Isolation and characterisation of wheat genes with early meiotic expression

by

Hayley Rebecca Jolly
B. Science (Hons.), The University of Adelaide

A thesis submitted for the degree of
Doctor of Philosophy

at

The University of Adelaide
Faculty of Sciences
School of Agriculture, Food and Wine
Waite Campus

June 2010

Table of Contents

List of Figures

List of Tables

Abstract

For all sexually reproducing organisms meiosis is an essential process. Cells undergoing meiosis go through one round of DNA replication followed by two successive rounds of chromosome segregation and cellular division, meiosis I and meiosis II. During chromosome alignment and subsequent segregation, recombination or the exchange of genetic information between homologous chromosomes takes place, which effectively maintains genetic diversity through new allelic combinations. The majority of meiotic research to date has been limited to diploids, such as yeast, mice and Arabidopsis. However, as more than 70% of all flowering plants are polyploid, exclusively using diploids could result in over simplifying the process in more complex species. A recent transcriptomics study in bread wheat conducted by Crismani and colleagues (2006) identified 1,350 transcripts which were temporally regulated during the early stages of meiosis. A number of the meiotically-regulated transcripts have annotated functions for chromatin condensation, synaptonemal complex formation, recombination and fertility; however 1,094 transcripts were deemed to have either no annotations or purely predicted annotations.

The meiotic gene *DMC1*, previously described in several diploid species, was one of the 1,350 transcripts having meiotic regulation in wheat. Isolation of the bread wheat *DMC1* homologue, and the genes encoding *Ta*HOP2 and *Ta*MND1 which are known to interact closely with DMC1, was successful. *TaDMC1* and *TaMND1* were located on chromosome group 5, while *TaHOP2* was located on chromosome group 4. All three genes have significant similarity to the rice (98%, 97%, 95%, respectively) and Arabidopsis (91%, 85%, 87%, respectively) homologues. *TaHOP2* and *TaMND1* transcript expression was found to be highly correlated (r = 0.98) across the germline and somatic tissues examined. Using an anti-*At*MND1 antibody, the *Ta*MND1 protein localised to meiotic chromosomes during pre-meiotic (PM) interphase through to pachytene, while localisation attempts for the *Ta*DMC1 protein with an anti-*At*DMC1 antibody proved unsuccessful. The functions of the *Ta*DMC1

and *Ta*HOP2 proteins were found to be conserved, with these proteins having the ability to bind preferentially to single-stranded DNA.

Another gene identified from the transcriptomics study, and within the 1,350 transcripts, was *TaASY1* (bread wheat *ASY*napsis*1*). While *TaASY1* is known to be involved in chromosome pairing and synapsis, the mechanics of how this gene is regulated still remains somewhat unknown. To elucidate potential protein partners of the *Ta*ASY1 homologue, a yeast two-hybrid approach was conducted. Three proteins were found to interact with the full length *Ta*ASY1 protein and the HORMA domain of the ASY1 protein, along with one protein which interacted only with the isolated HORMA domain. Characterisation of two of the proteins which interact with *Ta*ASY1 revealed a Replication Protein A (RPA14) homologue and a novel plant protein.

A further objective of the research conducted within this study was to isolate and characterise novel meiotic candidates to further the understanding of the process in bread wheat. From the previously identified 1,094 novel transcripts, those with either no or putative annotations underwent selective filtering. This resulted in 40 candidates being targeted for further analysis, after which eight candidates were selected for in-depth analysis. Arabidopsis homologue mutants were obtained for these candidates to determine whether a meiotic role could be assigned. Mutant plants from one candidate, *AtMN29,* showed altered phenotypic features that suggest this gene has a role during meiosis in Arabidopsis. Abnormalities included decreased silique length, decreased seeds per silique, decreased pollen viability and altered expression of the meiotic genes *ASY1* and *DMC1*.

Declaration

I declare that the work presented in this thesis contains no material which has been accepted for the award of any other degree or diploma in any University or other tertiary institution. To the best of my knowledge and belief, this thesis does not contain any material previously written or published by another person, except where due reference is made in the text.

I give consent to this copy of my thesis when deposited in the University Library, being available for loan and photocopying, subject to the provisions of the copyright Act 1968.

The author acknowledges that copyright of the published works contained within this thesis resides with the copyright holder(s) of those works.

Hayley R. Jolly

June 2010

Acknowledgments

Firstly I would like to thank my principal supervisor Dr Jason Able. Throughout my PhD you have encouraged and challenged me to better myself in both my research and personal life. Thank you for pushing me, and yes it feels good to be here finally. Thank you so much for all that you have done for me and I look forward to keeping in contact with you and possibly collaborating together at later stages in our research careers. I would also like to thank my co-supervisor Dr Andrew Milligan, for continuing in being my supervisor despite your move and busy schedule. Thank you for always providing a new perspective and new ideas for my project.

I would also like to thank all those who have provided specialised help to me during my PhD including: Dr Neil Shirley (ACPFG) for help with Q-PCR, Ms Margaret Pallota (ACPFG) for nulli-tetra membranes for Southern analysis and Ms Lyn Waterhouse (Adelaide Microscopy) for help and advice with fluorescent confocal microscopy.

Thank you to all past and current "Able" lab members. To Dr Scott Boden and Dr Wayne Crismani, firstly for providing me with a great platform of research to build off for my project and secondly for providing me with advice even after leaving for bigger and better things in your post doc careers. I would like to thank Kelvin Khoo, for being my "lab partner" during my PhD, you have been constant support throughout and I cannot thank you enough for being there for me. Thank you to Dr William Bovill and Caroline Abrahamse, for help and time during my time in the group and I look forward to keeping in contact in the future.

Thankyou very much to the past and current members of the Molecular Plant Breeding (MPB) CRC education team, Dr Amanda Able, Dr Heather Bray, Michael Mclean, Melanie Carew and Belinda Griffith. Thank you for all the great professional development and educational support which has been provided through the MPB CRC Student Retreats and Annual Research Meetings.

Thankyou very much to the University of Adelaide, the School of Agriculture, Food and Wine, the Australian Government and MPB CRC for providing me with the space and funding for conducting my PhD within the University at the Waite Campus, Adelaide. Thank you very much to the Australian Society of Plant Scientists and MPB CRC for enabling me to travel to conferences during my candidature, allowing me to gain experience in communicating and expanding my research and networking with fellow molecular biologists.

I would also like to thank friends and family who have provided me with never ending support during my PhD and throughout my university life. Thank you to Kieren Arthur, fellow PhD student, for always being an ear to listen and all the coffee and unwines. To Ana, Bianca and all the other WAGs, for your excitement and understanding of my PhD project and the time it took away from group gatherings especially lately; thanks and looking forward to more drinks on the weekends. Thank you to Sarah Fox, having known you since high school, I wholly thank you for still being my bestie; even through weeks of silence, yet being able to make it as if it had never been weeks but only days. A special thanks to Loris Siddons, for all the coffees after work with the dogs and helping me to spend my scholarship every time we went shopping. Thank you very much to all my family members, for all your support over all my schooling years; yes I am finally finishing studying and am getting a real job. Thank you : John, Lachlan, Dad, Sue, Toni, Zach, Nan and Pa Jolly, Nan Wharton, Ally and Pete and the girls; Rob and Jackie and the boys, Steve, Regan, Angus and Tylah, Peter, Kate and Brett. I would also like to thank Pa Wharton, for all the garden tours in your backyard, who was to know that it would foster my future and love in plant science; I miss you and wish you were still here to see and enjoy this day with me. Finally I would especially like to thank my mum, Leann Mclean, I would not be where I am today without you. You have and continue to support, encourage me and be excited for me, even if you don‟t fully understand half the things I talk about.

Last, but not in the least, Mark. I love you so, so much. I could have never been able to imagine that I would have made it through without you. For making sure that I took time

out to relax and making me laugh on the worst of the PhD days. Thank you for all the time that I was not able to spend with you, all the holidays I was not able to take when you had holidays and when what little time I did have had to be shared around with others. Again I cannot thankyou enough for all your encouragement and understanding. I love you so much and cannot wait to spend the rest of my life with you. With PhD done, wedding planning should be a breeze!

Glossary of Abbreviations

Abbreviation	Full term
3'	three prime
5'	five prime
9mer	9 base pair nucleotide
α-dCTP	alpha-deoxycytidine triphosphate
ºC	degrees Celcius
Amp	*Ampicillin*
ASY1	*ASYnapsis 1*
At	*Arabidopsis thaliana*
AUS	Australia
BCIP/NBT	5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium
BLAST	Basic Local Alignment and Search Tool
bp	base pair
BSA	Bovine Serum Albumin
BW26	Bobwhite 26 cultivar of bread wheat
cDNA	complimentary deoxyribonucleic acid
DABCO	diazabicyclo-[2,2,2] octane
DAPI	4',6-diamidino-2-phenylindole
DMC1	*Disrupted Meiotic cDNA 1*
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetra-acetic acid
EMS	ethane methyl sulfonate

EST	expressed sequence tag
g	gram
g	relative centrifugal force
GAPdH	*GlycerAldehyde-3-Phosphate DeHydrogenase*
His	histidine
HOP1	*HOmologous Pairing 1*
HOP2	*HOmologous Pairing 2*
HORMA	Hop1p, Rev7p, MAd2
IgG	immunoglobulin G
kb	kilobase
kDa	kilo Dalton
LB	Luria Bertani
M	molar
mg	milli gram
mM	milli molar
MND1	*Meiotic Nuclear Division 1*
mRNA	messenger ribonucleic acid
MRE11	*Meiotic REcombination 11*
MSH4/5	*Mut S Homologue 4/5*
MS/MS	tandem mass spectrometry
NCBI	National Center of Biotechnology Information
ng	nano gram
nm	nano metre
NT	nullisomic-tetrasomic
ORF	open reading frame
Os	*Oryza sativa*
p	probability

Ph1	*Pairing homoeologous 1*
PAIR2	*homologous Pairing Aberration In Rice meiosis 2*
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Q-PCR	quantitative real-time PCR
RAD51	*RADiation sensitive 51*
RNA	ribonucleic acid
RNAi	RNA interference
rpm	revolutions per minute
RT-PCR	reverse transcriptase PCR
Sc	*Saccharomyces cerevisiae*
SC	synaptonemal complex
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS - polyacrylamide gel electrophoresis
SSC	standard saline citrate
SPO11	*SPOrulation-deficient 11*
Ta	*Triticum aestivum*
Taq	*Thermus aquaticus*
TBS	Tris-buffered-saline
T-DNA	transfer DNA
TE	Tris EDTA solution
Tris	tris(hydroxymethyl)aminomethane
U	units
µL	micro litre
µg	micro gram
µM	micro molar
UV	ultra-violet

V	volts
v/v	volume/volume
w/v	weight/volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyrano-side

Chapter 1

Chapter 1 – Literature Review

1.1 – Meiosis and meiosis in bread wheat

For sexually reproducing organisms meiosis is an essential process. Meiotic cells undergo one round of DNA replication followed by two successive rounds of chromosome segregation and cellular division, meiosis I and meiosis II. During chromosome segregation the exchange of genetic information between homologous chromosomes can take place, which effectively maintains genetic diversity through new allelic combinations (Pawlowski and Cande, 2005).

Through each of the two successive rounds of division the cell proceeds through five main stages: prophase, metaphase, anaphase, telophase and cytokinesis (Figure 1.1). Prophase I can be further divided into five sub-stages: leptotene, zygotene, pachytene, diplotene and diakinesis. It is during these sub-stages of prophase I that chromosomes first come together and form a physical connection (synapsis) with their homologous pair, via the proteinaceous synaptonemal complex (SC), and subsequently undergo homologous recombination and crossing-over (Hamant *et al.*, 2006).

Investigation of meiotic processes has been extensive in model diploid species, including yeast (Zickler and Kleckner, 1998) and Arabidopsis (Armstrong *et al.*, 2003). These studies (and others) have produced detailed time-courses of both the cytogenetic and biochemical pathways in diploid meiosis. However, diploid meiosis studies can result in oversimplification (Martínez-Pérez *et al.*, 1999), as it is estimated that over 70% of flowering plants are polyploids (Masterson, 1994; Bowers *et al.*, 2003) including a majority of the important agricultural crops (Udall and Wendel, 2006). A consequence of polyploidisation is that a complex process such as meiosis becomes even more complex (Moore, 2002).

