels, J. (2010) Documentation system for plant transformation service and research. Plant Methods **6**: 4.

Koltunow, A. M., Truettner, J., Cox, K. H., Wallroth, M. and Goldberg, R. B. (1990) Different temporal and spatial gene expression patterns occur during anther development. The Plant Cell **2**: 1201-1224.

Koornneef, M. and Van der Veen, J. H. (1980) Induction and analysis of gibberellin sensitive mutants in *Arabidopsis thaliana* (L.) heynh. Genetics **58**: 257-263.

Koornneef, M., Cherry, D.W., Dean, J.M. and Rounsley, C. (1998) *Arabidopsis thaliana* a model plant for genome analysis. Science **282**: 662-682.

Kurata, T., Kawabata-Awai, C., Sakuradani, E., Shimizu, S., Okada, K. and Wada, T. (2003) The *YORE-YORE* gene regulates multiple aspects of epidermal cell differentiation in *Arabidopsis*. The Plant Journal **36**: 55-66.

Kuroda, H., Yanagawa, Y., Takahashi, N., Horii, Y. and Matsui, M. (2012) Analysis of interaction and localization of Arabidopsis SKP1-LIKE (ASK) and F-Box (FBX) proteins. PLoS ONE **7**: e50009.

Kuromori T, Ito T, Sugimoto E, Shinozaki K. (2011) Arabidopsis mutant of AtABCG26, an ABC transporter gene, is defective in pollen maturation. Journal of Plant Physiology **68**: 2001-5.

Kuromori, T., Sugimoto, E. and Shinozaki, K. (2011) Arabidopsis mutants of AtABCG22, an ABC transporter gene, and increase water transpiration and drought susceptibility. The Plant Journal **67**:885-94.

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, S. A., Lee, C. H., Nelson, W. D., Ploetz, L. Singh, S. Wensel, A. and Huala, E. (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acid Research **40**: D1202–D1210.

Lee, J. and Lee, D. (2003) Use of SAGE technology to reveal changes in gene expression in Proteome mapping of mature pollen of *Arabidopsis thaliana*. Proteomics **5**: 4864-4884.

Li, H, Yuan, Z, Vizcay-Barrena, G, Yang, C, Liang, W, Zong J, Wilson, Z. A and Zhang, D. (2011) PERSISTENT TAPETAL CELL1 encodes a PHD-finger protein that is required for tapetal cell death and pollen development in rice Plant Physiology **156**: 615-630.

Li, H. and Zhang, (2010) Biosynthesis of anther cuticle and pollen exine in rice. Plant Signal Behaviour **5**: 1121-1123.

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P. and Schreiber, L. (2010) Cytochrome P450 family member CYP704B2 catalyzes the omega-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice.The Plant Cell **22**: 173-190.

Li, N., Zhang, D. S. and Liu, H. S. (2006a) The rice *tapetum degeneration retardation* gene is required for tapetum degradation and anther development. The Plant Cell **18**: 2999-3014.

Li, X. and Zhang, Y. (2002) Reverse genetics by fast neutron mutagenesis in higher plants. Functional and Integrative Genomics **2**:254-258.

Lozano-Juste, J., Colom-Moreno, R. and León J. (2011) *In vivo* protein tyrosine nitration in *Arabidopsis thaliana*. Journal of Experimental Botany **62**: 3501-17.

M. Bienz (2006) The PHD finger, a nuclear protein-interaction domain. Trends Biochemical Sciences **31**: 35-40.

Ma, H. (2005) Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. Annual Review of Plant Biology **56**: 393-434.

Ma, X., Feng, B. and Ma, H. (2012) AMS-dependent and independent regulation of anther transcriptome and comparison with those affected by other Arabidopsis anther genes. BMC Plant Biology **12**: 23.

Malmström, S., Askerlund, P. and Palmgren M. G. (1997) A calmodulin-stimulated Ca^{2+}-ATPase from plant vacuolar membranes with a putative regulatory domain at its N-terminus. FEBS Letters **400**: 324-328.

Mariani, C., Debeuckeleer, M., Truettner, J., Leemans, J. and Goldberg, R. B. (1990) Induction of male-sterility in plants by a chimeric ribonuclease gene. Nature **347**: 737-741.

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. and Y. Gilad (2008) RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Research **18**: 1509-1517.

Martínez-Zapater, J. M. and Salinas, J. (1988) *Arabidopsis* Protocols. Biologia Plantarum **1**: 110.

Masi, A., Cicchi, R., Carloni, A., Pavone, F. S. and Arcangeli, A. (2010) Optical methods in the study of protein–protein interactions. Advances in Experimental Medicine and Biology **674**: 33-42.

Matsuno, M., Compagnon, V., Schoch, G. A., Schmitt, M., Debayle, D., Bassard, J. E., Pollet, B., Hehn, A., Heintz, D., Ullmann, P., Lapierre, C., Bernier, F., Ehlting, J. And Werck-Reichhart, D. (2009) Evolution of a novel phenolic pathway for pollen development Science **325**: 1688-92.

Mayfield, J. A., Fiebig, A., Johnstone, S. E and Preuss, D. (2001) Gene families from the *Arabidopsis thaliana* pollen coat proteome. Science **292**: 2482-85.

McCallum, C. M., Comai, L., Greene, E. A. and Henikoff, S. (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiology **123**: 439-442.

McCormick, S. (2003) Male gametophyte development. The Plant Cell **5**: 1265-1275.

McCormick, S. (2004) Control of male gametophyte development. The Plant Cell **16**: S142-S153.

Mei, Y., Gao, H. B., Yuan, M. and Xue, H. W. (2012) The Arabidopsis ARCP protein, CSI1, which is required for microtubule stability, is necessary for root and anther development. Plant Cell **24**: 1066-80.

Meinke, D. W., Cherry, J. M., Dean, C. Rounsley, S. D. and Koornneef, M. (1998) Arabidopsis thaliana: a model plant for genome analysis. Science **282**: 662-682.

Meshi, T. and Iwabuchi, M. (1995) Plant transcription factors. Plant Cell Physiology **36**: 1405-1420.

Meyerowitz, E. M. and Somerville, C. R. (1994) *Arabidopsis*. Cold Spring Harbor Laboratory Press Cold Spring Harbor, New York.

Michnick, S. W., Ear, P. H., Manderson, E. N., Remy, I. and Stefan, E. (2007) Universal strategies in research and drug discovery based on protein-fragment complementation assays. Nature Reviews Drug Discovery **6**: 569-582.

Millar, A. A. and Gubler F. (2005) The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. The Plant Cell **17**: 705-21.

Millar, A. A.,Clemens, S.,Zachgo, S.,Giblin, E. M.,Taylor, D. C.,Kunst, L. (1999) CUT1, an Arabidopsis gene required for cuticular wax biosynthesis and pollen fertility, encodes a very-long-chain fatty acid condensing enzyme. The plant cell **11**: 825.

Moens, P. D. J. And dos Remedios, C. G. (2001) Analysis of models of F-actin using fluorescence resonance energy transfer spectroscopy. Results and Problems in Cell Differentiation **32**: 59-77.

Moore, P. D., Webb, J. A. and Collinson, M. E. (1991) Pollen Analysis. Blackwell Scientific Publications, UK.

Morell, M., Espargaro, A., Aviles, F. X. and Ventura, S. (2008) Study and selection of in vivo protein interactions by coupling bimolecular fluorescence complementation and flow cytometry. Nature Protocols **3**: 22-33.

Morin, R. D., O'Connor, M. D. and Griffith, M. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Research **18**: 610-621.

Murgia, M., Charzynska, M., Rougier, M. and Cresti, M. (1991) Secretory tapetum of *Brassica oleracea*: polarity and ultrastructural features. Sexual Plant Reproduction **4**: 28-35.

Murmu, J., Bush, M. J., DeLong, C, Li, S., Xu, M., Khan, M., Malcolmson, C., Fobert, P. R., Zachgo, S., and Hepworth, S. R. (2010) *Arabidopsis* basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. Plant Physiology **154**: 1492-504.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, K. and Snyder, M. (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science **6**: 1344-9.

Nakamura, M. and Carninci, P. (2004) Cap analysis gene expression: CAGE. Tanpakushitsu Kakusan Koso **49**: 2688-2693.

Noir, S., Brautigam, A., Colby, T., Schmidt, J. and Panstruga, R. (2005) A reference map of the *Arabidopsis thaliana* mature pollen proteome. Biochemical and Biophysical Research Communications - Elsevier **337**: 1257-1266.

Nussbaumer, T., Martis, M. M., Roessner, S. K., Pfeifer, M., Bader, K. C., Sharma, S., Gundlach, H and Spannagl, M. (2013) MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Research **41**: D1144-51.

Obayashi T., Hayashi, S., Saeki, M., Ohta, H. and Kinoshita K. (2009) ATTED-II provides coexpressed gene networks for *Arabidopsis* Nucleic Acids Research. Plant journal **37**: D987-D991.

Okoniewski, M. J. and Miller, C. J. (2006) Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. BMC Bioinformatics **7**: 276.

Paxson-Sowders, D.M., Dodrill, C. H., Owen, H. A. and Makaroff, C. A. (2001) DEX1, a novel plant protein, is required for exine pattern formation during pollen development in *Arabidopsis*. Plant Physiology **127**: 1739-1749.

Peiffer, J. A, Kaushik, S., Sakai, H., Arteaga-Vazquez, M., Sanchez-Leon, N., Ghazal, H., Vielle-Calzada, J. P. and Meyers, B. C. (2008) A spatial dissection of the *Arabidopsis* floral transcriptome. Plant Biology **8**: 43.

Pelletier, G. and Budar, F. (2007) The molecular biology of cytoplasmically inherited male sterility and prospects for its engineering. Current Opinion in Biotechnology **18**: 121-125.

Peña, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., Zhao R. and Kutateladze, T. G. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature **442**: 100-103.

Peng, M., Bi, Y. M., Zhu, T., and Rothstein, S. J. (2007b) Genome-wide analysis of Arabidopsis responsive transcriptome to nitrogen limitation and its regulation by the ubiquitin ligase NLA. Plant Mol. Biol. **65**: 775-797.

Perehinec, T. M., Qazi, S. N., Gaddipati, S. R., Salisbury, V., Rees, C. E. and Hill**,** P. J. (2007) Construction and evaluation of multisite recombinatorial (Gateway) cloning vectors for Gram-positive bacteria. BMC Molecular Biology **8**: 1471-2199.

Perez-Amador, M. A., Lidder, P., Johnson, M. A., Landgraf, J., Wisman, E., and Green, P. J. (2001). New molecular phenotypes in the *dst* mutants of Arabidopsis revealed by DNA microarray analysis. The Plant Cell **13**: 2703-2717.

Pérez-Pérez, J. M., Candela, H. and Micol, J. L. (August 2009) Understanding synergy in genetic interactions. Trends in Genetics **25**: 368-76.

Phan, H. A., Iacuone, S., Li, S. F. and Parish, R.W. (2011) The MYB80 transcription factor is required for pollen development and the regulation of tapetal programmed cell death in Arabidopsis thaliana. Plant Cell **23**: 2209-24.