Figure 1.1 - Chromosome behaviour and structure during meiosis in maize. (**a**) Leptotene. Chromosomes start to condense into a thread-like form. The nucleolus is intact in the centre of the nucleus. (**b**) Leptotene-zygotene. Chromosomes condense further and the nucleolus has moved to the periphery of the nucleus. (**c**) Late zygotene. Chromosomes are partially synapsed. (**d**) Pachytene. Chromosomes are fully synapsed. (**e**) Diplotene. (**f**) Diakinesis. Ten bivalents can be distinguished. **(g)** Metaphase I. (**h**) Anaphase I. (**i**) Telophase I. (**j**) Prophase II. (**k**) Metaphase II. (**l**) Anaphase II. Scale bar = 5 μm. Figure adapted from Hamant *et al.*, 2006.

Bread wheat (*Triticum aestivum* L.) is an allohexaploid resulting from the fusion of three different species: *Triticum urartu*, *Aegilops speltoides* and *Triticum tauschii*, with each contributing seven pairs of chromosomes to the genetic complement (Petersen *et al.*, 2006). A chromosome pair from a given chromosome group and the same genome are termed homologues; whereas similar chromosomes from a chromosome group but from different genomes are referred to as homoeologues (Figure 1.2) (Moore, 2002).

NOTE:
This figure is included on page 3 of the print copy of the thesis held in the University of Adelaide Library.

Figure 1.2 - Chromosome relationships in bread wheat. Homologous chromosomes have identical gene order and repetitive DNA sequences, whereas homoeologous chromosomes share a similar gene order, with their respective repetitive DNA sequences differing (Able *et al.*, 2007). Chromosomes from different chromosome groups are referred to as non-homologous. Only group one chromosomes are shown here (with sister chromatids not depicted), yet this description holds true for all the seven chromosome groups in bread wheat. Figure adapted from Moore, 2002.

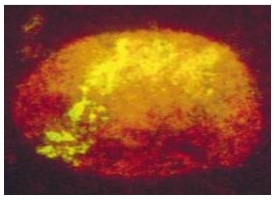
During meiotic prophase I in bread wheat, chromosome pairing is restricted to homologues; despite the significant conservation of gene content and order that occurs between homoeologues (Sears, 1952). Pairing of homoeologues results in multivalent associations and improper segregation of chromosomes at anaphase I; often producing unbalanced and non-viable gametes. Allopolyploid organisms avoid this by behaving like a diploid during meiosis, with strict homologous pairing. Yet interestingly, homoeologous chromosomes and even non-homologous chromosomes can be seen to have pre-meiotic associations, mainly at the centromeres (Aragón-Alcaide *et al.*, 1997). Therefore active processes which distinguish improper associations, those between homoeologous and non-homologous chromosomes, exist within polyploids.

1.2 – Three key processes that occur during prophase I

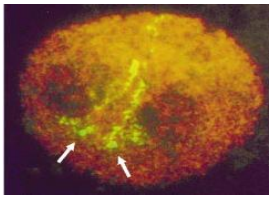
1.2.1 – Homologous chromosome pairing

Pairing is the process by which chromosomes are brought into close association for the search of sequence homology in the nucleotides of DNA (Harper *et al.*, 2004). Homologous chromosomes come into rough alignment during leptotene, before the onset of synapsis in zygotene (Figure 1.3 A) (Schwarzacher, 1997). This alignment may occur through a number of mechanisms but one model suggests that rough alignment is brought about by bringing key DNA sequences on homologous chromosomes together at the same „rendezvous" location (Wilson *et al.*, 2005). By having similar sequences (telomeres, centromeres and rRNA sequences) and key DNA sites (high transcription areas) along the chromosomes, rough alignment of the chromosomes can be readily achieved before synapsis (Figure 1.3 B) (Wilson *et al.*, 2005).

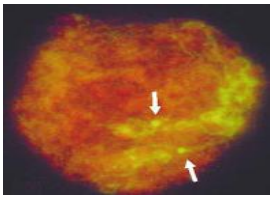
A



Early leptotene



Mid leptotene



Late leptotene

B

NOTE:
This figure is included on page 4 of the print copy of the thesis held in the University of Adelaide Library.

Figure 1.3 - Rough chromosome alignment during leptotene. (**A**) In the wheat-rye translocation line T5AS·5RL, chromosome domains on the translocated long arm of the rye chromosome 5R (*green-yellow*, FITC) can be seen to associate. This association is seen before the onset of synapsis, which occurs during zygotene. Arrows indicate the heterochromatin which occurs 80% of the way down the rye chromosome arm. Figure adapted from Schwarzacher, 1997. (**B**) Rough alignment of a chromosome could be achieved through the strategic placement of key DNA/high transcriptional sequences (depicted in grey circles), centromeric (depicted in green circle) and telomeric regions (small circles; with the red area representing the telomere bouquet) around the nucleus. Parallel lines represent a pair of homologues. Figure modified from Wilson *et al*., 2005.

1.2.2 Recombination of homologous chromosomes

Although the chromosomes are roughly associated during leptotene, for alignment at the metaphase I equator and correct segregation at anaphase I the chromosomes need to be closely connected. This close alignment (or juxtaposition) and then subsequent physical connection of homologous chromosomes occurs through the mechanisms of recombination and synapsis. Homologues in some organisms do have the ability to identify one another through mechanisms independent of recombination and synapsis, such as in fission yeast (Nabeshima *et al.*, 2001) and worm (*Caenorhabditis elegans*) (Dernburg *et al.*, 1998). However in the majority of organisms, including budding yeast, Arabidopsis and rye, the stable pairing of homologues during meiosis is dependent on recombination and/or synapsis (Grelon *et al.*, 2001; Mikhailova *et al.*, 2001; Peoples *et al.*, 2002).

Initiation of recombination is known to begin with programmed double-stranded breaks (DSBs) which in *S. cerevisiae* requires at least ten gene products: SPO11, RAD50, MRE11, XRS2, MEI2, MER2, REC102, REC104, REC114 and SKI8 (Keeny, 2001; Nag *et al.*, 2006). The biochemical role of many of these proteins in the formation of DSBs is not clearly understood but it is widely believed that DSBs are catalysed by the topoisomerase-related transesterase Spo11p (reviewed in Keeney, 2001). The leptotene DSBs are then processed by endonucleases to create single stranded overhangs (Wilson *et al*., 2005). These single-stranded DNA (ssDNA) molecules and associated proteins attract the well known recombination proteins RAD51 and DMC1, which are loaded onto the ssDNA (Shinohara and Shinohara, 2004). Both RAD51 and DMC1 promote strand invasion, with DMC1 required to facilitate inter-homologue recombination (Schwacha and Kleckner, 1997). This DMC1-mediated strand invasion into the homologous chromosome has been postulated by some as a mechanism for bringing homologous chromosomes even closer; „an intimate pairing" ready for the installation of the SC and complete synapsis (Figure 1.4), along with performing DNA repair (Chen *et al*., 2004).

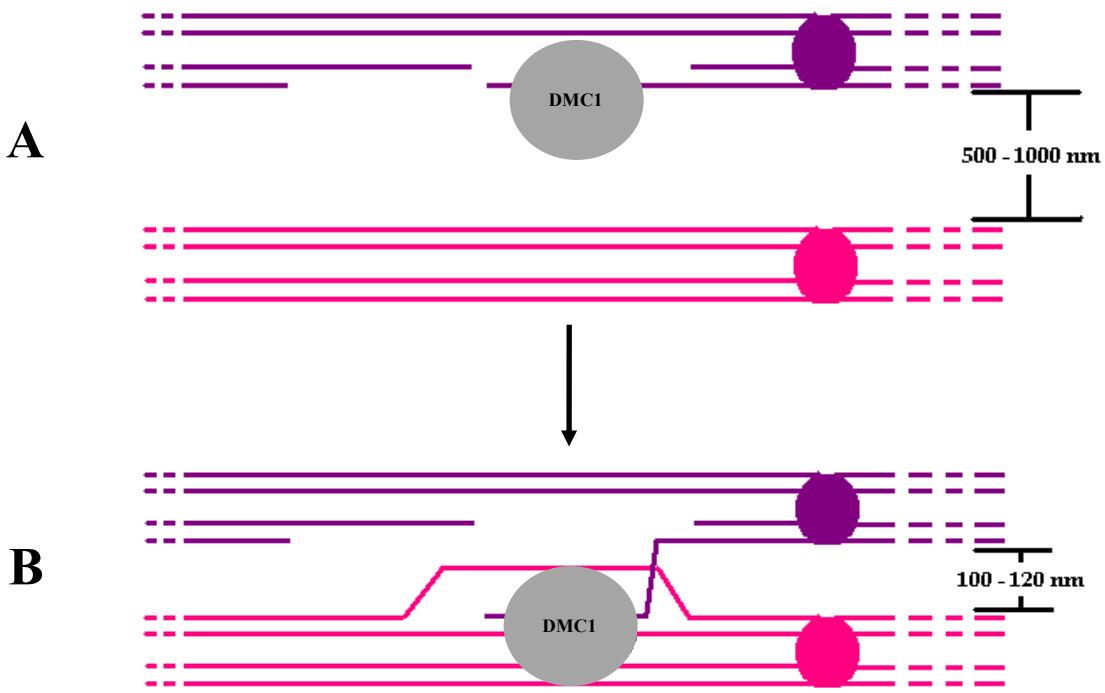


Figure 1.4 - DMC1-mediated strand invasion could act as a mechanism for chromosome juxtaposition. Homologous chromosomes in rough alignment undergo DSBs, end processing and loading of DMC1 (**A**). Strand invasion mediated by DMC1 could „pull" the processed chromosome during homologue searching and D-loop formation to form a close association with its homologous partner (**B**). Chromosome distances from Wilson *et al*., (2005).

1.2.3 Synapsis: forming an intimate association between homologues

Once homologous chromosomes are in juxtaposition during zygotene, full synapsis can occur. Synapsis is the formation of the tripartite synaptonemal complex (SC) along the entire length of the homologous chromosomes; consisting of axial elements (formed during leptotene), which are then later referred to as lateral elements, and of central transverse elements. Telomeric and centromeric regions are generally among the first chromosome regions to synapse in plants probably due to the close proximity of each, which occurs during the earlier pre-meiotic and rough alignment stages (Figure 1.5) (Schwarzacher, 1997; Pawlowski and Cande, 2005).

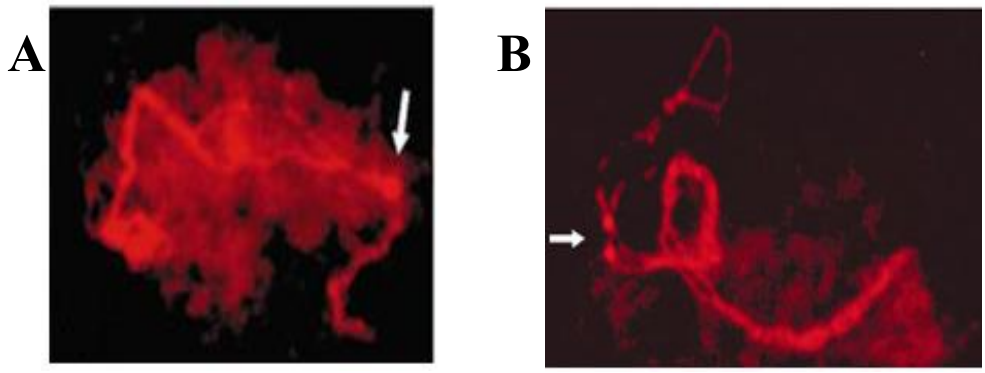


Figure 1.5 - Homologous chromosomes during zygotene. Often the first regions to be synapsed at zygotene are the telomeric and centromeric regions, with the middle of the chromosome arms not yet being intimately associated (**A**). During late zygotene to early pachytene, a chromosome pair is almost fully synapsed, with only small interstitial non-paired sites (**B**). Arrows designate the heterochromatin knobs on the rye 5RL arm. Figure adapted from Schwarzacher (1997).

The initiation of synapsis is believed to be linked with recombination nodules, and thus recombination (Pawlowski and Cande, 2005). Early recombination nodules contain DMC1 and RAD51, of which RAD51 has been shown to interact with an SC component, ZIP3, in yeast (Agarwal and Roeder, 2000). ZIP3 has also been shown to interact with recombination proteins MSH4 and MSH5, which localise to late recombination nodules. Although recombination and synapsis are separate processes, they are both occurring at similar times during prophase I, thus it is highly possible that synapsis and the SC form to stabilise chromosome associations during the process of recombination (Zickler and Kleckner, 1999).