Piffanelli P., Ross J. H. E., Murphy D. J. (1998) Biogenesis and function of the lipidic structures of pollen grains. Sexual Plant Reproduction **11**: 65-80.

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**: 744-756.

Piotrzkowski N, Schillberg S, Rasche S (2012) Tackling heterogeneity: A leaf disc-based assay for the high-throughput screening of transient gene expression in tobacco. PLoS ONE **9**: e45803.

Poehlman, J. M. and Sleper, D.A. (1995) Breeding field crops. 4th edition. Iowa State University Press, Ames.

Primrose, S. B. and Twyman, R. M. (2006) Principles of gene manipulation and genomics, 6th edition: Blackwell Publishing, Oxford, UK

Qu, N., Gan, W., Bi, D., Xia, S., Li, X. and Zhang, Y. (2010) Two BTB proteins function redundantly as negative regulators of defense against pathogens in *Arabidopsis.* Botany **88**: 953-960.

Rackham, O. and Brown, C. M. (2004) Visualization of RNA-protein interactions in living cells: FMRP and IMP1 interact on mRNAs. The *EMBO Journal* **23**: 3346-3355.

Raghavan, V. (1989) mRNAs and a cloned histone gene are differentially expressed during anther and pollen development in rice (Oryza sativa L.). Journal of Cell Science **92**: 217-29.

Reinartz, J., Bruyns, E., Lin, J. Z., Burcham, T., Brenner, S., Bowen, B., Kramer, M. and Woychik, R. (2002) Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. Briefings in Functional Genomics and Proteomics **1**: 95-104.

Richards, F. M. (1958) On the Enzymic Activity of Subtilisin-Modified Ribonuclease. Proceedings of the National Academy of Sciences **44**: 162-166.

Riechmann, J. L.,Heard, J.,Martin, G.,Reuber, L.,Jiang, C.,Keddie, J.,Adam, L.,Pineda, O.,Ratcliffe, O. J.,Samaha, R. R.,Creelman, R.,Pilgrim, M.,Broun, P.,Zhang, J. Z.,Ghandehari, D.,Sherman, B. K.,Yu, G. (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science **290**: 2105.

Rieder, B. and Neuhaus, H. E. (2012) Identification of an *Arabidopsis* plasma membrane-located ATP transporter important for anther development. Plant Cell **23**:1932-44.

Rizzo, M. A., Davidson, M. W. and Piston, D. W. (2009a) Fluorescent protein tracking and detection: fluorescent protein structure and color variants. Cold Spring Harbor Protocols doi:10.1101/pdb.top63.

Rizzo, M. A., Davidson, M. W. and Piston, D. W. (2009b) Fluorescent protein tracking and detection: Applications using fluorescent proteins in living cells. Cold Spring Harbor Protocols doi:10.1101/pdb.top64.

Ros, A., Faupel, M., Mees, H., Oostrum. J., Ferrigno, R., Reymond, F., Michel, P., Rossier, J. S. and Girault, H. H. (2002) Protein purification by Off-Gel electrophoresis. Proteomics **2**: 151-156.

Rosellini, D., Pezzotti, M. and Veronesi, F. (2001) Characterization of transgenic male sterility in *alfalfa*. Euphytica **3**: 313-319.

Rowland, O., Lee, R., Franke, R., Schreiber, L. and Kunst, L. (2007) The *CER3* wax biosynthetic gene from *Arabidopsis thaliana* is allelic to *WAX2*/*YRE*/*FLP1*. FEBS Letters **581**:3538-44.

Royce, T.E., Rozowsky, J. S., Bertone, P., Samanta, M., Stolc, V., Weissman, S., Snyder, M. and Gerstein, M. (2005) Issues in the analysis of oligonucleotide tiling microarrays for transcript mapping. Trends in Genetics **21**: 466-475.

Sanders, D., Brownlee, C. and Harper, J. F. (1999) Communicating with calcium. The Plant Cell **11**: 691-706.

Saurin, W., Hofnung, M. and Dassa, E. (1999) Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. Journal of Molecular Evolution **48**: 22-41.

Schneider, E. and Hunke, S. (1998) ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolysing subunits/domains. FEMS Microbiology Reviews **22**: 1-20.

Schýtze1, K., Klaus Harter, K. and Chaban, C. (2009) Bimolecular Fluorescence Complementation (BiFC) to Study Protein-protein Interactions in Living Plant Cells Methods in Molecular Biology. Springer Protocols **479**: 1-14

Scott, R. J., Guerineau, A., Canales-Holzeis, F. and Dickinson C. (2002) A novel extinction screen in *Arabidopsis thaliana* identifies mutant plants defective in early microsporangial development. The Plant Journal **29**: 581-594

Séraphin, B., Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M. and Mann, M. (1999) A generic protein purification method for protein complex characterization and proteome exploration. Nature Biotechnology **17**: 1030-1032.

Sheoran, I. S., Olson, D. J. H., Ross A. R. S. and Sawhney, V. K. (2005) Proteome analysis of embryo and endosperm of germinating tomato seeds. Proteomics **5**: 3752-3764.

Sheoran, I. S., Ross, A. R. S., Olson, D. J. H. and Sawhney, V. K. (2007) Proteomic analysis of tomato (*Lycopersicon esculentum*) pollen. Journal of Experimental Botany **58**: 3525-3535.

Sheoran, I. S., Sproule, K. A, Olson, D. J. H., Ross, A.R.S. and Sawhney, V. K. (2006) Proteome profile and functional classification of proteins in *Arabidopsis thaliana* (*Landsberg erecta*) mature pollen. Sexual Plant Reproduction **19**:185-196.

Shi, J., Tan, H., Yu, X. H., Liu, Y., Liang, W., Ranathunge, K., Franke, R. B., Schreiber, L., Wang, Y., Kai, G., Shanklin, J. and Ma, Zhang D. (2011) *Defective Pollen Wall* is required for anther and microspore development in rice and encodes a fatty acyl carrier protein reductase. The Plant Cell **23**: 2225-2246.

Shi, Y.H., Zhu, S.W., Mao, X. Z., Feng, J.X., Qin, Y.M., Zhang, L., Cheng, J., Wei, L. P., Wang, Z. Y. and Zhu, Y. X. (2006) Transcriptome profiling, molecular

biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. The Plant Cell **18**: 651-664.

Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T., Fukuda, S., Sasaki, D., Podhajska, A., Harbers, M., Kawai, J., Carninci, P. and Hayashizaki, Y. (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proceedings of the National Academy of Sciences **26**: 15776-81.

Sorensen, A. M., Krober, S., Unte, U. S., Hujser, P., Dekker, K. and Saedler, H. (2003) The Arabidopsis Aborted Microspores (AMS) gene encodes a MYC class transcription factor. The Plant Journal **33**: 413-423.

Sparkes, I. and Brandizzi, F. (2012) Fluorescent protein-based technologies: shedding new light on the plant endomembrane system. The Plant Journal **70**: 96-107.

Steiner, F. A., Talbert, P. B., Kasinathan, S., Deal, R. B. and Henikoff, S. (2012) Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. Genome Research **22**: 766-77.

Studier, F. W. and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology **189**: 113-30.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods in Enzymology **185**: 60-89.

Suwabe, K., Suzuki, G., Takahashi, H., Shiono, K. and Endo, M. (2008) Separated transcriptomes of male gametophyte and tapetum in rice: validity of a laser microdissection (LM) microarray. Plant Cell Physiology **49**:1407-1416.

Suzuki, T., Masaoka, K., Nishi, M., Nakamura, K. and Ishiguro S. (2012) Identification of *kaonashi* mutants showing abnormal pollen exine structure in *Arabidopsis thaliana*. Plant Physiology **158**: 264-272.

Tang, L. K., Chu, H., Yip, W. K, Yeung, E. C. and Lo C. (2009) An anther-specific *dihydroflavonol 4-reductase-like* gene (*DRL1*) is essential for male fertility in *Arabidopsis*. New Phytology **181**:576-87.

The Royal Society (2009) Reaping the benefits: Science and the sustainable intensification of global agriculture. 72pp. ISBN: 978-0-85403-784-1.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L. A., Rhee, S. Y. and Stitt, M. (2004) MapMan: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant Journal **37**: 914-939.

Tsien, R. Y. (1998) The green fluorescent protein. Annual Review of Biochemistry **67**: 509-544.

Tsuda K, Qi Y, Nguyen le V, Bethke G, Tsuda Y, Glazebrook J, Katagiri F. (2012) An efficient Agrobacterium-mediated transient transformation of Arabidopsis. The Plant Journal **69**:713-9.

Varshney, R. K., Nayak, S. N., May, G. D. and Jackson S. A. (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends in Biotechnology 9: 522-530

Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K. (1995) Serial analysis of gene expression. Science **270**: 484-487.

Vidal, M., Brachmann, R. K., Fattaey, A., Harlow, E., and Boeke, J. D. (1996) Proceedings of the National Academy of Sciences **93**: 10315.

Vinod, K. K. (2005) Cytoplasmic genetic male sterility in plants: A molecular prospective. In: Proceedings of the training programme on "Advances and Accomplishments in Heterosis Breeding", Tamil Nadu Agricultural University, Coimbatore, India. 147-162.

Vizcay-Barrena, G. and Wilson, Z. A. (2006) Altered tapetal PCD and pollen wall development in the *Arabidopsis ms1* mutant. Journal of Experimental Botany **57**: 2709-2717.

Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. The Plant Journal **33**: 949-956.

Voulgaridou, G. P., Mantso, T., Chlichlia, K., Panayiotidis, M. I. and Pappa, A. (2013) Efficient *E. coli* expression strategies for production of soluble human crystallin ALDH3A1. PLoS ONE 8(2): e56582. doi:10.1371/journal.pone.0056582

Wang, C., Marshall, A., Zhang, D. and Wilson, Z. A. (2012) ANAP: an integrated knowledge base for *Arabidopsis* protein interaction network analysis. Plant Physiology **158**: 1523-1533.

Wang, j., Qiu, X., Li, Y., Deng, Y. and Shi, T. (2011) A transcriptional dynamic network during *Arabidopsis thaliana* pollen development. BMC Systems Biology **5**: S8.

Wang, X. T., Yuan, C., Yuan, T. T. and Cui S. J. (2012) The *Arabidopsis LFR* gene is required for the formation of anther cell layers and normal expression of key regulatory genes. Molecular Plant **5**: 993-1000.

Wang, y., Zhang, W.Z., Song, L.F., Zou, J.J., Su, Z. and Wu, W.H. (2008) Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in *Arabidopsis*. Plant Physiology **148**: 1201-1211.

Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics **10**: 57-63.

Wijeratne, A.J., Zhang, W., Sun, Y., Liu, W., Albert, R., Zheng, Z., Oppenheimer, D. G., Zhao, D. and Ma, H. (2007) Differential gene expression in Arabidopsis wild-type and mutant anthers: insights into anther cell differentiation and regulatory networks. The Plant Journal **52**: 14-29.

Wilson, Z. A, Song, J., Taylor, B. and Yang C. (2011) The final split: the regulation of anther dehiscence. Journal of Experimental Botany **62**: 1633-49.

Wilson, Z. A, Yang, C. (2004) Plant gametogenesis: conservation and contrasts in development. Reproduction **128**: 483-492.

Wilson, Z. A. (2001) 'The Arabidopsis MALE STERILITY1 (*MS1*) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. The Plant Journal **28**:27-39.

Wilson, Z. A. and Zhang, D. (2009) From *Arabidopsis* to rice: pathways in pollen development Journal of Experimental Botany **60**: 1479-1492

Wilson, Z. A., Morroll, S. M., Dawson, J., Swarup R. and Tighe P. J. (2001) The Arabidopsis MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. The Plant Journal **28**: 27-39.

Wu, R., Cao, J., Huang, Z., Wang, Z., Gai, J. and Vallejos E. (2011) Systems mapping: how to improve the genetic mapping of complex traits through design principles of biological systems. BMC Systems Biology **5**: 84.

Wuchty, S. (2002) Interaction and domain networks of yeast. Proteomics **2**:1715-1723.

Xing, S. and Zachgo, S. (2007) *ROXY1* and *ROXY2*, two Arabidopsis glutaredoxin genes, are required for anther development. The Plant Journal **53**:790-801.

Xu, J., Yang, C., Yuan, Z., Zhang, D., Gondwe, M.Y., Ding, Z., Liang, W., Zhang, D. and Wilson, Z. A. (2010) The ABORTED MICROSPORES regulatory network is required for postmeiotic male reproductive development in Arabidopsis thaliana. The Plant Cell **22**: 91-107.

Yamada, T., Tozawa, Y., Hasegawa, H., Terakawa, T., Ohkawa, Y. and Wakasa, K. (2004) Use of a feedback-insensitive a-subunit of anthranilate synthase as a selectable marker for transformation of rice and potato. Molecular Breeding **14**: 363-373.

Yang, C., Vizcay-Barrena, G., Conner, K. and Zoe A. Wilson (2007) MALE STERILITY1 is required for tapetal development and pollen wall biosynthesis. The Plant Cell **19**: 3530-3548.

Yang, C., Zheng X., Song, J., Conner, K., Barrena, G, V. and Wilson Z. A. (2007) *Arabidopsis* MYB26/MALE STERILE35 regulates secondary thickening in the endothecium and is essential for anther dehiscence. The Plant Cell **19**: 534-548.

Yuan, T., Fujioka, S., Takatsuto, S., Matsumoto, S., Gou, X., He, K., Russell, S. D. and Li J. (2007) BEN1, a gene encoding a dihydroflavonol 4-reductase (DFR)-like protein, regulates the levels of brassinosteroids in *Arabidopsis thaliana*. The Plant Journal **51**: 220-33.

Yuan, Y. R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P. J. and Hunt, J. F. (2001) The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. The Journal of Biological Chemistry **276**: 32313-21.

Zhang, W., Sun, Y., Timofejeva, L., Changbin, C., Grossniklaus, U. and Ma, H. (2006) Regulation of *Arabidopsis* tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) Encoding a Putative bHLH transcription factor. Development **16**: 3085-3095.

Zhang, W., Sun, Y., Timofejeva, L., Chen, C., Grossniklaus, U. and Ma, H. (2006) Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor. Development **133**: 3085-3095.

Zhang, Y., Liu, S., Dai, S. Y. and Yuan, J. S. (2012) Integration of shot-gun proteomics and bioinformatics analysis to explore plant hormone responses. BMC Bioinformatics **13**: S8.

Zhang, Z. B., Zhu, J., Gao, J. F., Wang, C. and Li H. (2007) Transcription factor *AtMYB103* is required for anther development by regulating tapetum development, callose dissolution and exine formation in *Arabidopsis*. The Plant Journal **52**: 528-38.

Zheng L, Liu G, Meng X, Li Y, Wang Y. (2012) A versatile Agrobacterium-mediated transient gene expression system for herbaceous plants and trees. Biochemical Genetics **50**: 761-9.

Zhou, Z., Dun, X., Xia, S., Shi, D., Qin, M., Yi, B., Wen, J., Shen, J., Ma, C., Tu, J. and Fu, T. (2012) BnMs3 is required for tapetal differentiation and degradation, microspore separation, and pollen-wall biosynthesis in Brassica napus. Journal of Experimental Botany **63**: 2041-58.

Zhu, J., Chen H., Li, H., Gao, J. F., Jiang, H., Wang, C., Guan, Y-F. and Yang Z-N. (2008) Defective in tapetal development and function 1 is essential for anther development and tapetal function for microspore maturation in Arabidopsis. The Plant Journal **55**: 266–277

Zhu, J., Lou, Y., Xu, X. and Yang, Z. N. (2011) A genetic pathway for tapetum development and function in Arabidopsis. Journal of Integrative Plant Biology **53**: 892-900.

Zhu, J., Zhang, G., Chang, Y., Li, X., Yang, J., Huang, X., Yu, Q., Chen, H., Wu, T. and Yang, Z. (2010) AtMYB103 is a crucial regulator of several pathways affecting Arabidopsis anther development. Science China Life Sciences **53**:1112-22.

CHAPTER 8

APPENDIX

8. APPENDIX

8.1 DNA MOLECULAR MARKERS (Bioline.com)

LADDER	SEPARATION RANGE (bp)	LOADING DYE COLOUR	HIGH INTENSITY BANDS (bp)
HyperLadder™ 1kb (formerly HyperLadder I)	200 - 10,037	Blue	1,000 & 10,000
HyperLadder™ 50bp (formerly HyperLadder II)	50 - 2,000	Blue	300, 1,000 & 3,500
HyperLadder™ 25bp (formerly HyperLadder V)	25 - 500	Blue	100 & 200

8.2 IMAGES OF DNA MOLECULAR MARKERS UTILISED (Bioline.com)

HyperLadder I

SIZE (bp)	ng/BAND
10037	100
8000	80
6000	60
5000	50
4000	40
3000	30
2500	25
2000	20
1500/1517	15/15
1000	100
800	80
600	60
400	40
200	20

1% agarose gel
5μl per lane

HyperLadder II

SIZE (bp)	ng/BAND
2000	50
1800	20
1600	20
1400	20
1200	20
1000	100
800	30
700	30
600	30
500	30
400	30
300	100
200	40
100	40
50	40

1.5% agarose gel
5μl per lane

HyperLadder V

SIZE (bp)	ng/BAND
500	60
400	60
300	60
250	60
200	120
175	80
150	80
125	80
100	120
75	80
50	80
25	80

3.5% agarose gel
5μl per lane

244

8.3 PROTEIN LADDER

SDS-PAGE band profile of the Thermo Scientific Spectra Multicolor Broad Range Protein Ladder. The Image is from a 4-20% Tris-glycine gel (SDS-PAGE) and subsequent transfer to membrane.
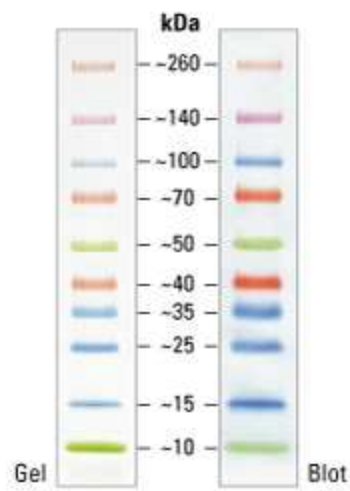
8.4 TABLE OF X² VALUE VERSUS P-VALUE

The table below gives a number of p-values matching to χ^2 for the first 10 degrees of freedom. A p-value of 0.05 or less is usually regarded as statistically significant, i.e. the observed deviation from the null hypothesis is significant.

Degrees of freedom (df)	χ^2 value										
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88
10	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59
P value (Probability)	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001
	Nonsignificant								Significant		