Upon completion of recombination during pachytene, the SC disassembles (Zickler and Kleckner, 1999), and chiasmata are then the only physical links holding homologous chromosome pairs together through the later stages of meiosis, metaphase I and chromosome segregation at anaphase I (Moore and Orr-Weaver, 1998). The importance of synapsis on the later stages of pairing has been shown by many; such as in asynaptic yeast mutants where

homologous pairing, while still occurring, was observed to be greatly reduced (Loidl *et al.*, 1994). Pairing and bivalent formation at metaphase I was also shown to be highly disrupted in Arabidopsis when the SC transverse element ZYP1 (a homologue of yeast ZIP1) was mutated (Higgins *et al.*, 2005).

1.3 – Tools for studying meiosis

As within many fields of life science, there are a spectrum of molecular techniques ranging from the analyses of DNA, RNA and protein that can be utilised to answer biological questions (Figure 1.6). Invariably, researchers often use many of these techniques in concert or combination to answer the functions of particular genes; and in the past decade there has been heightened interest in furthering our understanding of meiosis in a broad range of organisms (McKim *et al.*, 1998; Azumi *et al.*, 2002; Libby *et al.*, 2002; MacQueen *et al.*, 2002; Chen *et al.*, 2005). Of particular interest to the plant science community, have been recent studies focussing on the use of mutants to discover whether the absence of a particular gene disrupts other protein(s), thus indicating a likely link between the two genes/proteins. As such many genes are usually taken through an „analytical pipeline" that includes isolating a given gene, determining the transcript expression and protein expression/location analysis in a „nomal" wild-type setting, which is then often followed by analysis of a mutant that lacks a functional copy of the gene under study.

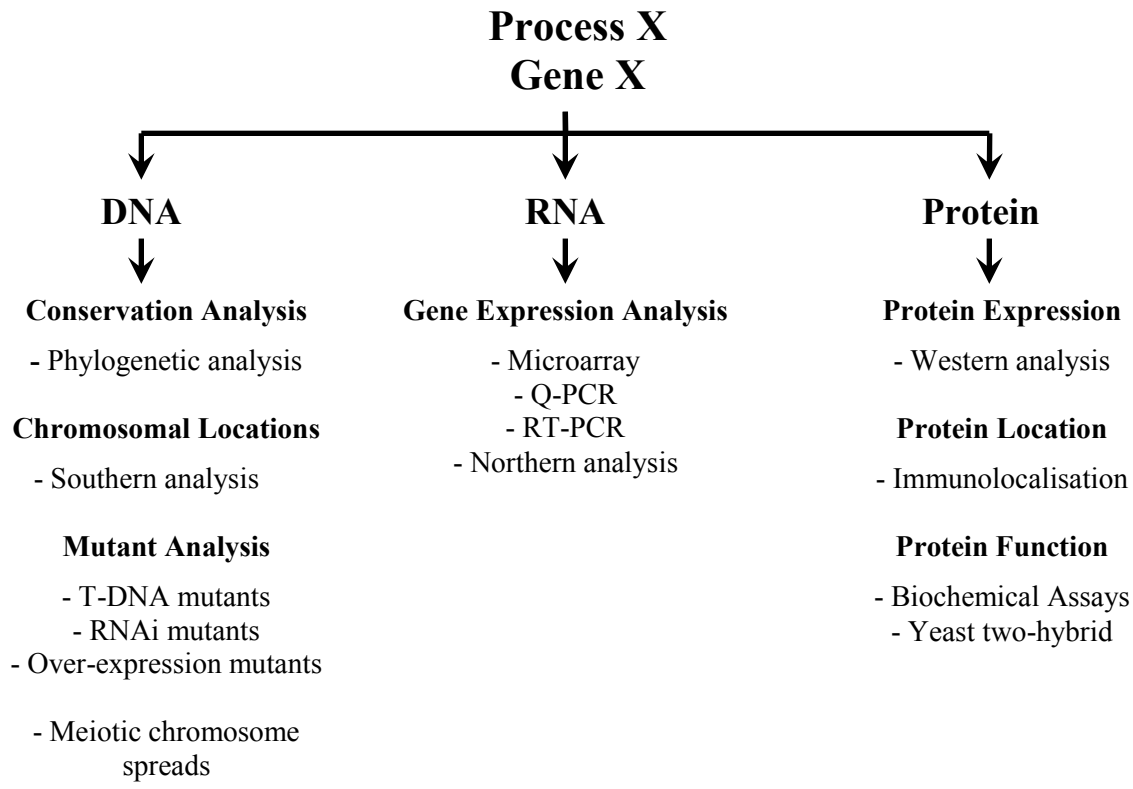


Figure 1.6 - Tools for studying meiosis. The biological questions of meiosis and meiotic gene function can be answered using a wide-range of DNA, RNA and protein techniques.

1.3.1 – Using microarrays as a tool to study meiosis

As highlighted in Figure 1.6, there are a number of molecular techniques which can be used to study gene expression; including northern analysis, quantitative real time PCR (Q-PCR) and reverse transcriptase PCR (RT-PCR). These techniques allow for detecting and quantifying any differences in gene expression. However a limitation of these particular techniques is that they can only measure differences of one gene at a time. While this is useful in some instances, such as the measurement of gene expression of a given gene in a knock-out or over-expression mutant; it does not allow for high throughput analysis of gene expression. In recent years, DNA microarray or DNA chips have evolved into a powerful tool used primarily for high throughput gene expression analysis, as well as having applications in other areas such as genotyping and protein interactions (Schena, 2003). Microarrays enable the expression levels of hundreds, or even thousands of genes to be measured simultaneously;

with the enormous amount of information generated requiring the use of statistics and powerful computers (Hanai *et al.*, 2006).

Typically, microarray experiments can include analysis of expression profiles of different treatments; expression profiles of developmental stages in certain tissues or differences between tissue types, for example diseased tissues compared to normal tissues or leaf expression compared to flower expression profiles. Indeed, in the past five years, numerous studies have been conducted using microarrays to understand the processes of meiosis and germ cell development. While many of these studies have been completed in the model species of *S. cerevisiae* (Rabitsch *et al.*, 2001), *S. pombe* (Mata and Bahler, 2003), *C. elegans* (Jiang *et al.*, 2001), *Drosophila* (Andrews *et al.*, 2000), mouse (Pang *et al.*, 2006), Arabidopsis (Yin *et al.*, 2002) and rice (Wang *et al.*, 2005), there has also been pioneering studies that have examined polyploid species (for example, Crismani *et al.*, 2006).

An early study from the late 1990ʻs highlighted just how powerful such a technique can be in identifying new gene targets for understanding processes such as meiosis. Using cDNA microarrays containing 97% of the known *S. cerevisiae* genes, knowledge of meiotically-regulated genes was increased from the previously known ~150 (identified through conventional methods) to over 1,000 genes showing differential meiotic expression (Chu *et al.*, 1998). Further microarray analyses in yeast have investigated meiotic expression and the regulation of expression, with many novel genes or genes not previously known to be expressed in meiosis identified (Primig *et al.*, 2000; Mata *et al.*, 2002). This cache of novel meiotic expressed genes have since been a platform for many studies on how these genes function, and interact with one another and other gene networks, in budding yeast (Gerton *et al.*, 2000; Sarkar *et al.*, 2002; Averbeck *et al.*, 2005). While the research findings from such studies have been immeasurable for moving the knowledge-base of yeast meiosis forward, it has not necessarily meant that this is transferrable to more complex organisms. Consequently, we have started to see a shift in other research groups that work on non-model systems to undertake their own analysis (even though processes like meiosis are evolutionary well-

conserved). Highlighting this is a recent study by Crismani and colleagues (2006) who used the Affymetrix wheat GeneChip to investigate gene expression over seven stages of development (which included five stages during meiosis). Bearing in mind that the array did not contain representations for every gene from within the bread wheat genome (it was predicted to cover about 60% of the genome), the analysis still identified 1,350 transcripts which were temporally regulated during the early stages of meiosis, including some with annotated functions of chromatin condensation, synaptonemal complex formation, recombination and fertility (Crismani *et al.*, 2006). Using hierarchical clustering, Crismani and colleagues (2006) observed that 36.15% of the 1,350 transcripts fell under one of five expression profile clusters. Two cluster groups showed low expression over the beginning of meiosis and then underwent a sharp increase at the onset of mature anther development, with many of the transcripts in these two groups showing homology to genes known to have roles in flower development and associated pollen development (Crismani *et al.*, 2006).

The other three clusters (consisting of 350 transcripts) appeared to be meiotically-regulated with expression changes not as dramatic across the stages, suggesting that meiotic genes undergo subtle fluctuations rather than sharply increasing or decreasing (Crismani *et al.*, 2006). In particular, cluster groups IV and V (Figure 1.7) showed an increase during early meiosis with expression of these transcripts then tapering off as meiosis progressed (Crismani *et al.*, 2006). Of the 350 transcripts that are clustered into the three groups of transcripts with meiotic regulation, transcripts with no hits or annotations within the public databases had the highest representation (40.57%), followed by transcripts which have homology to genes with reported roles in meiosis or cell cycle processes (17.1%) (Crismani *et al.*, 2006). Consequently, these transcripts represent a source of novel wheat meiotic candidates that can be investigated further (Crismani *et al.*, 2006).

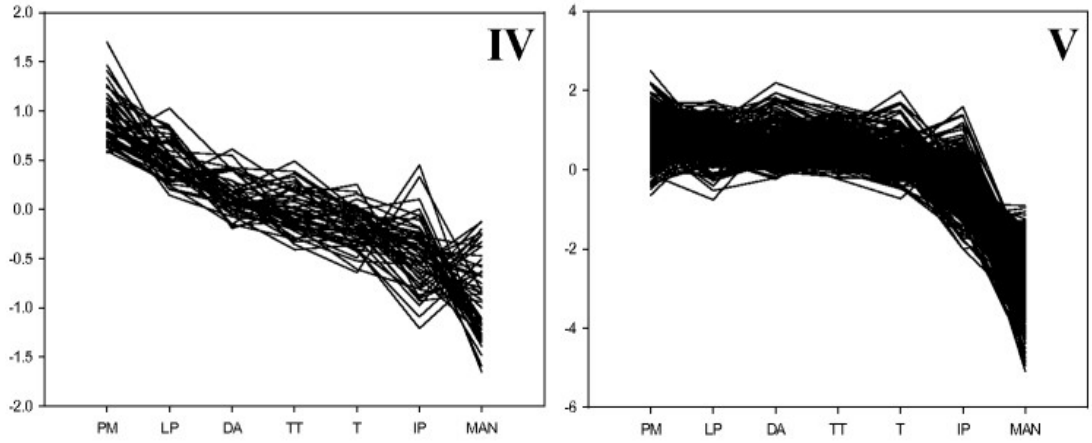


Figure 1.7 - Hierarchical clustering of meiotically-regulated transcripts from bread wheat. Clusters IV and V represent early meiotic-regulated genes, approximately 40% of which have no annotation (Crismani *et al*., 2006). Expression was profiled over meiotic stages of pre-meiosis (PM), leptotene to pachytene (LP), diplotene to anaphase I (DA), telophase I to telophase II (TT), tetrads (T), immature pollen (IP) and mature anthers (MAN). The Y axis for cluster groups is centered, log base 2, RMA normalised values. Figure adapted from Crismani *et al*., 2006.

1.3.2 – Dissecting gene function using gene mutants

After selecting gene candidates for further investigative research, one of the more common experimental avenues pursued is to study the proposed meiotic gene in a „nomal" or wild-type setting, by isolating the gene, and conducting gene expression analysis and protein localisation. Protein localisation can be determined several ways and typically depends on the level of resolution required. While western blot analysis is less informative than microscopy-based techniques, it still nonetheless enables the tissue specificity (or lack of) to be determined. Immunolocalisation on the other hand, is a very powerful method to determine if your protein of interest is locating to the meiotic chromosomes as many meiotic genes have been found to do (Armstrong *et al.*, 2002; Higgins *et al.*, 2005; Boden *et al.*, 2007; Boden, 2008; Boden *et al.*, 2009). Irrespective of these two techniques, both of them require a specific antibody for the protein of interest.

After establishing the profile of the gene and protein in wild-type, studying mutant lines for the gene of interest is a powerful way to further elucidate the biological significance. A number of different strategies are available to generate mutants in a particular organism.