MS1 cDNA Assembly-CLONE3
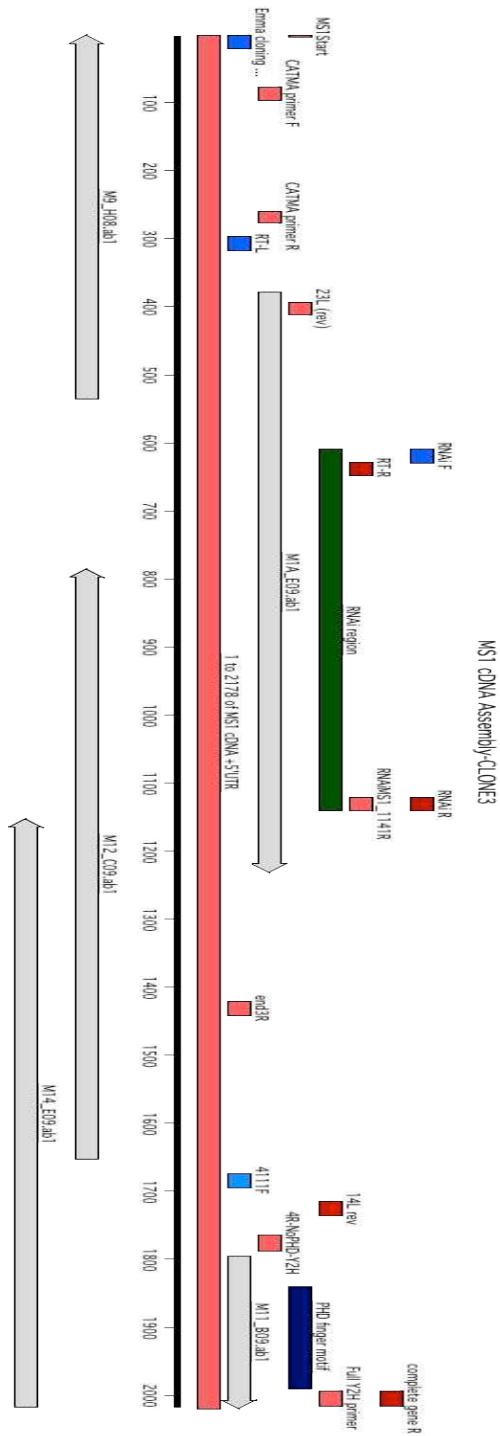
Formatted Alignments

```
Translation_of_Consensus_of_MS1_cDNA_Alignments-2    1  M A N L I R T D H Q Q H I P K K R K R G E S R V F R L K T F G E S G H   35
Translation_of_MS1_cDNA                              1  M A N L I R T D H Q Q H I P K K R K R G E S R V F R L K T F G E S G H   35
                                                        M A N L I R T D H Q Q H I P K K R K R G E S R V F R L K T F G E S G H

Translation_of_Consensus_of_MS1_cDNA_Alignments-2   36  P A E M N E L S F R D N L A K L L E F G H F E S S G L M G S W S F Q L   70
Translation_of_MS1_cDNA                             36  P A E M N E L S F R D N L A K L L E F G H F E S S G L M G S W S F Q L   70
                                                        P A E M N E L S F R D N L A K L L E F G H F E S S G L M G S W S F Q L

Translation_of_Consensus_of_MS1_cDNA_Alignments-2   71  E I Q R N P N P L Y V L L F V V E E P I E A S L N L R C N H C Q Y V G  105
Translation_of_MS1_cDNA                             71  E I Q R N P N P L Y V L L F V V E E P I E A S L N L R C N H C Q Y V G  105
                                                        E I Q R N P N P L Y V L L F V V E E P I E A S L N L R C N H C Q Y V G

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  106  W G N Q M I C N K K Y H F V I P S K E T M A A F L K L E G G G Y A F P  140
Translation_of_MS1_cDNA                            106  W G N Q M I C N K K Y H F V I P S K E T M A A F L K L E G G G Y A F P  140
                                                        W G N Q M I C N K K Y H F V I P S K E T M A A F L K L E G G G Y A F P

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  141  E K E S F S H L V E L Q G H V L H G F F H S N G F G H L L S L N G I E  175
Translation_of_MS1_cDNA                            141  E K E S F S H L V E L Q G H V L H G F F H S N G F G H L L S L N G I E  175
                                                        E K E S F S H L V E L Q G H V L H G F F H S N G F G H L L S L N G I E

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  176  T G S D L T G H Q V M D L W D R L C T G L K A R K I G L N D A S H K K  210
Translation_of_MS1_cDNA                            176  T G S D L T G H Q V M D L W D R L C T G L K A R K I G L N D A S H K K  210
                                                        T G S D L T G H Q V M D L W D R L C T G L K A R K I G L N D A S H K K

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  211  G M E L R L L H G V A K G E P W F G R W G Y R F G S G T Y G V T Q K I  245
Translation_of_MS1_cDNA                            211  G M E L R L L H G V A K G E P W F G R W G Y R F G S G T Y G V T Q K I  245
                                                        G M E L R L L H G V A K G E P W F G R W G Y R F G S G T Y G V T Q K I

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  246  Y E K A L E S V R N I P L C L L N H H L T S L N R E T P I L L S K Y Q  280
Translation_of_MS1_cDNA                            246  Y E K A L E S V R N I P L C L L N H H L T S L N R E T P I L L S K Y Q  280
                                                        Y E K A L E S V R N I P L C L L N H H L T S L N R E T P I L L S K Y Q

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  281  S L S T E P L I T L S D L F R F M L H L H S R L P R D N Y M S N S R N  315
Translation_of_MS1_cDNA                            281  S L S T E P L I T L S D L F R F M L H L H S R L P R D N Y M S N S R N  315
                                                        S L S T E P L I T L S D L F R F M L H L H S R L P R D N Y M S N S R N

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  316  Q I I S I D S T N C R W S Q K R I Q M A I K V V I E S L K R V E Y R W  350
Translation_of_MS1_cDNA                            316  Q I I S I D S T N C R W S Q K R I Q M A I K V V I E S L K R V E Y R W  350
                                                        Q I I S I D S T N C R W S Q K R I Q M A I K V V I E S L K R V E Y R W

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  351  I S R Q E V R D A A R N Y I G D T G L L D F V L K S L G N Q V V G N Y  385
Translation_of_MS1_cDNA                            351  I S R Q E V R D A A R N Y I G D T G L L D F V L K S L G N Q V V G N Y  385
                                                        I S R Q E V R D A A R N Y I G D T G L L D F V L K S L G N Q V V G N Y

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  386  L V R R S L N P V K K V L E Y S L E D I S N L L P S G N N E L I T L Q  420
Translation_of_MS1_cDNA                            386  L V R R S L N P V K K V L E Y S L E D I S N L L P S G N N E L I T L Q  420
                                                        L V R R S L N P V K K V L E Y S L E D I S N L L P S G N N E L I T L Q

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  421  N Q N S M G K M A T N G H N K I T R G Q V M K D M F Y F Y K H I L M D  455
Translation_of_MS1_cDNA                            421  N Q N S M G K M A T N G H N K I T R G Q V M K D M F Y F Y K H I L M D  455
                                                        N Q N S M G K M A T N G H N K I T R G Q V M K D M F Y F Y K H I L M D

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  456  Y K G V L G P I G G I L N Q I G M A S R A I L D A K Y F I K E Y H Y I  490
Translation_of_MS1_cDNA                            456  Y K G V L G P I G G I L N Q I G M A S R A I L D A K Y F I K E Y H Y I  490
                                                        Y K G V L G P I G G I L N Q I G M A S R A I L D A K Y F I K E Y H Y I

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  491  R D T S A K T L H L D R G E E L G I F C T I A W K C H H H N N E I K V  525
Translation_of_MS1_cDNA                            491  R D T S A K T L H L D R G E E L G I F C T I A W K C H H H N N E I K V  525
                                                        R D T S A K T L H L D R G E E L G I F C T I A W K C H H H N N E I K V

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  526  P P Q E C I V V K K D A T L S E V Y G E A E R V F R D I Y W E L R D V  560
Translation_of_MS1_cDNA                            526  P P Q E C I V V K K D A T L S E V Y G E A E R V F R D I Y W E L R D V  560
                                                        P P Q E C I V V K K D A T L S E V Y G E A E R V F R D I Y W E L R D V

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  561  V V E S V V G G Q I E I T R V D E M A L N G N K G L V L E G N V G M M  595
Translation_of_MS1_cDNA                            561  V V E S V V G G Q I E I T R V D E M A L N G N K G L V L E G N V G M M  595
                                                        V V E S V V G G Q I E I T R V D E M A L N G N K G L V L E G N V G M M

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  596  M N I E V T K C Y E D D D K K K D K R I E C E C G A T E E D G E R M V  630
Translation_of_MS1_cDNA                            596  M N I E V T K C Y E D D D K K K D K R I E C E C G A T E E D G E R M V  630
                                                        M N I E V T K C Y E D D D K K K D K R I E C E C G A T E E D G E R M V

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  631  C C D I C E V W Q H T R C V G V Q H N E E V P R I F L C Q S C D Q H L  665
Translation_of_MS1_cDNA                            631  C C D I C E V W Q H T R C V G V Q H N E E V P R I F L C Q S C D Q H L  665
                                                        C C D I C E V W Q H T R C V G V Q H N E E V P R I F L C Q S C D Q H L

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  666  I P L S F L P X P A F L Y K V V F D P G S G C Q S P K G S V G C C H R  700
Translation_of_MS1_cDNA                            666  I P L S F L P - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        I P L S F L P X P A F L Y K V V F D P G S G C Q S P K G S V G C C H R

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  701  A I T S I T P W G L T G L E G F F A E R R N Y I R I S T G R V W S P S  735
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        A I T S I T P W G L T G L E G F F A E R R N Y I R I S T G R V W S P S

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  736  R S R W L Q V A K R A G L G G G Q S G R T V L R E R V R I E I A S T H  770
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        R S R W L Q V A K R A G L G G G Q S G R T V L R E R V R I E I A S T H

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  771  I A L A A R H S D W R C C R N G R Y P A E A R Q Y R H N Q A Y A Y S I  805
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        I A L A A R H S D W R C C R N G R Y P A E A R Q Y R H N Q A Y A Y S I

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  806  Q G D G A E D D D E R I V R F H T P G L T A L A I S D E M L R K E G E  840
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        Q G D G A E D D D E R I V R F H T P G L T A L A I S D E M L R K E G E

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  841  N R V V W S N G R R W R E N G V L Y L S M A T H K V C W C S T Q G S A  875
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        N R V V W S N G R R W R E N G V L Y L S M A T H K V C W C S T Q G S A

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  876  S H F S L S K L S T S Y S S L F F T R P S F L V Q S G L I R P G I R L  910
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        S H F S L S K L S T S Y S S L F F T R P S F L V Q S G L I R P G I R L

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  911  L T K P E R K L S W L L P P L S N N H N P L G P L N G S G V F C K E E  945
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        L T K P E R K L S W L L P P L S N N H N P L G P L N G S G V F C K E E

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  946  L Y P D I H R T G V V A M I A S I V A P S S E A S R T G R R P K R S D  980
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        L Y P D I H R T G V V A M I A S I V A P S S E A S R T G R R P K R S D

Translation_of_Consensus_of_MS1_cDNA_Alignments-2  981  S A P R T G A H R N C I N A Y S A S S T P L A M L S E W T I S R R G P 1015
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  672
                                                        S A P R T G A H R N C I N A Y S A S S T P L A M L S E W T I S R R G P

Translation_of_Consensus_of_MS1_cDNA_Alignments-2 1016  A V P A P S L C L Q H P G R C R G R A H C I S Y T G P D C V S N L   L1048
Translation_of_MS1_cDNA                            673  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -    -672
                                                        A V P A P S L C L Q H P G R C R G R A H C I S Y T G P D C V S N L
```