Many mutant plant populations have been produced using either chemical (ethane methyl sulfonate [EMS] or nitrosomethyl-urea) or insertional mutagenesis (transposons, transfer DNA [T-DNA]). These populations have been used extensively in forward genetic strategies; with screens for a phenotype of interest, such as reduced fertility, and the resulting mutation identified and investigated (Hamant *et al.*, 2006). Reverse genetic approaches can also be used to decipher gene function, whereby RNAi or T-DNA mutagenesis results in disrupting the gene of interest. If available for your target species and gene T-DNA mutagenesis is a popular approach, providing easy and fast access to mutations in your gene(s) of interest. The SALK institute, for example, has over 88,000 publically available T-DNA mutants covering roughly 73% of the expected 29,000 genes in Arabidopsis (Alonso *et al.*, 2003). Insertional mutagenesis also allows for a wider-range of effects on the gene of interest to be created; allowing for knock-out, knock-downs, knock-ons, etc., all of which will produce different effects (Figure 1.8) (Krysan *et al.*, 1999).

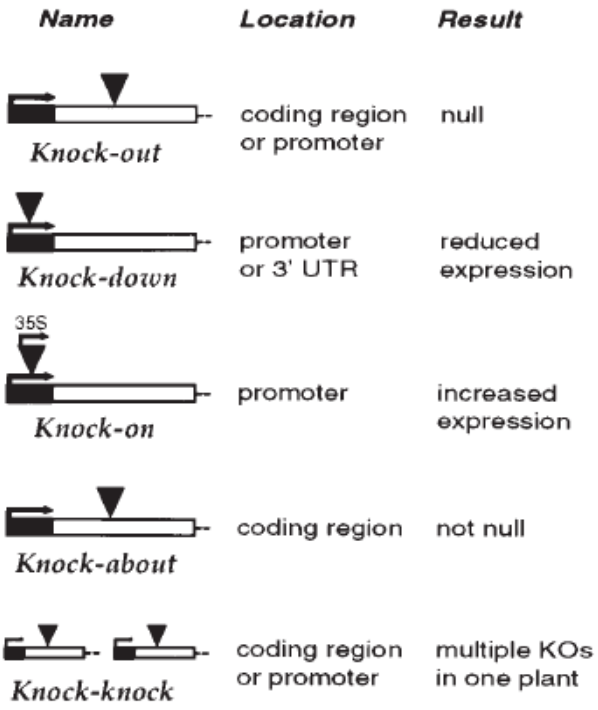


Figure 1.8 - "Knockology" of T-DNA insertional mutatagenesis. Insertion of T-DNA elements into different regions of a particular gene can lead to many different outcomes. Coding regions of genes are shown as a white box, while the promoter is the black region, with the arrow and the T-DNA element represented by the black triangle. KOs = knockouts, UTR = untranslated region. Figure from Krysan *et al*., (1999).

Techniques like those listed above are especially useful if there is limited or no prior knowledge on the gene being studied. Using mutations in a gene of interest enable researchers to observe the effects of the absence of the gene or reduction of gene expression; which allows for the gene function to be determined or at least inferred. Visualisation of the DNA (using stains like DAPI) within mutants allows for the determination of where a particular gene is acting based on the defect in meiosis. For example if chromosome fragmentation is observed, the gene then has a high probability of being involved in recombination and DNA repair mechanisms, such as *AtXRCC3* (Bleuyard and White, 2004). Likewise, mutants where genes are involved in synapsis display defects such as unpaired and unsynapsed chromosomes during early meiosis and ultimately results in univalents and uneven chromosome segregation during anaphase I, such as seen in *Atasy1* mutants (Ross *et al.*, 1997).

Further understanding of a meiotic gene of interest can be gained by looking at other known meiotic genes in the mutant plant. An example of this is the recent investigation of the known meiotic genes RAD50 and RAD51 in the *phs1* mutants in maize and Arabidopsis (Pawlowski *et al.*, 2004; Ronceret *et al.*, 2009). Within the *phs1* mutants these two known meiotic proteins, while still being present, have been shown not to enter the nucleus as they do in wild-type meiosis. This provided evidence of a role for PHS1 in the processes of shuttling these two proteins into the nucleus and onto the chromosomes (Pawlowski *et al.*, 2004; Ronceret *et al.*, 2009). Therefore the investigation of other known proteins in a mutant background can provide strong evidence of the mutated genes function during the process of meiosis.

1.3.3 – Understanding function by protein partners

As previously mentioned, using meiotic mutants and observing effects of other genes allows for the formation of protein networks such as RAD50 or RAD51 and PHS1; where in the *phs1* maize mutant disrupted loading of RAD50 and RAD51 is observed, thus providing evidence of a link between PHS1 and two meiotic proteins (Pawlowski *et al.*, 2004). While useful, the suggestion of a link between the proteins does not determine whether there is an actual physical interaction between PHS1 and RAD50 or PHS1 and RAD51. Using protein localisation methods, such as fluorescent immunolocalisation, will allow the identification of whether two proteins are co-located to the same area. However, just because two proteins are co-located does not mean that they physically interact. For example the inner SC protein ZYP1 and the intra-sister chromatid cohesion protein REC8 (ADF1) are co-located to meiotic chromosomes in Arabidopsis (Ronceret *et al.*, 2009) yet they have no recorded physical link to date. Having information of physical protein interactions and protein partners gives stronger evidence for protein functionality, for example genes/proteins which interact with a recombinase protein is likely to also be involved in recombination or a very similar process. Therefore the investigation of protein interactions, and large scale protein interaction networks, provides a powerful tool for determining the functionality of novel genes with meiotic expression/regulation.

There are a number of different techniques for determining physical protein interactions, including affinity and molecular-sized based chromatography, and split-GFP assays (Lalonde *et al.*, 2008). However these techniques fail to analyse large-scale protein networks effectively as they are based on interactions between two known and isolated proteins (Mukherjee *et al.*, 2001). A technique which has become popular for the identification and characterisation of large-scale protein networks is the eukaryotic *in vivo* yeast two-hybrid approach. This technique allows the user to screen a cDNA library with a single protein which enables the identification of multiple protein interactions in the one

experiment (Guan and Kiss-Toth, 2008), thus being a powerful tool for the identification of novel protein interactors. In conjunction, one of the most appealing features of the yeast two-hybrid system is that in the process of confirming an interacting protein, this results in the isolation and cloning of at least part of the corresponding gene (Mukherjee *et al.*, 2001).

1.4 – Known meiotic genes

As already discussed, the processes of homologous chromosome pairing, recombination and synapsis are important and closely linked. While these processes and the genes/proteins involved have been extensively studied in diploids, such as yeast and Arabidopsis (Thompson and Stahl, 1999; Keeny, 2001; Hamant *et al*., 2006), the mechanics of these processes remains somewhat unresolved for more complex organisms like bread wheat. During meiosis in wheat, strict pairing between homologous chromosomes occurs to the exclusion of the highly similar homoeologous chromosomes. This system therefore provides a unique opportunity to further characterise and understand how the genes involved in homologous recognition and recombination function. The following sections focus on two meiotic proteins, ASY1 and DMC1 (along with the DMC1 protein partner complex of HOP2:MND1), which play a role in the processes of homology recognition, chromosome pairing and recombination in higher eukaryotes and could be expected to have a similar role in these processes in wheat.

1.4.1 – *Asynapsis 1 (ASY1)*

In a forward genetics yeast screen designed to detect mutations defective in homologous chromosome pairing, a recessive mutation termed *hop1* was identified; where electron microscopy revealed that the homozygous *hop1* yeast failed to form the SC structure between chromosomes, which as previously mentioned provides a structural basis for homologous pairing (Hollingsworth and Byers, 1989). The *HOP1* gene was then further found to be transcriptionally-regulated during sporulation and the protein was found to

localise along the length of the meiotic yeast chromosomes (Hollingsworth *et al.*, 1990). Concurrently, a separate yeast mutation *red1* was found to belong to the same epitasis group as the *hop1* mutation, suggesting an additional gene in the same pathway as HOP1 (Rockmill and Roeder, 1990). These two proteins have since been shown to physically interact with one another (Niu *et al.*, 2005).

While several years after the Hollingsworth and Byers (1989) discovery, a homologue of HOP1 (known as ASY1) from plants was identified in Arabidopsis using T-DNA insertional mutagenesis (Ross *et al.*, 1997). The full *At*ASY1 homologue was isolated and found to have 28% identity and 51% similarity to *Sc*HOP1 with the highest homology being located to the N-terminal end of the protein (roughly 250 amino acids) (Caryl *et al.*, 2000). While not meiotic-specific, *AtASY1* transcript expression was up-regulated in bud (meiotic) tissues (Caryl *et al.*, 2000); however protein presence was only detected in buds undergoing early meiosis (Armstrong *et al.*, 2002). The *At*ASY1 protein in meiotic cells was found to localise to meiotic chromosomes; starting with punctuate foci at meiotic interphase (Figure 1.9 A), and developing during leptotene to appear as discontinuous stretches of *At*ASY1 signal. By zygotene (where the meiotic chromosomes are almost completely synapsed) the *At*ASY1 signal is fully associated along the entire length of the lateral elements of the SC (Figure 1.9 B) (Armstrong *et al.*, 2002).

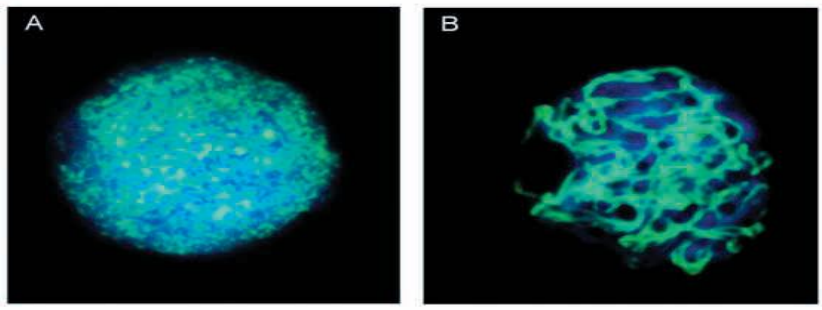


Figure 1.9 - *At*ASY1 protein localisation on meiotic chromosomes. Immunolocalisation of *At*ASY1 (*green* signal) on spread Arabidopsis pollen mother cells with DNA counterstained using DAPI (4′, 6′-diamidino-2-phenylindole, [DAPI], *blue* signal). (**A**) Interphase/early leptotene showing numerous punctuate foci. (**B**) Zygotene showing a continuous *At*ASY1 signal along the chromosome axes. Figure adapted from Armstrong *et al*., 2002.

ASY1 homologues have also been identified in rice (*Oryza sativa*) and bread wheat (*Triticum aestivum*) (Nonomura *et al.*, 2004; Boden *et al.*, 2007). Similar to the yeast *HOP1* gene, the rice homologue *OsPAIR2* (*p*airing *a*berration *i*n *r*ice meiosis *2*) was identified using mutation screens; with the *pair2* mutant displaying a complete lack of homologous chromosome pairing in pachytene and diakinesis due to a failure to undergo extensive synapsis (Nonomura *et al.*, 2004). The isolated *PAIR2* gene and predicted protein was shown to have 52.9% identity to *At*ASY1 and 16.1% identity to *Sc*HOP1. The isolated wheat *ASY1* (*TaASY1*) gene and the predicted protein were also found to have high similarity to the other *ASY1* homologues, with the wheat protein displaying 80% sequence identity to *Os*PAIR2, 53.8% to *At*ASY1 and 16.5% to *Sc*HOP1 (Boden *et al.*, 2007). As with the yeast and Arabidopsis homologues of *ASY1*, the *Ta*ASY1 protein was found to locate to meiotic chromosomes and specifically the axial elements of the SC (Figure 1.10) (Boden *et al.*, 2007; Boden *et al.*, 2009). A common feature of all the ASY1 homologues is the presence of the HORMA protein domain (HOP1, REV7, MAD2) (Aravind and Koonin, 1998), which facilitates interaction with chromatin during mitosis and meiosis, including chromatin with DNA double stranded breaks (Hollingsworth and Byers, 1989; Hollingsworth *et al.*, 1990; Zetka *et al.*, 1999; Armstrong *et al.*, 2002; Boden *et al.*, 2007).

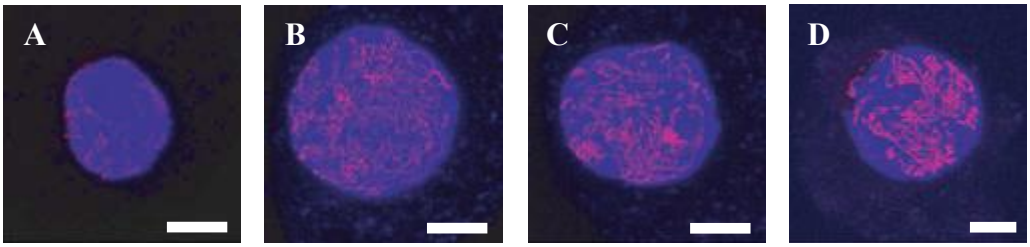


Figure 1.10 - *Ta*ASY1 localises to meiotic chromosomes. *Ta*ASY1 (*red*) localises to chromosomes (DNA counterstained with DAPI, *blue*) during the meiotic stages of (**A**) leptotene, (**B**) early zygotene, (**C**) late zygotene, (**D**) pachytene. Bars, 10 μm. From Boden *et al.*, (2009).