248

```
M9_H08-ZW_REV    1  GACTTCATGTTGTATGACGCTCTTGATGTTGTTTT   35
MS1_cDNA         1  -----------------------------------    0
                    G A C T T C A T G T T G T A T G A C G C T C T T G A T G T T G T T T T

M9_H08-ZW_REV   36  ATACATGGACCCAATGTGCCTGGATGCGTTCCCAA   70
MS1_cDNA         1  -----------------------------------    0
                    A T A C A T G G A C C C A A T G T G C C T G G A T G C G T T C C C A A

M9_H08-ZW_REV   71  AATTAGTTTGTTTTAAAAAACGTATTGAAGCTATC  105
MS1_cDNA         1  -----------------------------------    0
                    A A T T A G T T T G T T T T A A A A A A C G T A T T G A A G C T A T C

M9_H08-ZW_REV  106  CCACAAATTGATAAGTACTTGAAATCCAGCAAGTA  140
MS1_cDNA         1  -----------------------------------    0
                    C C A C A A A T T G A T A A G T A C T T G A A A T C C A G C A A G T A

M9_H08-ZW_REV  141  TATAGCATGGCCTNNGCAGGGCTGGCAAGCCACGT  175
MS1_cDNA         1  -----------------------------------    0
                    T A T A G C A T G G C C T     G C A G G G C T G G C A A G C C A C G T

M9_H08-ZW_REV  176  TTGGTGGTGGCGACCATCCTCCAAAATCGGATCTG  210
MS1_cDNA         1  -----------------------------------    0
                    T T G G T G G T G G C G A C C A T C C T C C A A A A T C G G A T C T G

M9_H08-ZW_REV  211  GTTCCGCGTCCATGGTCGAATCAAACAAGTTTGTA  245
MS1_cDNA         1  -----------------------------------    0
                    G T T C C G C G T C C A T G G T C G A A T C A A A C A A G T T T G T A

M9_H08-ZW_REV  246  CAAAAAAGCAGGCTTCTCAAGATCAAACCAAATTT  280
MS1_cDNA         1  -----------------------------------    0
                    C A A A A A A G C A G G C T T C T C A A G A T C A A A C C A A A T T T

M9_H08-ZW_REV  281  CTGATTCGATGGCGAATCTGATTCGAACAGACCAT  315
MS1_cDNA         1  -------ATGGCGAATCTGATTCGAACAGACCAT   27
                    C T G A T T C G A T G G C G A A T C T G A T T C G A A C A G A C C A T

M9_H08-ZW_REV  316  CAGCAACATATTCCAAAGAAGAGGAAGAGAGGGGA  350
MS1_cDNA        28  CAGCAACATATTCCAAAGAAGAGGAAGAGAGGGGA   62
                    C A G C A A C A T A T T C C A A A G A A G A G G A A G A G A G G G G A

M9_H08-ZW_REV  351  AAGTAGAGTTTTTAGGCTGAAGACGTTCGGAGAGT  385
MS1_cDNA        63  AAGTAGAGTTTTTAGGCTGAAGACGTTCGGAGAGT   97
                    A A G T A G A G T T T T T A G G C T G A A G A C G T T C G G A G A G T

M9_H08-ZW_REV  386  CTGGACATCCAGCTGAGATGAACGAGTTGTCTTTT  420
MS1_cDNA        98  CTGGACATCCAGCTGAGATGAACGAGTTGTCTTTT  132
                    C T G G A C A T C C A G C T G A G A T G A A C G A G T T G T C T T T T

M9_H08-ZW_REV  421  CGAGATAACCTCGCTAAACTACTTGAGTTTGGTCA  455
MS1_cDNA       133  CGAGATAACCTCGCTAAACTACTTGAGTTTGGTCA  167
                    C G A G A T A A C C T C G C T A A A C T A C T T G A G T T T G G T C A

M9_H08-ZW_REV  456  CTTTGAGAGCTCCGGTCTAATGGGAAGTTGGTCTT  490
MS1_cDNA       168  CTTTGAGAGCTCCGGTCTAATGGGAAGTTGGTCTT  202
                    C T T T G A G A G C T C C G G T C T A A T G G G A A G T T G G T C T T

M9_H08-ZW_REV  491  TTCAGCTCGAGATTCAACGAAATCCAAATCCTCTC  525
MS1_cDNA       203  TTCAGCTCGAGATTCAACGAAATCCAAATCCTCTC  237
                    T T C A G C T C G A G A T T C A A C G A A A T C C A A A T C C T C T C

M9_H08-ZW_REV  526  TATGTTCTTCTCTTTGTCGTAGAAGAGCCCATCGA  560
MS1_cDNA       238  TATGTTCTTCTCTTTGTCGTAGAAGAGCCCATCGA  272
                    T A T G T T C T T C T C T T T G T C G T A G A A G A G C C C A T C G A

M9_H08-ZW_REV  561  AGCCTCTCTCAATCTCCGTTGCAACCATTGCCAAT  595
MS1_cDNA       273  AGCCTCTCTCAATCTCCGTTGCAACCATTGCCAAT  307
                    A G C C T C T C T C A A T C T C C G T T G C A A C C A T T G C C A A T

M9_H08-ZW_REV  596  ATGTTGGTTGGGGAAATCAAATGATATGCAACAAG  630
MS1_cDNA       308  ATGTTGGTTGGGGAAATCAAATGATATGCAACAAG  342
                    A T G T T G G T T G G G G A A A T C A A A T G A T A T G C A A C A A G

M9_H08-ZW_REV  631  AAGTACCATTTCGTGATCCCCTCAAAGGAAACAAT  665
MS1_cDNA       343  AAGTACCATTTCGTGATCCCCTCAAAGGAAACAAT  377
                    A A G T A C C A T T T C G T G A T C C C C T C A A A G G A A A C A A T

M9_H08-ZW_REV  666  GGCGGCTTTTTTAAAACTGGAAGGTGGAGGCTACG  700
MS1_cDNA       378  GGCGGCTTTTTTAAAACTGGAAGGTGGAGGCTACG  412
                    G G C G G C T T T T T T A A A A C T G G A A G G T G G A G G C T A C G

M9_H08-ZW_REV  701  CTTTTCCCGAAAAGGAAAGTTTCTCCCATCTTGTG  735
MS1_cDNA       413  CTTTTCCCGAAAAGGAAAGTTTCTCCCATCTTGTG  447
                    C T T T T C C C G A A A A G G A A A G T T T C T C C C A T C T T G T G

M9_H08-ZW_REV  736  GAGCTTCAAGGCCATGTCCTTCACGGCTTCTTTCA  770
MS1_cDNA       448  GAGCTTCAAGGCCATGTCCTTCACGGCTTCTTTCA  482
                    G A G C T T C A A G G C C A T G T C C T T C A C G G C T T C T T T C A

M9_H08-ZW_REV  771  CTCCAACGGATTTGGTCACTTGCTCTCTCTCAACG  805
MS1_cDNA       483  CTCCAACGGATTTGGTCACTTGCTCTCTCTCAACG  517
                    C T C C A A C G G A T T T G G T C A C T T G C T C T C T C T C A A C G

M9_H08-ZW_REV  806  GCATTGAAACCGGCTCCGAN-----------------  825
MS1_cDNA       518  GCATTGAAACCGGCTCCGACTTAACCGGTCATCAG  552
                    G C A T T G A A A C C G G C T C C G A C T T A A C C G G T C A T C A G

M9_H08-ZW_REV  826  -----------------------------------  825
MS1_cDNA       553  GTCATGGATTTGTGGGACCGGCTCTGCACCGGTTT  587
                    G T C A T G G A T T T G T G G G A C C G G C T C T G C A C C G G T T T
```

249

```
Translation_of_MS1_cDNA        1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -    0
Translation_of_M9_H08-ZW_REV   1   D F M L Y D A L D V V L Y M D P M C L D A F P K L V C F K K R I E A I   35
                                   D F M L Y D A L D V V L Y M D P M C L D A F P K L V C F K K R I E A I

Translation_of_MS1_cDNA        1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -    0
Translation_of_M9_H08-ZW_REV   36  P Q I D K Y L K S S K Y I A W P X Q G W Q A T F G G G D H P P K S D L   70
                                   P Q I D K Y L K S S K Y I A W P X Q G W Q A T F G G G D H P P K S D L

Translation_of_MS1_cDNA        1   - - - - - - - - - - - - - - - - - - - - - - - - M A N L I R T D H    9
Translation_of_M9_H08-ZW_REV   71  V P R P W S N Q T S L Y K K A G F S R S N Q I S D S M A N L I R T D H   105
                                   V P R P W S N Q T S L Y K K A G F S R S N Q I S D S M A N L I R T D H

Translation_of_MS1_cDNA        10  Q Q H I P K K R K R G E S R V F R L K T F G E S G H P A E M N E L S F   44
Translation_of_M9_H08-ZW_REV   106 Q Q H I P K K R K R G E S R V F R L K T F G E S G H P A E M N E L S F   140
                                   Q Q H I P K K R K R G E S R V F R L K T F G E S G H P A E M N E L S F

Translation_of_MS1_cDNA        45  R D N L A K L L E F G H F E S S G L M G S W S F Q L E I Q R N P N P L   79
Translation_of_M9_H08-ZW_REV   141 R D N L A K L L E F G H F E S S G L M G S W S F Q L E I Q R N P N P L   175
                                   R D N L A K L L E F G H F E S S G L M G S W S F Q L E I Q R N P N P L

Translation_of_MS1_cDNA        80  Y V L L F V V E E P I E A S L N L R C N H C Q Y V G W G N Q M I C N K   114
Translation_of_M9_H08-ZW_REV   176 Y V L L F V V E E P I E A S L N L R C N H C Q Y V G W G N Q M I C N K   210
                                   Y V L L F V V E E P I E A S L N L R C N H C Q Y V G W G N Q M I C N K

Translation_of_MS1_cDNA        115 K Y H F V I P S K E T M A A F L K L E G G G Y A F P E K E S F S H L V   149
Translation_of_M9_H08-ZW_REV   211 K Y H F V I P S K E T M A A F L K L E G G G Y A F P E K E S F S H L V   245
                                   K Y H F V I P S K E T M A A F L K L E G G G Y A F P E K E S F S H L V

Translation_of_MS1_cDNA        150 E L Q G H V L H G F F H S N G F G H L L S L N G I E T G S D L T G H Q   184
Translation_of_M9_H08-ZW_REV   246 E L Q G H V L H G F F H S N G F G H L L S L N G I E T G S X - - - - -   275
                                   E L Q G H V L H G F F H S N G F G H L L S L N G I E T G S   L T G H Q

Translation_of_MS1_cDNA        185 V M D L W D R L C T G L K A R K I G L N D A S H K K G M E L R L L H G   219
Translation_of_M9_H08-ZW_REV   276 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   275
                                   V M D L W D R L C T G L K A R K I G L N D A S H K K G M E L R L L H G

Translation_of_MS1_cDNA        220 V A K G E P W F G R W G Y R F G S G T Y G V T Q K I Y E K A L E S V R   254
Translation_of_M9_H08-ZW_REV   276 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   275
```