As previously mentioned, yeast HOP1 was found to physically interact with another meiotic protein, RED1, and this interaction in turn results in the phosphorylation of HOP1 by the kinase protein MEK1 (with phosphorylation of HOP1 being dependent on the presence of DSBs) (Niu *et al.*, 2005). The phosphorylation of the HOP1 protein is critical for the functionality/activation of the protein in successfully maintaining the barrier to sister chromatid repair or the DMC1 protein independent recombination/repair pathway (Niu *et al.*, 2005). A role for ASY1/HOP1 in the DMC1-dependent inter-homologue recombination pathway was further confirmed when Arabidopsis ASY1 was found to facilitate the presence of the DMC1 protein on the meiotic chromosomes; with rapid reduction of DMC1 foci on the chromatin; thus suggesting that DMC1 fails to form stable associations in the absence of ASY1 (Sanchez-Moran *et al.*, 2007).

The hypothesis of ASY1 promoting interactions between homologous chromosomes in plants has recently been substantiated further by Boden and colleagues (2009). In transgenic wheat plants where reduced *TaASY1* expression was observed (transgenics were generated through RNAi), multiple chromosome associations were observed at metaphase I (Boden *et al.*, 2009). As previously mentioned, the absence of the yeast *HOP1* homologue results in decreased inter-homologue recombination conducted by DMC1 but increased sister chromatid repair and non-homologous chromosome associations (Carballo *et al.*, 2008). Therefore in the context of wheat and the reduction of *TaASY1*, it could be possible that the sister chromatid repair is confusing the homoeologous sister chromatid with a homologous partner‟s sister chromatid, thus leading to the observed multiple chromosome associations (Boden *et al.*, 2009).

Interestingly, the chromosome associations observed in the *Taasy1* RNAi mutants is a phenocopy of the well known wheat chromosome pairing mutant *ph1b* (*P*airing *h*omoeologous *1*); however surprisingly *TaASY1* expression was actually found to be increased in the *ph1b* mutant (Figure 1.11) (Boden *et al*., 2009). How is it that both a reduction and an increase in the same protein cause a similar phenotype? As previously

mentioned the reduction in *TaASY1* could be allowing sister chromatid repair to occur, giving rise to multiple chromosome associations (Boden *et al.*, 2009). In contrast, an increase in *Ta*ASY1 could be leading to the over „stimulation" of inter-homologue recombination. This over stimulation of *TaASY1* may be promoting recombination with any chromosome which is close enough in similarity, given the high similarity and the close proximity of homoeologous chromosomes through associations at the onset of meiosis in wheat (Martínez-Pérez *et al.*, 2003).

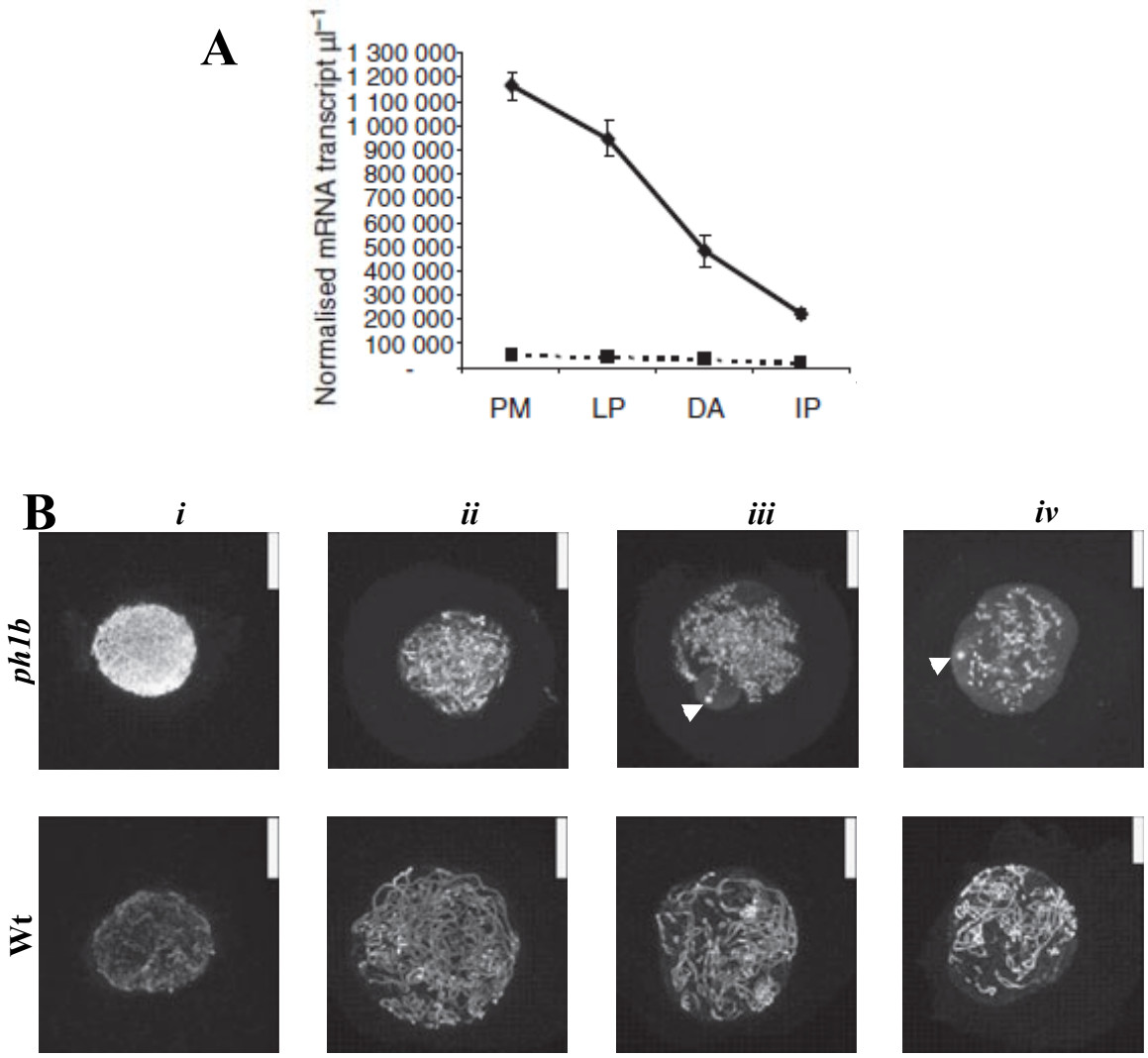


Figure 1.11 - Disruption of *Ta*ASY1 in the *ph1b* mutant. (**A**) *TaASY1* transcription is up-regulated 20-fold in the *ph1b* mutant (solid line) compared to Wt (wild-type) wheat (broken line). (**B**) Increased and disordered loading of the *TaASY1* (white signal) protein in the *ph1b* mutant during leptotene (*i*), early zygotene (*ii*), late zygotene (*iii*) and pachytene (*iv*). White arrowheads indicate ASY1 polycomplexes in the *ph1b* mutant. Bars = 10µm. From Boden *et al*. 2009.

1.4.2 – *Disrupted Meiotic cDNA 1 (DMC1)*

The recombinase protein DMC1 is very important for recombination between homologous chromosomes (Schwacha and Kleckner, 1997). Originally discovered and characterised in a yeast mutant screen, DMC1 is a homologue of the bacterial recombination protein RecA; with *dmc1* mutants lacking correct resolution of DSBs and correct formation of the tripartite SC (Bishop *et al.*, 1992). The yeast DMC1 gene is meiotic specific, with induction of expression occurring three hours into meiosis (Bishop *et al.*, 1992). In contrast, the DMC1 gene in plant species is not meiosis specific; although expression during meiosis is significantly higher when compared to vegetative tissues (Doutriaux *et al.*, 1998; Ding *et al.*, 2001). Subsequent research on the DMC1 protein has shown that it is associated with chromatin during meiosis, forming punctuate foci (in contrast to the linearisation of the ASY1 protein) which peak during zygotene (Figure 1.12) (Bishop, 1994; Terasawa *et al.*, 1995; Tarsounas *et al.*, 1999).

Figure 1.12 - DMC1 localisation on meiotic chromosomes during zygotene in *Lilium longiflorum*. Punctuated foci of DMC1 protein (*red* signal) on meiotic chromosomes during the prophase I sub stage zygotene (**A**). (**B**) and (**C**) expanded portions of (**A**). Chromosomes (*blue* signal) stained with DAPI. Figure adapted from Terasawa *et al.*, 1995.

The *DMC1* gene and protein is highly conserved across many eukaryotic species, and highly related to another recombinase protein RAD51 (Lin *et al.*, 2006; Khoo *et al.*, 2008). Functional studies in yeast further confirmed that the DMC1 protein had a highly conserved

function with the bacterial RecA protein; with DMC1 binding preferentially to single-stranded DNA (ssDNA) and the ability to conduct strand invasion independently of other proteins, the assimilation of single-stranded DNA into homologous super-coiled double-stranded DNA, which is representative of recombination between homologous chromosomes during meiosis (Figure 1.13) (Hong *et al.*, 2001; Kant *et al.*, 2005; Henry *et al.*, 2006). Conservation of the DMC1 protein sequence and function has been found in other higher eukaryotic species, including humans (Sehorn *et al.*, 2004) and rice (Kant *et al.*, 2005). DNA repair and/or resolution via recombination requires the DMC1 protein to interact and cooperate with a number of other proteins; including other recombination/repair proteins such as RAD51, BRCA2, MSH4, TopoII and Rad54B (Neyton *et al.*, 2004; Siaud *et al.*, 2004; Iwabata *et al.*, 2005; Sarai *et al.*, 2006) and other genes which do not have a particular known function in recombination or repair function such as HOP2 and MND1 (Chen *et al.*, 2004; Petukhova *et al.*, 2005; Enomoto *et al.*, 2006; Pezza *et al.*, 2006).

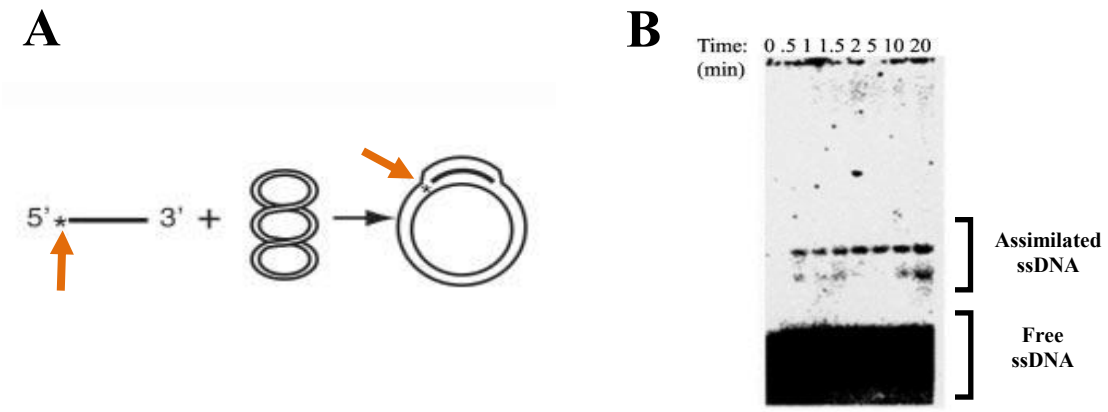


Figure 1.13 - *Sc*DMC1 is able to conduct homologous strand invasion. (**A**) Diagram of the D-loop formation assay. Asterisks (indicated by orange arrows) represent the ^{32}P-labeled end of the ssDNA. Figure adapted from Enomoto *et al.*, 2006. (**B**) Time course analysis of strand assimilation by *Sc*DMC1. Figure adapted from Hong *et al.*, 2001.

1.4.3 – The HOP2:MND1 meiotic complex

1.4.3.1 *Ho*mologous *p*airing 2 (*HOP2*)

In 1998, Leu and colleagues reported on the characterisation of mutants from a novel budding yeast gene, termed *HOP2* (for *ho*mologous *p*airing). The *hop2* null mutant produced a phenotype distinct from other yeast meiotic mutants characterised previously; with chromosomes undergoing nearly wild-type levels of synapsis but occurring mostly between non-homologous chromosomes and that multiple chromosome associations were seen (Leu *et al.*, 1998). *HOP2* homologues have been identified in other eukaryotes including mouse *TBPIP* (*T*ATA box *b*inding *p*rotein *i*nteracting *p*rotein) (Tanaka *et al.*, 1997), fission yeast *MEU13* (for *m*eiotic *e*xpression *u*pregulated) (Nabeshima *et al.*, 2001) and Arabidopsis *AHP2* (*A*rabidopsis *h*omologue *p*airing 2) (Schommer *et al.*, 2003); with the fission yeast and Arabidopsis mutants displaying similar phenotypes to the originally described budding yeast mutant. Within the *hop2* mutant DSBs were still occurring but had deficiencies in being repaired, while DMC1, though loaded correctly onto the chromosomes failed to disassociate from the chromatin (Figure 1.14) (Leu *et al.*, 1998). This provided strong evidence that the *HOP2* homologue played a role in DMC1-mediated recombination processes.

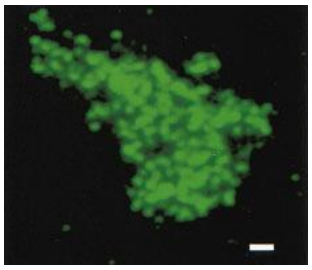


Figure 1.14 - DMC1 protein localisation at pachytene in a yeast *hop2* mutant. Anti-DMC1 antibody (*green* signal) shows DMC1 being associated with chromatin at pachytene in the *hop2* mutant, which is later than the stage at which DMC1 is normally associated. In wild-type cells, DMC1 would have already been removed from the chromatin by this stage. Scale bar = 1μm. Figure adapted from Leu *et al*., (1998).

Yeast *HOP2* expression was shown to be meiosis-specific, with expression increasing rapidly upon the induction of meiosis and reaching maximum expression eight hours after induction (Leu *et al.*, 1998). In contrast to *HOP2* in yeast but similar to other plant meiotic genes, *AHP2/HOP2* was expressed in all tissues but had highest expression in meiotic tissues (Schommer *et al.*, 2003). Through the use of an anti-HOP2 protein antibody, Leu and colleagues (1998) were able to show that the HOP2 protein associated with chromatin at zygotene and pachytene (Figure 1.15) and is independent of DSB formation.

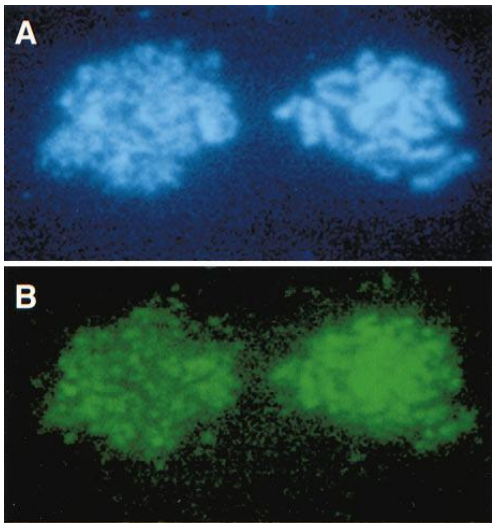


Figure 1.15 - The HOP2 protein is shown to localise to meiotic chromosomes. Wild-type meiotic chromosomes at zygotene (*left*) and pachytene (*right*) are stained with DAPI (*blue*, **A**) to show chromatin structure. Anti-HOP2 protein antibody (*green*, **B**) shows HOP2 localising to chromatin. Figure adapted from Leu *et al*., (1998).

1.4.3.2 *M*eiotic *n*uclear *d*ivision (*MND1*)

A previously known meiotic protein from budding yeast, termed MND1 (*m*eiotic *n*uclear *d*ivision 1), was found to interact with the aforementioned HOP2 protein (Tsubouchi and Roeder, 2002). In yeast *mnd1* mutants, chromosome pairing frequencies are much lower than in wild-type nuclei, which is similar to *MND1*‟s protein partner *hop2* mutant and the double mutant *hop2/mnd1* (Tsubouchi and Roeder, 2002). This phenotype is also observed in the Arabidopsis *mnd1* mutants (Kerzendorfer *et al.*, 2006; Panoli *et al.*, 2006). Comparing

yeast *mnd1* mutants with the *hop2* mutants, synapsis was not observed at the same levels with the SC protein ZIP1 not being loaded in the *mnd1* mutants to the same extent as in the *hop2* mutants (Figure 1.16) (Tsubouchi and Roeder, 2002). Recombination within the *mnd1* mutants was also found to be reduced (Tsubouchi and Roeder, 2002) but was not due to the absence of RAD51, which was found to actually accumulate on the chromosomes (Figure 1.16) (Tsubouchi and Roeder, 2002); indicating that the MND1 protein is also involved in the recombination pathway.

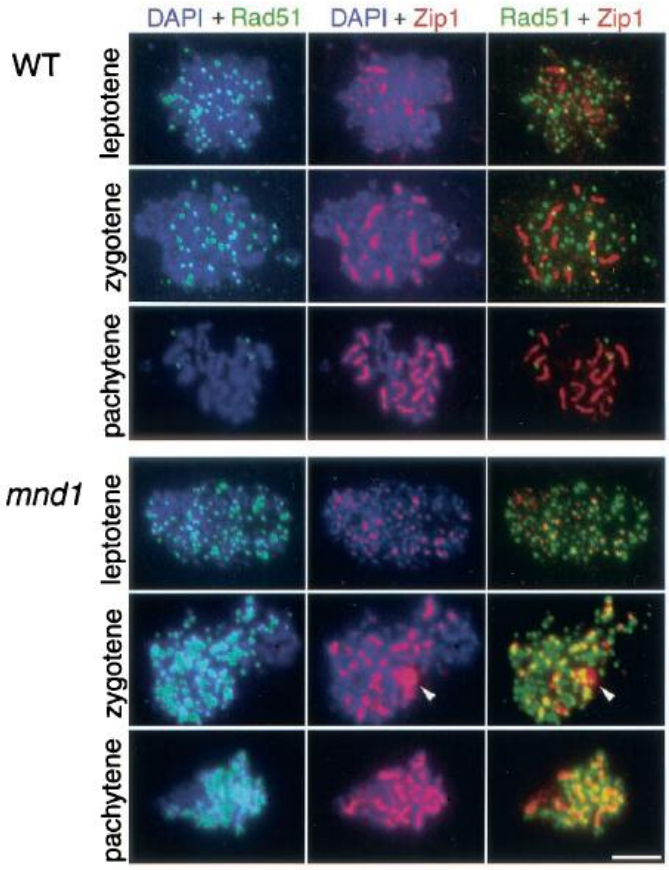


Figure 1.16 - Disrupted loading of RAD51 and ZIP1 on meiotic chromosomes in yeast *mnd1* mutants. RAD51 (*green* signal) and ZIP1 (*red* signal) proteins were affected in yeast *mnd1* mutants; with RAD51 accumulating on chromosomes (DAPI, *blue* signal) while ZIP1 was not loaded onto chromosomes to the same extent as wild-type (WT). Arrows indicate a polycomplex. Figure adapted from Tsubouchi and Roeder (2002).

As with *HOP2*, *MND1* has a similar expression profile with induced expression during meiosis (Chu *et al.,* 1998; Rabitsch *et al.*, 2001). Expression of the *MND1* Arabidopsis homologue was found to be similar to the expression of *AtAHP2/HOP2*, with higher expression in meiotic tissues and lower expression in vegetative tissues (Kerzendorfer *et al.*, 2006; Panoli *et al.*, 2006). The MND1 protein was also shown to be localised to chromatin in yeast, and to co-localise with the HOP2 protein (Figure 1.17 A) (Tsubouchi and Roeder, 2002). In the absence of either gene, the localisation of the other corresponding proteins was affected (Figure 1.17 B); thus indicating the loading of either of these two proteins onto the chromatin is dependent on the presence of the other (Tsubouchi and Roeder, 2002). Interaction between the MND1 and HOP2 yeast proteins was confirmed when immunoprecipitation experiments using meiotic cells isolated both proteins (Tsubouchi and Roeder, 2002). Direct interaction was further confirmed by Chen and colleagues (2004) in yeast via cross-linking assays and by Kerzendorfer and colleagues (2006) between the Arabidopsis plant homologues of MND1 and HOP2/AHP2 proteins through a yeast two-hybrid assay; thus confirming the physical interaction between the two proteins and the formation of the complex observed in other species.

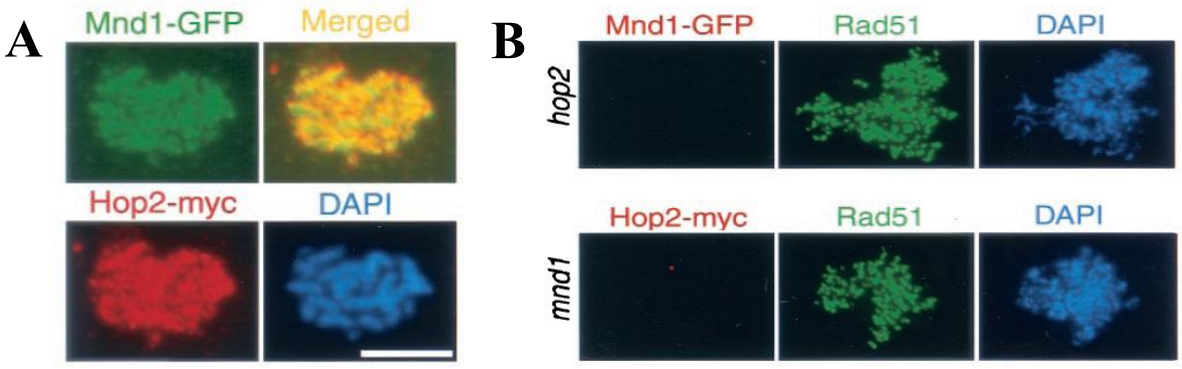


Figure 1.17 - MND1 protein co-localises with HOP2 on meiotic chromosomes. Meiotic nuclei stained with antibodies for the tagged MND1 (GFP, *green*) and HOP2 (myc, *red*) proteins, showing that both localise to chromatin (DAPI, *blue*). Merged, MND1 and HOP2 show that the two proteins co-localise on chromatin (**A**). Loading of either MND1 or HOP2 onto chromatin is dependent on the presence of both proteins (**B**, *displaying hop2 and mnd1 mutants*). Figure adapted from Tsubouchi and Roeder (2002).

1.4.3.3 HOP2 and MND1 proteins act as a complex during meiotic processes

In budding yeast the HOP2 and MND1 proteins are known to form a stable heterodimer which binds directly to DNA, with a preference for double stranded DNA (dsDNA) (Figure 1.18 A) (Chen *et al.*, 2004). This preference for dsDNA is also observed in the human form of the HOP2-MND1 complex (Figure 1.18 B) (Enomoto *et al.*, 2006). In contrast the mouse HOP2:MND1 complex binds preferentially to ssDNA (Figure 1.18 C) (Petukhova *et al.*, 2005; Pezza *et al.*, 2006). Incidentally, Pezza and colleagues (2006) also showed that HOP2 and MND1 (independently) showed a preference for ssDNA.

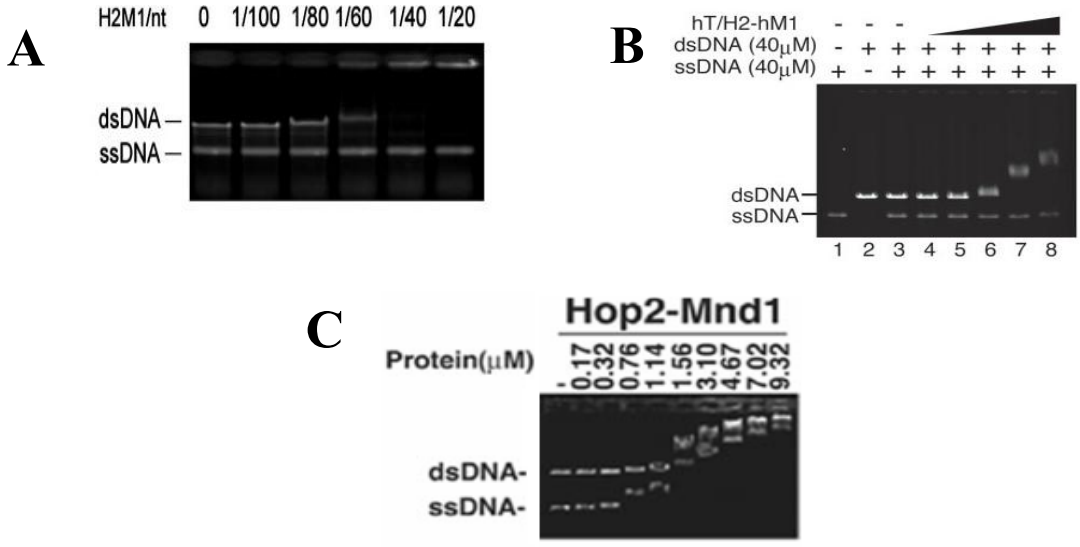


Figure 1.18 - HOP2:MND1 complex DNA binding ability tested using electrophoresis mobility assay. The HOP2:MND1 protein complex in yeast (**A**, Chen *et al*., 2002) and humans (**B**, Enomoto *et al*., 2006) binds almost exclusively to dsDNA. In contrast, the mouse HOP2:MND1 protein complex (**C**) binds both forms of DNA with a preference to ssDNA. Figure adapted from Pezza *et al*., (2006).