250
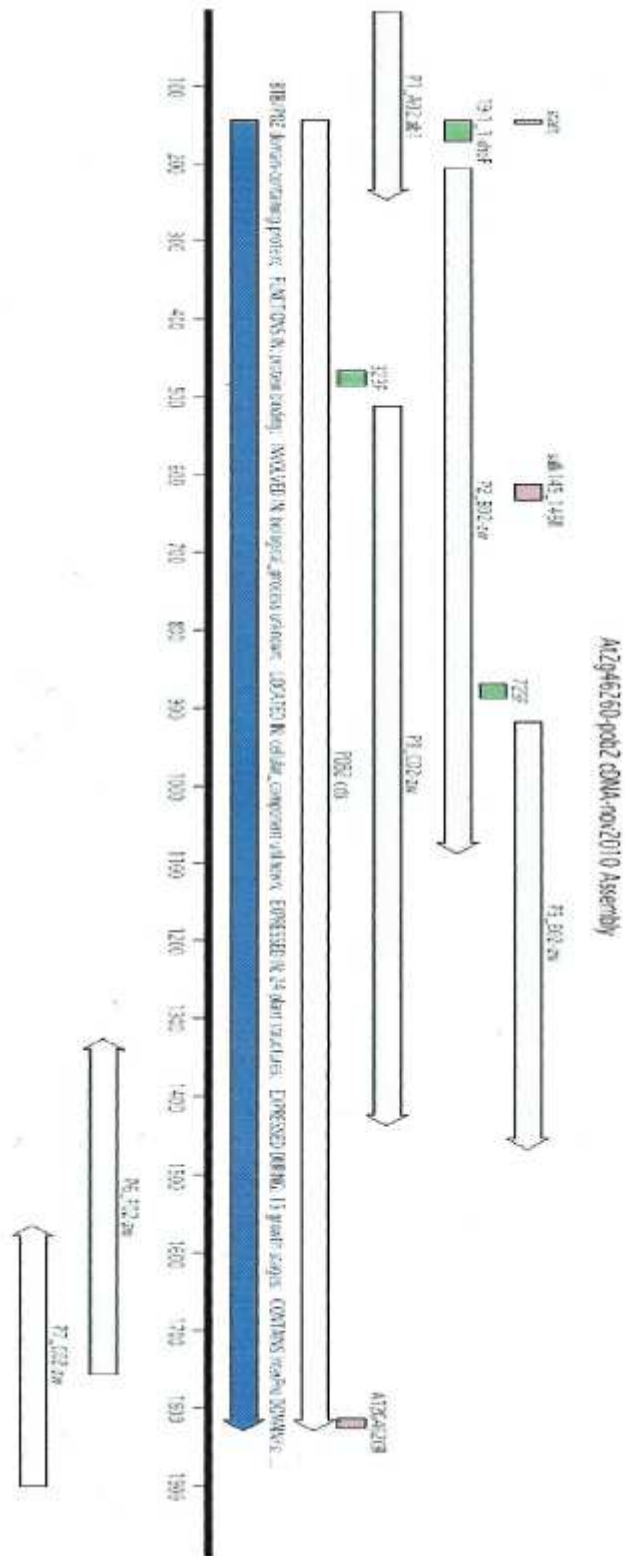
Formatted Alignments

```
POB2                    1   M R G S N N T D L F D P K T E M D S N F S R H G S S S E G D F G F A F   35
pDEST17-POB23'region    1   M R G S N N T D L F D P K T E M D S N F S R H G S S S E G D F G F A F   35
                            M R G S N N T D L F D P K T E M D S N F S R H G S S S E G D F G F A F

POB2                   36   N D S N F S D R L L R I E I L G G P S D S R S D A E G C T S I A D W A   70
pDEST17-POB23'region   36   N D S N F S D R L L R I E I L G G P S D S R S D A E G C T S I A D W A   70
                            N D S N F S D R L L R I E I L G G P S D S R S D A E G C T S I A D W A

POB2                   71   R H R K R R R E D N K K D N G V A I S D I V A C A E E Q I L T D N N Q   105
pDEST17-POB23'region   71   R H R K R R R E D N K K D N G V A I S D I V A C A E E Q I L T D N N Q   105
                            R H R K R R R E D N K K D N G V A I S D I V A C A E E Q I L T D N N Q

POB2                  106   P D M D D A P G G D N L D D E G E A M V E E A L S G D D D A S S E P N   140
pDEST17-POB23'region  106   P D M D D A P G G D N L D D E G E A M V E E A L S G D D D A S S E P N   140
                            P D M D D A P G G D N L D D E G E A M V E E A L S G D D D A S S E P N

POB2                  141   W G I D C S T V V R V K E L H I S S P I L A A K S P F F Y K L F S N G   175
pDEST17-POB23'region  141   W G I D C S T V V R V K E L H I S S P I L A A K S P F F Y K L F S N G   175
                            W G I D C S T V V R V K E L H I S S P I L A A K S P F F Y K L F S N G

POB2                  176   M R E S E Q R H V T L R I S A Q E E G A L M E L L N F M Y S N S L S V   210
pDEST17-POB23'region  176   M R E S E Q R H V T L R I S A Q E E G A L M E L L N F M Y S N S L S V   210
                            M R E S E Q R H V T L R I S A Q E E G A L M E L L N F M Y S N S L S V

POB2                  211   T T A P A L L D V L M A A D K F E V A S C M R Y C S R L L R N M P M T   245
pDEST17-POB23'region  211   T T A P A L L D V L M A A D K F E V A S C M R Y C S R L L R N M P M T   245
                            T T A P A L L D V L M A A D K F E V A S C M R Y C S R L L R N M P M T

POB2                  246   P D S A L L Y L E L P S S V L M A E A V Q P L T D A A K Q F L A S R Y   280
pDEST17-POB23'region  246   P D S A L L Y L E L P S S V L M A E A V Q P L T D A A K Q F L A S R Y   280
                            P D S A L L Y L E L P S S V L M A E A V Q P L T D A A K Q F L A S R Y

POB2                  281   K D I T K F H D E V M A L P L A G I E A I L S S D D L Q I A S E D A V   315
pDEST17-POB23'region  281   K D I T K F H D E V M A L P L A G I E A I L S S D D L Q I A S E D A V   315
                            K D I T K F H D E V M A L P L A G I E A I L S S D D L Q I A S E D A V

POB2                  316   Y D F V L K W A R G Q Y S S L E D R R E I L G S R L A L Y I R F P Y M   350
pDEST17-POB23'region  316   Y D F V L K W A R G Q Y S S L E D R R E I L G S R L A L Y I R F P Y M   350
                            Y D F V L K W A R G Q Y S S L E D R R E I L G S R L A L Y I R F P Y M

POB2                  351   T C R K L K K V L T C S D F E H E V A S K Q V L E A L F F K A E A P H   385
pDEST17-POB23'region  351   T C R K L K K V L T C S D F E H E V A S K Q V L E A L F F K A E A P H   385
                            T C R K L K K V L T C S D F E H E V A S K Q V L E A L F F K A E A P H

POB2                  386   R Q R I L A A E G S D S M N R R F I E R A Y K Y R P V K V V E F E L P   420
pDEST17-POB23'region  386   R Q R I L A A E G S D S M N R R F I E R A Y K Y R P V K V V E F E L P   420
                            R Q R I L A A E G S D S M N R R F I E R A Y K Y R P V K V V E F E L P

POB2                  421   R P Q C V V Y L D L K R E E C A G L F P S G R V Y S Q A F H L G G Q G   455
pDEST17-POB23'region  421   R P Q C V V Y L D L K R E E C A G L F P S G R V Y S Q A F H L G G Q G   455
                            R P Q C V V Y L D L K R E E C A G L F P S G R V Y S Q A F H L G G Q G

POB2                  456   F F L S A H C N M D Q Q S S F H C F G L F L G M Q E K G A V S F G V D   490
pDEST17-POB23'region  456   F F L S A H C N M D Q Q S S F H C F G L F L G M Q E K G A V S F G V D   490
                            F F L S A H C N M D Q Q S S F H C F G L F L G M Q E K G A V S F G V D

POB2                  491   Y E F A A R D K S T K E E Y V S K Y K G N Y T F T G G K A V G Y R N L   525
pDEST17-POB23'region  491   Y E F A A R D K S T K E E Y V S K Y K G N Y T F T G G K A V G Y R N L   525
                            Y E F A A R D K S T K E E Y V S K Y K G N Y T F T G G K A V G Y R N L

POB2                  526   F G I P W T S F I A E D S Q H F I N G I L H L R A E L T I K R S S D L   560
pDEST17-POB23'region  526   F G I P W T S F I A E D S Q H F I N G I L H L R A E L T I K R S S D L   560
                            F G I P W T S F I A E D S Q H F I N G I L H L R A E L T I K R S S D L

POB2                  561   H K G G R A D P A F L Y K V L I R G C * Q S P K G S   586
pDEST17-POB23'region  561   H - - - - - - - - - - - - - - - - - - - - - - - - -   561
                            H K G G R A D P A F L Y K V L I R G C   Q S P K G S
```

252

8.7 MS1 (AT5G22260) FULL LENGTH CDNA

1ACTTTTCGGT CTCTTCTTAT AACAAACCCC ATCATAATAT ATATCCAAAT
51 CTATCATATT CCTATTCCCT CTCTTCTCAA GATCAAACCA AATTTCTGAT
101 TCGATGGCGA ATCTGATTCG AACAGACCAT CAGCAACATA TTCCAAAGAA
151 GAGGAAGAGA GGGGAAAGTA GAGTTTTTAG GCTGAAGACG TTCGGAGAGT
201 CTGGACATCC AGCTGAGATG AACGAGTTGT CTTTTCGAGA TAACCTCGCT
251 AAACTACTTG AGTTTGGTCA CTTTGAGAGC TCCGGTCTAA TGGGAAGTTG
301 GTCTTTTCAG CTCGAGATTC AACGAAATCC AAATCCTCTC TATGTTCTTC
351 TCTTTGTCGT AGAAGAGCCC ATCGAAGCCT CTCTCAATCT CCGTTGCAAC
401 CATTGCCAAT ATGTTGGTTG GGGAAATCAA ATGATATGCA ACAAGAAGTA
451 CCATTTCGTG ATCCCCTCAA AGGAAACAAT GGCGGCTTTT TTAAAACTGG
501 AAGGTGGAGG CTACGCTTTT CCCGAAAAGG AAAGTTTCTC CCATCTTGTG
551 GAGCTTCAAG GCCATGTCCT TCACGGCTTC TTTCACTCCA ACGGATTTGG
601 TCACTTGCTC TCTCTCAACG GCATTGAAAC CGGCTCCGAC TTAACCGGTC
651 ATCAGGTCAT GGATTTGTGG GACCGGCTCT GCACCGGTTT AAAGGCCAGG
701 AAAATAGGGT TGAATGACGC GTCGCACAAA AAAGGAATGG AGTTGAGGCT
751 GCTGCATGGG GTAGCAAAAG GAGAGCCATG GTTCGGTCGT TGGGGCTACC
801 GGTTCGGGTC AGGGACATAC GGAGTGACTC AAAAGATTTA CGAGAAGGCA
851 CTTGAGTCGG TCCGTAACAT ACCCTTGTGC TTGCTTAACC ATCACCTAAC
901 CAGCCTTAAC CGAGAAACTC CAATCCTCTT GTCAAAGTAC CAAAGTTTAT
951 CCACCGAGCC ATTGATCACT CTCAGTGACC TCTTCAGGTT CATGCTTCAT
1001 CTCCATTCAC GTCTTCCAAG AGATAACTAC ATGAGTAACT CCCGAAACCA
1051 AATCATCTCC ATTGATAGTA CCAACTGCAG ATGGTCTCAA AAACGGATCC
1101 AAATGGCTAT CAAAGTGGTC ATAGAGTCAC TGAAAAGAGT CGAATACCGA
1151 TGGATATCGA GACAAGAAGT GAGGGATGCA GCTAGAAATT ACATTGGGGA
1201 CACTGGTTTG CTTGATTTTG TGTTGAAGTC GCTTGGGAAC CAGGTGGTTG
1251 GAAACTATTT GGTCCGACGT AGTCTAAATC CGGTGAAGAA AGTGCTAGAG
1301 TATTCCTTGG AGGATATATC AAATTTGTTA CCAAGTAGTA ACAATGAACT
1351 CATAACCCTT CAAAACCAAA ACTCAATGGG GAAGATGGCG ACAAACGGTC
1401 ACAATAAGAT CACAAGAGGT CAAGTTATGA AAGACATGTT TTATTTTTAC
1451 AAACACATTC TCATGGACTA CAAGGGAGTG TTAGGCCCCA TAGGAGGTAT
1501 ATTGAACCAA ATCGGAATGG CTTCAAGAGC AATCCTCGAC GCTAAGTACT
1551 TCATCAAAGA GTATCACTAC ATTAGAGATA CATCGGCGAA AACGTTACAC
1601 TTAGATCGAG GGAAGAATT AGGAATATTC TGCACGATCG CGTGGAAATG
1651 TCATCATCAT AACAACGAGA TAAAAGTTCC TCCACAAGAA TGCATTGTAG
1701 TGAAGAAAGA TGCAACATTG AGTGAAGTGT ACGGAGAGGC AGAAAGAGTG
1751 TTTAGAGATA TCTATTGGGA ACTAAGAGAC GTCGTGGTGG AGTCAGTGGT
1801 GGGTGGTCAA ATAGAGATCA CAAGGGTCGA TGAAATGGCC TTGAATGGGA
1851 ATAAGGGATT GGTGTTAGAA GGGAACGTAG GAATGATGAT GAACATTGAA
1901 GTGACGAAAT GTTATGAAGA TGATGATAAA AAGAAGGATA AGAGAATAGA
1951 GTGTGAGTGT GGAGCAACGG AAGAAGATGG AGAGAGAATG GTGTGTTGTG
2001 ATATTTGTGA AGTATGGCAA CACACAAGGT GTGTTGGTGT CAACACAAT
2051 GAGGAAGTGC CTCGCATTTT TCTTTGTCAA AGTTGTGATC AACATCTTAT
2101 TCCTCTCTCT TTTTTACCCT AAATCATTCA TGTTATATTA AACCTTCTTT
2151 TTTTCTATCA TTACGTCATA ATACATAAAT TATACTCCTC GATC