As previously mentioned the DMC1 protein has been shown to bind preferentially to single-stranded DNA (ssDNA) and to conduct strand invasion independently of other proteins (Figure 1.19 A) (Kant *et al.*, 2005; Henry *et al.*, 2006). The addition of the HOP2:MND1 complex into the *in vitro* strand invasion assays shows that the presence of the yeast HOP2:MND1 complex increases the amount of DMC1-mediated strand invasion (D-loop

formation as per figure 1.13 A) (Figure 1.19 B) (Chen *et al.*, 2004); with an increase also observed when using the mouse HOP2:MND1 complex (Petukhova *et al.*, 2005). Interestingly the mouse HOP2 protein alone was also able to form D-loops in *in vitro* assays which was not the case for either MND1 or the HOP2:MND1 complex, thus suggesting that the complex alters the activity of the HOP2 protein (Petukhova *et al.*, 2005). However, HOP2 activation of cross-overs, and thus strand invasion, is not seen in budding yeast as *dmc1/mnd1* double mutants do not show a level of cross-over formation (Henry *et al.*, 2006).

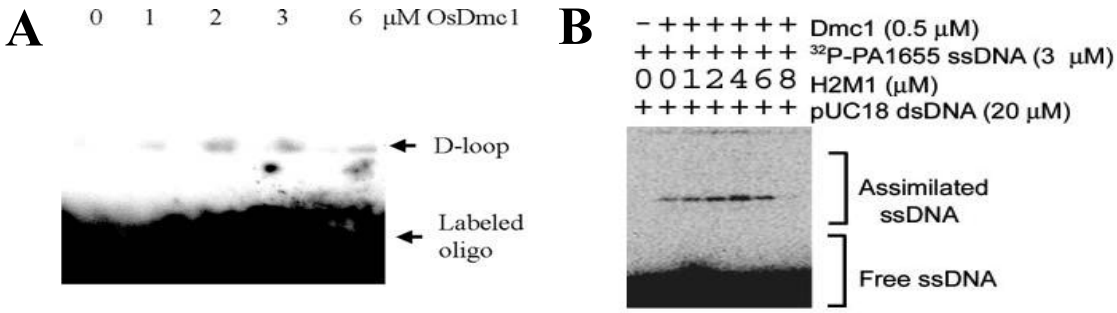


Figure 1.19 - HOP2:MND1 complex stimulates the DMC1-mediated D-loop formation. DMC1 in rice (*Os*DMC1) has been shown to be all that is required for D-loop formation (**A**). Figure from Kant *et al*. (2005). This has also been reported in yeast (**B**, Lane 2), however with the addition of the HOP2:MND1 complex (up to 4μM) the efficiency of the DMC1-mediated D-loop formation was greatly increased (**B**, Lanes 3-5). Figure from Chen *et al*., (2004).

Henry and colleagues (2006) have shown that HOP2:MND1 and RAD51 are important for preventing single strand invasion recombination between non-homologous sequences/chromosomes. In yeast there are two particular genes, *BUD23* and *EBP2* on chromosomes III and XI, respectively, that contain an imperfect CTT repeat site and which undergo very low levels of ectopic recombination naturally (Henry *et al.*, 2006). When the wild-type level of ectopic recombination was compared to levels which occurred in yeast strains with *mnd1*, *hop2* and/or *rad51* and *dmc1* mutated, it was found that the levels of ectopic recombination greatly increased when the HOP2:MND1 complex was disrupted and when Rad51 was absent (Figure 1.20) (Henry *et al.*, 2006). This implies that the HOP2:MND1 complex is functioning in an aspect of homology recognition and/or

stabilisation of interactions between homologous chromosomes during the important strand invasion process conducted by the DMC1 protein (Henry *et al.*, 2006).

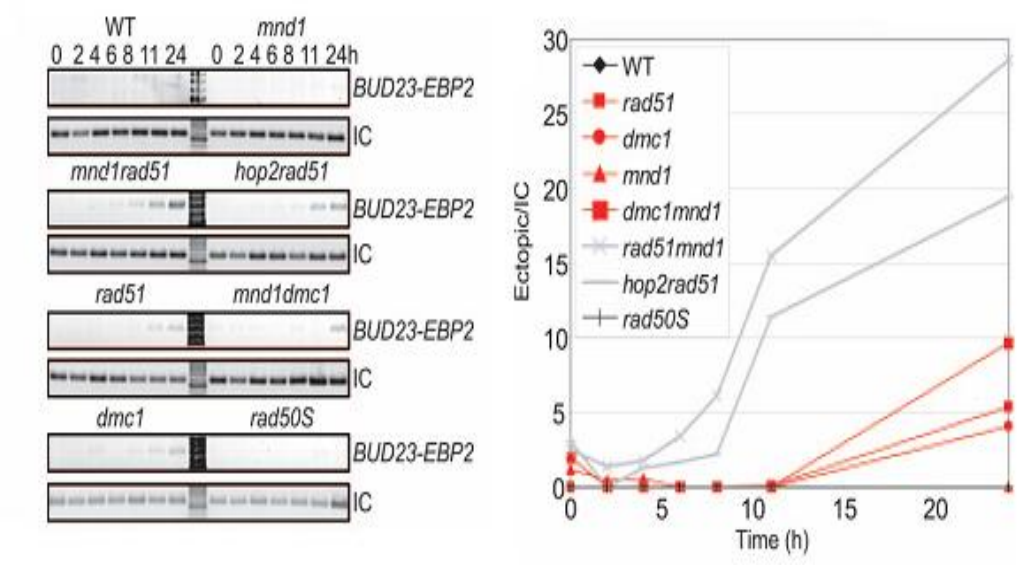


Figure 1.20 - Ectopic recombination occurs between non-homologous chromosomes when MND1, HOP2 and/or RAD51 are absent. Equal amounts of DNA collected at various time points after induction of meiosis were used as template for a 33 cycle PCR that amplified stable strand invasion events between BUD23 and EBP2 (*left panel*). Quantitation of the results from the left panel shown in the graph, were derived by dividing the amount of input DNA as measured by their Internal Control (IC) to normalise for the amount of DNA added to the reaction. The graph shows that when either MND1 or HOP2 and RAD51 are absent, ectopic recombination occurs at increased levels. This indicates that the HOP2:MND1 complex and RAD51 can prevent recombination events between non-homologous chromosomes which is unable to be done with DMC1 alone. Figure adapted from Henry *et al*., (2006).

1.5 – Rationale of current study

As highlighted throughout this literature analysis meiosis is a complex and integrated process. Furthering our understanding of meiosis over the next decade will therefore require a significant investment of time and money across a wide-range of organisms. In further developing our understanding of meiosis in bread wheat, which has three closely related genomes, it is anticipated that genes and proteins involved in homology recognition, recombination and chromosomes pairing during prophase I may provide more insight into gene and protein function which may not be possible in diploid organisms. Using bread

wheat, along with Arabidopsis, this current study seeks to take a number of meiotic candidates through the „analytical pipeline" to achieve multiple goals.

The first component aims to build off the solid platform of microarray analysis previously conducted by Crismani and colleagues (2006), in characterising novel candidates with meiotic expression. These candidates will be taken through a number of filtering steps, including expression comparisons to the two known meiotic genes ASY1 and DMC1, which will result in the identification of a small number of novel candidates which have a high probability in having a role during meiosis. The resulting novel wheat candidates will be investigated in Arabidopsis using the T-DNA insertion derived mutants of putatively homologous genes. Investigation of these T-DNA mutants via a number of analysis steps, including reproductive morphology and meiotic gene expression, will lead to understanding whether these novel gene candidates have a role during meiosis in bread wheat.

The second component of this dissertation is to investigate a small number of key meiotic genes which are known to be involved during the early stages of prophase I in other diploid species to further the understanding of their activity and roles during meiosis in bread wheat. The first of which is the important recombination protein DMC1, and its protein partner complex HOP2:MND1, while the second candidate is the known pairing/synapsis gene *TaASY1*. While transcript expression, protein expression and localisation, and RNAi analysis have confirmed that *Ta*ASY1 plays an important role during meiosis in bread wheat (Boden *et al.*, 2007; Boden *et al.*, 2009); little is known about the proteins that interact with this.

It is envisaged that as a result of the knowledge gleaned from this dissertation, our understanding on a number of meiotic genes and novel candidates with meiotic gene expression in bread wheat, will be enhanced. Broadly, the research presented here aspires to contribute to the application of accelerating agricultural-based breeding programmes; which will in time, enable introgression of new traits rapidly and effectively between species and allow for the production of fertile hybrid offspring.

Chapter 2

Chapter 2 - Microarray transcription profiling is effective for identifying novel genes with meiotic expression in bread wheat

2.1 – Introduction

Identification of genes which are differentially expressed during meiosis has been achieved using a variety of approaches (Burns *et al.*, 1994; Kupiec *et al.*, 1997). Classically the technique of choice has involved untargeted mutations and screening by meiotic recombination assays (Rockmill and Roeder, 1990), or visual assays (Esposito *et al.*, 1972; Xu *et al.*, 1995). Using such assays the yeast meiosis community has identified over 150 genes involved in meiosis (Kupiec *et al.*, 1997). However with the introduction of microarray technology, the expression and regulation of thousands of genes can be observed during meiosis. One of the first examples reporting the power of DNA microarray technology contained 97% of the known or predicted yeast genes (Chu *et al.*, 1998). This array enabled the temporal program of gene expression during meiosis and spore formation to be investigated and reported that more than 1000 out of 6200 protein encoding genes had significant changes in mRNA levels during sporulation (Chu *et al.*, 1998). A subsequent study of the core meiotic transcriptome in yeast, using transcript expression variation occurring during meiosis and sporulation in two genetically distinct strains, identified a further 600 genes which were also meiotically regulated (Primig *et al.*, 2000). It has been through this technology platform that the number of genes shown to be involved in yeast meiosis has increased from approximately 150 to more than 1600, thus highlighting the value of such a powerful approach. Subsequent studies have since used microarrays to identify meiotic transcripts and/or understand the transcriptional regulation of meiosis in their species of interest; for example *Drosophila* (Andrews *et al.,* 2000), *Caenorhabiditis elegans* (Reinke *et*

al., 2000), rat (Schlecht *et al.*, 2004), rice (Wang *et al.*, 2005), maize (Ma *et al.*, 2008) and wheat (Crismani *et al.*, 2006).

The recent study conducted by Crismani and colleagues (2006) used the Affymetrix wheat GeneChip to observe gene expression over seven stages of reproductive development (which included five stages from meiosis). Although the array is predicted to be representative of approximately 60% of the wheat genome (given that this species has not been fully sequenced), Crismani *et al.* (2006) identified 1,350 transcripts which were temporally regulated during the early stages of meiosis. A number of the meiotically regulated transcripts have annotated functions for chromatin condensation, synaptonemal complex formation, recombination and fertility (Crismani *et al.*, 2006). However an interesting observation from this study was that there were 1094 transcripts deemed to have either no annotations or purely predicted annotations.

Based on this solid platform to work from, the objective of the research in this chapter was to reanalyse the wheat GeneChip data (given that since the original analysis was conducted, sequence contributions to databases have increased exponentially) and identify genes that still have either no or putative annotations, which may play a role in wheat meiosis. This was achieved by identifying novel candidates, selectively filtering these candidates, scrutinising the individual expression profiles, locating where these candidates reside in the wheat genome using Southern analysis, and two rounds of sequence similarity searching in order to obtain a small number of candidates for further in-depth analysis.