8.8 AMS (AT2G16910) FULL LENGTH cDNA

1 ATGGAGAGTA ATATGCAAAA CTTGTTGGAG AAATTGAGGC CTCTTGTGGG
51 TGCAAGGGCT TGGGATTATT GTGTTCTTTG GAGACTAAAC GAAGACCAAA
101 GGTTTGTGAA GTGGATGGGA TGTTGTTGTG GTGGGACAGA ACTCATAGCG
151 GAGAATGGAA CAGAAGAGTT TAGCTACGGA GGTTGTAGGG ATGTTATGTT
201 TCATCATCCT CGGACCAAAT CTTGTGAATT TCTTTCCCAT CTTCCAGCTT

253

251 CCATACCTCT TGATTCCGGG ATATATGCGG AGACTCTTCT GACTAACCAG
301 ACTGGTTGGT TGAGTGAGAG CTCAGAACCA AGTTTTATGC AGGAAACAAT
351 CTGTACTAGG GTTTTGATTC CTATACCGGG AGGACTAGTG GAGCTCTTTG
401 CGACCAGACA TGTCGCTGAA GATCAGAACG TGGTGGATTT TGTAATGGGA
451 CATTGCAACA TGTTGATGGA CGATTCTGTA ACGATAAACA TGATGGTAGC
501 AGACGAGGTC GAATCGAAGC CGTACGGGAT GTTATCCGGT GACATCCAGC
551 AAAAGGGTTC CAAAGAAGAA GATATGATGA ATCTCCCTTC GTCTTACGAT
601 ATCTCTGCTG ATCAGATCCG ACTCAATTTC TTGCCTCAGA TGAGTGATTA
651 CGAGACACAA CACTTGAAGA TGAAGAGTGA TTATCATCAT CAAGCCTTGG
701 GGTATCTCCC AGAGAATGGT AACAAGGAAA TGATGGGTAT GAACCCATTT
751 AACACGGTGG AAGAAGATGG GATTCCAGTG ATTGGAGAGC CTAGCTTGCT
801 TGTGAATGAA CAGCAAGTTG TTAACGATAA GGATATGAAT GAGAACGGTA
851 GGGTGGATTC AGGGTCGGAT TGCAGCGACC AGATTGATGA TGAAGATGAT
901 CCAAAGTACA AGAAGAAGTC AGGAAAAGGA TCTCAAGCCA AGAACCTGAT
951 GGCTGAGAGA CGGAGGAGGA AGAAGCTAAA CGATAGGCTT TATGCTCTCC
1001 GGTCGCTTGT TCCCAGGATA ACCAAGTTGG ACAGGGCGTC GATTCTTGGC
1051 GATGCAATCA ACTACGTTAA GGAGTTGCAG AATGAGGCAA AGGAGCTTCA
1101 AGACGAGCTT GAAGAGAACT CAGAGACTGA GGATGGATCT AATAGGCCAC
1151 AAGGAGGGAT GAGCCTGAAC GGGACAGTGG TCACTGGGTT TCACCCGGGG
1201 CTTTCATGTA ACTCCAACGT TCCTAGCGTG AAGCAAGATG TTGATCTTGA
1251 GAATTCCAAT GATAAAGGAC AAGAAATGGA GCCACAGGTG GACGTGGCTC
1301 AGTTAGATGG CAGAGAGTTT TTCGTAAAGG TGATTTGCGA ATACAAACCA
1351 GGAGGCTTCA CAAGGCTAAT GGAGGCACTG GATTCTCTTG GACTAGAAGT
1401 CACAAATGCT AACACGACTC GCTACCTCAG CCTGGTCTCC AATGTCTTCA
1451 AAGTCGAGAA AAATGATAAC GAGATGGTCC AAGCTGAACA CGTAAGGAAC
1501 TCGTTGCTGG AAATAACCCG GAACACATCT AGAGGATGGC AAGATGATCA
1551 GATGGCTACA GGCTCTATGC AAAACGAAAA GAACGAAGTT GATTATCAAC
1601 ACTATGATGA TCACCAGCAT CACAATGGTC ATCATCACCC ATTTGATCAC
1651 CAGATGAATC AGAGCGCTCA TCATCATCAC CACCACCAAC ACATCAACCA
1701 TTACCACAAC CAATAAGGAC CAATTCCCCA AAGCTATAAA GATGTCTGCT
1751 TTACAAGCAA AAGAAACATT CTGCATCAAC AAACTTTGAA GATCTTTTCA
1801 TTTCAAGATA TATAAGGATA AAAGGAATGT CAGTTCGAGA ACAGTTCCTG
1851 TATCCAATAC CAGGAATAAA CCAAAGCGAA ACAACTTTAC TTA

8.9 ABC TRANSPORTER / ABCG26/ ATABCG26/MSR02/ATP-BINDING CASSETTE G26 (AT3G13220) FULL LENGTH cDNA

1 ATGGAGATCA GAAGATCAAC AGAAGAAGTT GAAGAGAATC ATGTAATGCA
51 GATTACAGGA AGCAATGGCA TAGTTCACAA CATGGAGTTC ATGCCACAAG
101 CTTACCTCAG AAACCAATAC TCATCCGAGA TTGACATTGA TGAAGAGTTC
151 GTTTCCACTT ATCCTCTCGA AGATGCCCCT CTTCCCATCT TTCTCAAGTT
201 TGAAGATGTG GAGTATAAGG TGAGAAACAG TCATGCTAGC TCAGCAAACC
251 TGGTGAAAAC AATGGTGTCA AAGGTGGTAA CACATACAAA TCCGGATCCA
301 GATGGATACA AGCACATTTT GAAAGGCATA ACAGGAAGCA CAGGTCCTGG
351 CGAAATCCTT GCGCTAATGG GGCCTTCGGG TAGTGGCAAA ACAACTCTTC
401 TGAAGATAAT GGGAGGAAGA TTGACAGATA ATGTCAAGGG GAAGCTTACC
451 TATAATGACA TTCCTTACAG TCCATCTGTT AAGAGGAGAA TTGGCTTTGT
501 GACGCAAGAC GATGTTCTGT TGCCACAACT TACTGTGGAA GAGACCTTGG
551 CATTTGCTGC GTTTTTAAGA CTTCCAAGTA GCATGAGCAA GGAACAAAAA
601 TACGCCAAAA TAGAGATGAT CATCAAAGAG CTCGGCCTCG AAAGATGCCG
651 TCGGACAAGA GTTGGAGGAG GTTTTGTCAA GGGAATATCA GGTGGAGAAA
701 GAAAAAGAGC AAGTATAGCT TATGAGATTC TCGTAGATCC TTCTCTGTTG
751 CTTCTAGACG AGCCAACTTC TGGACTTGAC TCCACTTCTG CCACCAAACT
801 TCTTCATATT CTTCAAGGAG TAGCTAAGGC AGGCAGGACC GTTATAACGA

254

851 CAATTCACCA ACCATCAAGC AGAATGTTTC ACATGTTTGA CAAGCTCTTA
901 CTGATATCAG AAGGACACCC TGCCTTTTAC GGCAAAGCCA GAGAATCTAT
951 GGAGTATTTC TCATCTTTGA GAATCTTACC TGAGATTGCA ATGAACCCAG
1001 CAGAGTTCTT GCTAGACTTG GCAACCGGAC AAGTTTCAGA CATTAGTCTA
1051 CCCGACGAAC TATTGGCTGC CAAAACTGCA CAGCCTGATT CTGAGGAAGT
1101 TTTACTCAAA TACTTGAAAC AAAGGTACAA GACAGACTTG GAGCCAAAAG
1151 AGAAGGAAGA AAATCATAGA AATAGAAAGG CCCCAGAGCA TCTTCAAATA
1201 GCAATCCAAG TGAAGAAGGA CTGGACTCTT TCATGGTGGG ATCAGTTCTT
1251 GATTCTTTCA AGAAGAACAT TCAGAGAGCG GCGTAGAGAT TACTTTGATA
1301 AGCTAAGGCT AGTTCAATCA CTTGGTGTGG CCGTCGTGTT GGGTCTTCTC
1351 TGGTGGAAGT CAAAGACAGA TACAGAGGCT CACCTTAGAG ACCAAGTTGG
1401 TTTGATGTTC TATATCTGCA TCTTCTGGAC ATCGTCATCA CTCTTTGGAG
1451 CAGTATACGT GTTCCCCTTC GAGAAGATCT ACCTGGTGAA AGAAAGAAAA
1501 GCAGAAATGT ACAGACTTAG TGTGTATTAC GTATGCAGTA CACTATGTGA
1551 TATGGTGGCA CACGTTTTGT ACCCGACATT CTTCATGATC ATTGTCTACT
1601 TCATGGCTGA GTTCAATAGG AACATCCCCT GCTTCCTTTT CACCGTGTTG
1651 ACAATACTAC TGATAGCCAT CACTAGCCAA GGTGCAGGAG AGTTCTTAGG
1701 AGCTTCAGTA TTGAGCATTA AGAGAGCTGG TATGATCGCC TCTTTGGTAC
1751 TAATGCTGTT TCTACTAACC GGAGGTTACT ATGTCCAGCA CATACCAAAA
1801 TTCATGCAGT GGTTGAAGTA CTTGTCATTC ATGCACTACG CTTCAGGCT
1851 ACTTTTGAAA GTTCAATACT CAGCTGACCA ACTTTTCGAA TGTGGGAGCA
1901 AAGGCGGCTG CAGGACACTG CAAAGCTCAT CGTCCTTTGA CACGATAAAC
1951 TTGAATGGCG GCTTACAAGA ACTGTGGGTT CTGCTTGCCA TGGCATTCGG
2001 CTACCGTCTC TGTGCATACT TTTGCCTTCG CAAGAAAATA AGCATTTGCC
2051 ATCTTTGA

8.10   POB1 (AT3G61600) FULL LENGTH cDNA

1 ACTTTGGTCT CGTTGAGCCT TTCGTCTATC TCTCACCTGA AAGAAAGACG
51 TTCAGCTTTC TCTTTCCAAA ATCAAAAGCT CTCCGATCAG TTTCTCCGGC
101 TTAGGAAATT AATTGAAGGA AGGTGATGAG AGGTACTACT GAGAATACGG
151 ATCTGTTCGA TCCCAAAACC CAAATGGATC CTGACTTCAC TCGTCATGGT
201 TCTTCCTCCG ACGGCGATTT TGGATTCGCT TTTAACGACA GTAACTTCTC
251 TGATCGTTTG CTTCGGATCG AGATCATGGG TGGACCTTCG GATTCTCGGT
301 CTGAAGTTGA AGGGTGTACG AGTATCGCTG ATTGGGCTCG TCATCGCAAG
351 AGAAGAAGAG AAGATATCAA GAAGGAATCT GGTGTCACGA TTTCAGACAT
401 TGTGGCATGT CCTGAGGAGC AGATTTTAAC TGATGAACAA CCTGACATGG
451 ATGGATGTCC TGGTGGTGAG AATCCTGATG ATGAAGGAGG AGAGGCAATG
501 GTTGAAGAAG CTTTATCAGG TGATGAAGAG AAACGTCTA GTGAGCCAAA
551 CTGGGGAATG GATTGTTCTA CAGTTGTTAG GGTTAAAGAA CTTCATATTA
601 GTTCTCCTAT TTTAGCTGCC AAAAGCCCTT TCTTTTACAA GTTGTTCTCC
651 AATGGAATGA GGGAATCTGA GCAAAGGCAT GTCACCCTTA GGATTAATGC
701 ATCAGAGGAA GCTGCTTTGA TGGAGCTTTT AAACTTTATG TATAGCAATG
751 CGGTATCTGT CACCACAGCA CCTGCCTTAT TAGATGTGTT GATGGCTGCT
801 GATAAGTTTG AAGTTGCCTC TTGTATGAGG TACTGCAGTA GACTTCTCCG
851 TAATATGCCT ATGACTCCAG AGTCTGCCCT GCTCTATCTC GAGCTTCCCT
901 CTAGTGTTCT AATGGCCAAA GCTGTTCAGC CTTTAACTGA TGCTGCAAAA
951 CAGTTCCTTG CAGCCCGCTA CAAGGATATT ACCAAGTTTC ATGAGGAGGT
1001 AATGTCTCTA CCTTTGGCTG GAATTGAGGC AATTCTATCA AGCGATGAAC
1051 TCCAAATTGC ATCAGAGGAT GCAGTTTATG ATTTTATCTT GAAGTGGGCA
1101 AGGGCGCAAT ACCCTTGTTT GGAAGAGCGA AGAGAGATTC TCGGGTCACG
1151 CCTTGCACTC TCCATCCGCT TCCCATTCAT GACATGCCGA AAGCTGAAGA
1201 AAGTGCTGAC TTGCAGTGAC TTTGAGCATG AAATAGCATC AAAGCTTGTT
1251 CTAGAAGCTC TTTTCTTCAA AGCAGAAGCC CCACACAGAC AACGTAGCCT
1301 AGCCTCCGAA GAATCTGCAT CCCTGAACCG CCGCCTGATA GAGAGGGCTT
1351 ACAAATACAG ACCCGTCAAA GTGGTCGAGT TTGAGCTTCC TAGACCGCAG

1401 TGTGTAGTCT ACCTAGACTT GAAAAGAGAA GAATGTGGGG GACTGTTCCC
1451 GTCGGGTAGA GTGTATTCGC AGGCCTTTCA CTTGGGAGGT CAAGGGTTTT
1501 TCCTGTCAGC TCACTGCAAC ATGGACCAAC AGAGCTCGTT CCACTGTTTC
1551 GGGCTGTTCC TAGGGATGCA GGAGAAAGGG TCGGTGAGTT TCGGAGTGGA
1601 CTATGAATTC TCGGCAAGGT CAAAGCCCGC AGAGGATTTC ATAAGCAAAT
1651 ACAAAGGGAA CTACACATTC ACAGGAGGGA AAGCAGTAGG TTACAGAAAC
1701 CTGTTTGGGG TCCCATGGAC GTCTTTTATA GCGGAAGATA GTCAATACTT
1751 CATCAATGGC ATTCTCCATC TCAGAGCAGA GCTTACCATC AAAAGGTCTA
1801 CAGATCCTTA GTGACCGTTT ACTTCTCAAT TGTCACTGTC TCTCTTCCT
1851 TTACCTGCTG CTAAATCTCT CGCTCTCTCT AACTTGCCTT TTTTTTCTTT
1901 TCTTTTATAA GGCTTGAAGA AGACTAATAA AGCTTATAAA AGTAAGGTTA
1951 AAAACGTTTC GACGACAACG TTATAAGAAG AATGTAAATG CGTTCTTTTT
2001 TTCTATCTAC CCTTACTCTT TTGGCTTTGG ATTATAAGTT TAACTCTTCA
2051 ATGAAATTCT TAGCGTAGAG ATTCTTGTAT TATGAACGAA GTTGTGAACA
2101 GAAAAGAACG AACGTAACTC TCTCGTCTCA TTTAACTAAT TTATCTAGTT
2151 GGATTGAAAT AGACATGACT TTTCGATT

8.11   POB2 (AT2G46260) FULL LENGTH cDNA

1 AGAGAAAGAC TTTTTTAGTG TTTCTTTGTC ATTTTCGTTC CTCAAAAAAA
51 CCTTTTCGAA TTTGACCCTT ATTGTTATCC TGATTCTTCT GCTCGGTTCT
101 CGCCGGAGAG AAGACACGGA GAAGAGGGTT TTCGAAAAGG AGGATGAGAG
151 GTTCCAATAA CACCGATCTA TTTGACCCTA AGACGGAGAT GGATTCCAAT
201 TTCTCTCGCC ATGGTTCTTC CTCCGAAGGT GATTTTGGTT TCGCTTTCAA
251 TGACAGTAAC TTCTCCGATC GTTTGCTCCG GATCGAGATC TTGGGTGGGC
301 CTTCGGATTC TAGGTCTGAT GCTGAAGGAT GTACGAGTAT TGCCGATTGG
351 GCTCGTCATC GCAAGAGGAG AAGAGAGGAC AACAAGAAGG ACAATGGTGT
401 TGCGATTTCA GACATTGTGG CATGTGCTGA AGAACAGATT TTAACCGATA
451 ACAACCAACC TGATATGGAT GATGCTCCTG GTGGTGATAA TCTTGACGAT
501 GAAGGAGAGG CAATGGTTGA AGAGGCTTTA TCAGGTGATG ATGATGCATC
551 TAGTGAGCCA AACTGGGGTA TTGATTGTTC TACTGTTGTT AGAGTTAAAG
601 AACTGCATAT TAGTTCTCCT ATCTTGGCCG CAAAAAGCCC CTTTTTCTAC
651 AAGCTGTTTT CTAATGGAAT GAGGGAATCA GAACAAAGAC ATGTAACCCT
701 TAGAATTAGT GCACAAGAGG AAGGTGCTTT GATGGAGCTT TTAAACTTTA
751 TGTATAGCAA CTCTCTAAGT GTCACAACAG CACCCGCTTT ATTAGATGTT
801 CTTATGGCTG CTGACAAGTT TGAGGTTGCT TCCTGTATGA GGTATTGTAG
851 TAGACTACTG CGAAATATGC CCATGACCCC TGATTCTGCT TTGCTCTATC
901 TTGAGCTGCC CTCTAGTGTT TTAATGGCTG AAGCAGTGCA ACCTCTAACC
951 GATGCAGCAA AGCAGTTCCT TGCCTCGCGC TACAAGGATA TTACCAAGTT
1001 TCATGATGAG GTTATGGCCT TACCATTGGC TGGAATAGAG GCGATACTAT
1051 CGAGCGATGA TCTCCAAATT GCTTCTGAGG ATGCTGTTTA TGATTTTGTC
1101 TTGAAATGGG CAAGGGGACA GTATAGTTCA TTGGAAGATC GTAGAGAGAT
1151 TCTGGGTTCA CGCCTTGCAC TCTACATCCG CTTCCCATAC ATGACGTGTC
1201 GAAAGCTGAA AAAGGTACTA ACGTGCAGTG ATTTTGAGCA TGAAGTAGCA
1251 TCAAAGCAAG TATTAGAAGC GCTCTTCTTC AAAGCAGAAG CCCCACACAG
1301 GCAACGCATC CTCGCAGCTG AAGGATCGGA CTCCATGAAC CGCCGTTTCA
1351 TAGAGAGGGC TTACAAATAC AGACCCGTAA AAGTGGTGGA GTTTGAGCTT
1401 CCTCGCCCGC AATGTGTAGT GTACCTGGAC TTGAAGCGAG AGGAATGCGC
1451 AGGACTATTC CCTTCAGGAA GAGTCTATTC TCAGGCTTTC CACTTAGGAG
1501 GTCAAGGTTT CTTCCTCTCA GCACACTGCA ACATGGACCA GCAAAGCTCG
1551 TTCCACTGCT TTGGGCTATT CCTCGGAATG CAAGAGAAAG GGGCTGTGAG
1601 TTTTGGTGTG GACTACGAAT TTGCGGCAAG AGATAAAAGT ACAAAGGAAG
1651 AATACGTAAG CAAATACAAA GGAAACTACA CATTCACAGG GGGGAAAGCG
1701 GTTGGTTACA GAAACCTTTT CGGGATTCCA TGGACTTCGT TCATAGCCGA
1751 GGACAGTCAA CATTTCATCA ATGGCATTCT CCATCTCAGA GCCGAGCTCA

1801 CCATTAAACG TTCCTCAGAC CTGCACTGAC TGTTTTCTTC TCTCTCTCTC
1851 TCTATGTCGT CACATTTAAG TTGTATAATC TCGTTATGTA GCAGCTCTAA
1901 CCACGAGTGT CCTTTGTACA ATAATTGCCT AAGGAGCAAA AGTATGGCTT
1951 AAAGAGTACC TCATTCTGTA AATTCTTTGA TTGGATTTTT TTACTTTTC

8.12   LB (LURIA BROTH)

For 1 litre of Medium	
Yeast Extract	0.5%
Tryptone	1%
NaCl	10 g
Bacto-agar	1.5%
ddH$_2$O	Complete to 1lt

Before being sterilized by autoclaving at 120˚C for 16 minutes pH was adjusted to 7.0.

8.13   PLASMID DNA EXTRACTION SOLUTION I, II AND III.

Solution I (Per litre)		Solution II (Per litre)		Solution III (Per litre)	
50% Glucose	18 ml	10% SDS	100 ml	5M K Acetate	600 ml
1M Tris-HCl (pH 8.0)	25 ml	1N NaOH	200 ml	Glacial Acetic acid	115 ml
0.5 M EDTA (pH 8.0)	20 ml	ddH$_2$O	Complete to 1 litre	ddH$_2$O	Complete to 1 litre
ddH$_2$O	Complete to 1 litre				

8.14   TAE BUFFER.

For 1 litre of 50x Buffer	
Tris Base	242 g
Glacial Acetic Acid	57.1 ml
0.5 M EDTA (pH 8)	100 ml
ddH$_2$O	Complete to 1lt

8.15   TB BUFFER

For any volume of buffer	
PIPES	10 mM
CaCl$_2$	15 mM
KCl	250 mM

Adjust the pH to 6.7 and then add MnCl$_2$ to 55 mM, complete to final volume with ddH$_2$O, sterilize by filtration and store at 4 ˚

8.16   TBE BUFFER

For 1 litre of 10x Buffer	
Tris base	108 g
Boric acid	55 g
Na$_2$EDTA 0.5M (pH 8.0)	40 ml
ddH$_2$O	Complete to 1 litre