2.2 – Material and Methods

2.2.1 – Identification of novel candidates

Crismani and colleagues (2006) previously classified 1350 transcripts to be meiotically regulated over the first four stages of the meiotic time course investigated; pre-meiosis (PM), leptotene to pachytene (LP), diplotene to anaphase I (DA) and telophase I to

telophase II (TT). From sequence similarity known at the time of investigation, it was found that 1094 of these transcripts had either no annotation (826), or predicted annotations (268). These candidates were therefore used as the focus of data filtration to significantly reduce the number to a subset anticipated to include novel transcripts that would have roles during meiosis in bread wheat.

2.2.2 – Comparative expression profiling

To determine the transcript‟s expression trend rather than to focus on the absolute values, the data was centred by removing the average expression value for each transcript across the time-course. *K*-means clustering was conducted on the 1094 transcripts, using 10 to 30 partitions (k = 10 to 30) (Acuity 4.0, Axon Instruments, USA). *K*-means clusters were scrutinised for partitions which returned profiles with increased average expression during pre-meiosis (PM), leptotene to pachytene (LP), and diplotene to anaphase I (DA). These clusters were then further scrutinised, identifying those groups which showed abundant levels of expression (above 9, log base 2, Robust Multichip Average (RMA) normalised) during these early stages.

Transcript profiles from the three partitions (41, 51 and 75 transcripts respectively in each of the partitions) in the *k*-means cluster (where k = 20) were manually scrutinised based on three criteria: 1) greater than 2-fold expression change over PM-LP-DA-TT; 2) a high level of expression over the time-course, thus indicating an increased copy number and therefore easier isolation via PCR; and 3) transcript levels that were moderate during and beyond immature pollen compared to expression levels during the earlier stages of meiosis.

Transcripts chosen for further analysis based on their expression profile were clustered with *ASY1* (Boden *et al.*, 2007; Boden *et al.*, 2009) and *DMC1* (Chapter 3) via Hierarchical clustering using the Euclidean squared similarity metric and an average linkage method to examine if the novel candidates shared a similar expression profile to these previously reported meiotic genes.

2.2.3 – Primary sequence analysis

As highlighted in Figure 2.1, Basic Local Alignment Search Tool (BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 160.0) searches against the public domain databases were performed with the expressed sequence tags (ESTs) which corresponded to the transcripts of interest shown in Appendix 2.1. The wheat UniGene database (http://www.ncbi.nlm.nih.gov/sites/entrez?d_b=unigene) was also used when no sequence similarity was found with the BLASTx protein sequence search function.

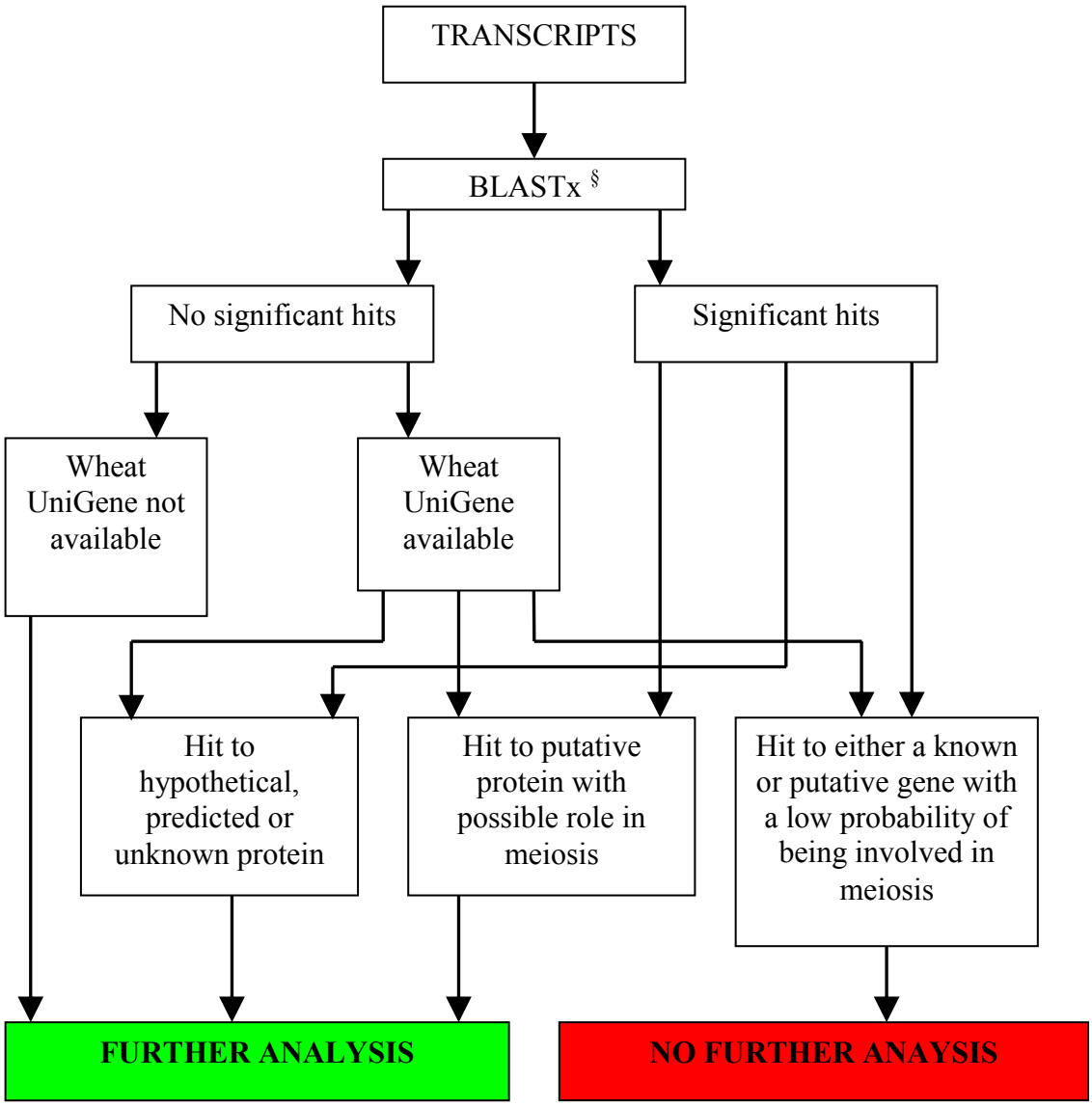


Figure 2.1 - Primary sequence similarity filtration of 39 transcripts. Using the selection criteria stated in the figure, sequence similarity searches using the NCBI database BLASTx function were conducted. § = BLASTx search with no organism restriction. E-value $\geq E^{-20}$ was used as the minimum threshold for sequence similarity.

2.2.4 – PCR amplification of candidates

Using the wheat cultivar Bob White 26, RNA from inflorescence undergoing meiosis was extracted using Trizol® Reagent (Invitrogen, Australia) as per the manufacturer‟s protocol. The extracted RNA was then used as a template for cDNA synthesis, using the iScript® Reverse Transcriptase cDNA synthesis kit (BioRad, Australia). This cDNA was subsequently used as a template for isolating the novel meiotic candidates.

Oligonucleotides used in the PCR described herein are listed in Appendix 2.2. Sequence amplification was conducted in a total volume of 20 µL using 0.5 µL of the cDNA template (from above), 2 µL of 10x PCR buffer (Invitrogen, Australia), 0.7 µL of 50 mM $MgCl_2$, 3.2 µL of 1.25 mM dNTPs, 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM), 0.2 µL of *Taq* Polymerase (Invitrogen, 5 U/µL) and 11.4 µL of sterile milli-Q water. PCR conditions were an initial denaturation for 7 minutes at 95°C, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 60°C, 30 seconds at 72°C, with a final extension at 72°C for 10 minutes. Post PCR reactions were analysed by gel electrophoresis (2% agarose) and visualised using ethidium bromide and a UV transilluminator (FirstLight™ UV illuminator, UVP, USA).

PCR products were isolated using the Invitrogen PureLink™ PCR purification kit as per the manufacturer‟s instructions; with the exception for Microarray Novels 1 and 22, where PCR products where extracted from the agarose gel using the Invitrogen PureLink™ quick gel extraction kit as per the manufacturer‟s instructions.

2.2.5 – Southern blot analysis to determine chromosome locations

Southern blot analysis using wheat nulli-tetra membranes were performed to determine the chromosome locations of the microarray novels within the wheat genome. Each nylon membrane contained digested (individually with one of the following enzymes: *Bam*HI, *Dra*I, *Eco*RI, *Eco*RV or *Xba*I) genomic DNA of Chinese Spring (CS) and 21 CS derivatives which are nullisomic for a particular chromosome and tetra compensated with a

homoeologous chromosome. For example, the absence of the two copies of chromosome 1A is compensated with the presence of another 2 copies of 1B chromosomes (courtesy of Margie Pallotta, Australian Centre for Plant Functional Genomics, Australia).

2.2.5.1 – Pre-hybridisation of nylon membranes

Both the pre-hybridisation and hybridisation solutions were made according to the following (per 100 mL): 5 mL of nanopure water, 5 mL of salmon sperm DNA (5 mg mL^{-1}), 30 mL of 5X HSB Buffer (3M NaCl, 100mM PIPES, 25mM Na_2EDTA, pH6.8), 30 mL Denhardt‟s III reagent (2% BSA, 2% Ficoll 400, 2% polyvinyl-pyrrolidone360, 10% sodium dodecyl sulfate (SDS), pre-heated to 65°C), and 30 mL of 25% Dextran Sulphate. All components of the solution were mixed together and heated at 65°C for 10 minutes. Membranes were individually placed into hybridisation bottles, to which 10 mL of the pre-hybridisation solution was added (initially). Membranes were then incubated at 65°C overnight with constant rotation in a hybridisation oven.

2.2.5.2 – Probe labelling and hybridisation

PCR products isolated in section 2.2.4 were used as a template for the synthesis of labelled probes. In individual reactions, 9 µL (approximately 150 ng) of each of the microarray PCR fragments were combined with 6 µL of 9-mer random oligo mix (10 µM); denatured at 100°C for 5 minutes, with immediate incubation on ice for 5 minutes. 25 µL of oligo-buffer, 2 µL Klenow polymerase (Roche, Australia) and 8 µL of radioactive 32^P (α-dCTP, 10 mCi mL-1, Perkin-Elmer, Australia) was added and mixed well. Synthesis reactions were incubated at 37°C for 60 minutes.

Synthesised probe was purified using the Invitrogen PureLink™ PCR purification kit as per manufacturer‟s instructions. Labelled probe was eluted using 60 µL of sterile milli-Q water and denatured at 100°C for 5 minutes, before incubating on ice for 5 minutes. Purified labelled probe was then added into the hybridisation bottle containing 10 mL of freshly added

hybridisation solution (section 2.2.5.1), and incubated at 65°C overnight with constant rotation in a hybridisation oven.

2.2.5.3 – Membrane washes

Membranes, contained within hybridisation bottles, were washed with consecutively increasing stringent washes of standard saline citrate (20x SSC; 3M NaCl, 0.3M Tri-sodium citrate) and sodium dodecyl sulfate (SDS) to remove excess labelled probe to decrease the amount of background. Hybridisation solution, containing excess probe, was initially decanted from the hybridisation bottle before adding 40 mL of 2x SSC with 0.1% SDS (w/v) and incubating at 65°C. Subsequent washes were also carried out for 20 minutes at 65°C with 40 mL of 1x SSC containing 0.1% SDS (w/v) until an appropriate level of radioactivity (1-2 counts per second) was recorded.

2.2.5.4 – Autoradiography

Hybridised membranes were sealed in plastic sleeves and placed into an X-ray cassette along with X-ray film (Kodak Biomax, Kodak, USA). Cassettes were stored at -80°C for 7 to 14 days depending upon the level of background radiation detected before sealing. X-ray films were developed using an AGFA CP1000 Developer (AGFA, Belgium).

2.2.6 – Secondary sequence analysis

Transcripts which were able to be mapped to wheat chromosomes were reanalysed for putative sequence homology, specifically to *Arabidopsis thaliana* (*At*) and *Oryza sativa* (*Os*, Rice), given the passing of time between the primary sequence analysis conducted previously (section 2.2.3). BLASTn and BLASTx (http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 166.0) analyses was performed with the corresponding ESTs for the 20 unique transcripts against the public domain databases with an *E*-value $\geq E^{-20}$ used as a significance threshold (as described in Figure 2.2). Transcripts which were found to have either no annotation, or an annotation that suggested that the transcript could be involved in

meiosis (based on the identification of a functional homologue in Arabidopsis) were then used for further characterisation. The availability of Arabidopsis mutant stocks for each of these candidates was crucial. As shown in Chapter 5, these mutants were used as a „proof of concept" in determining whether the gene of interest had a role during meiosis.

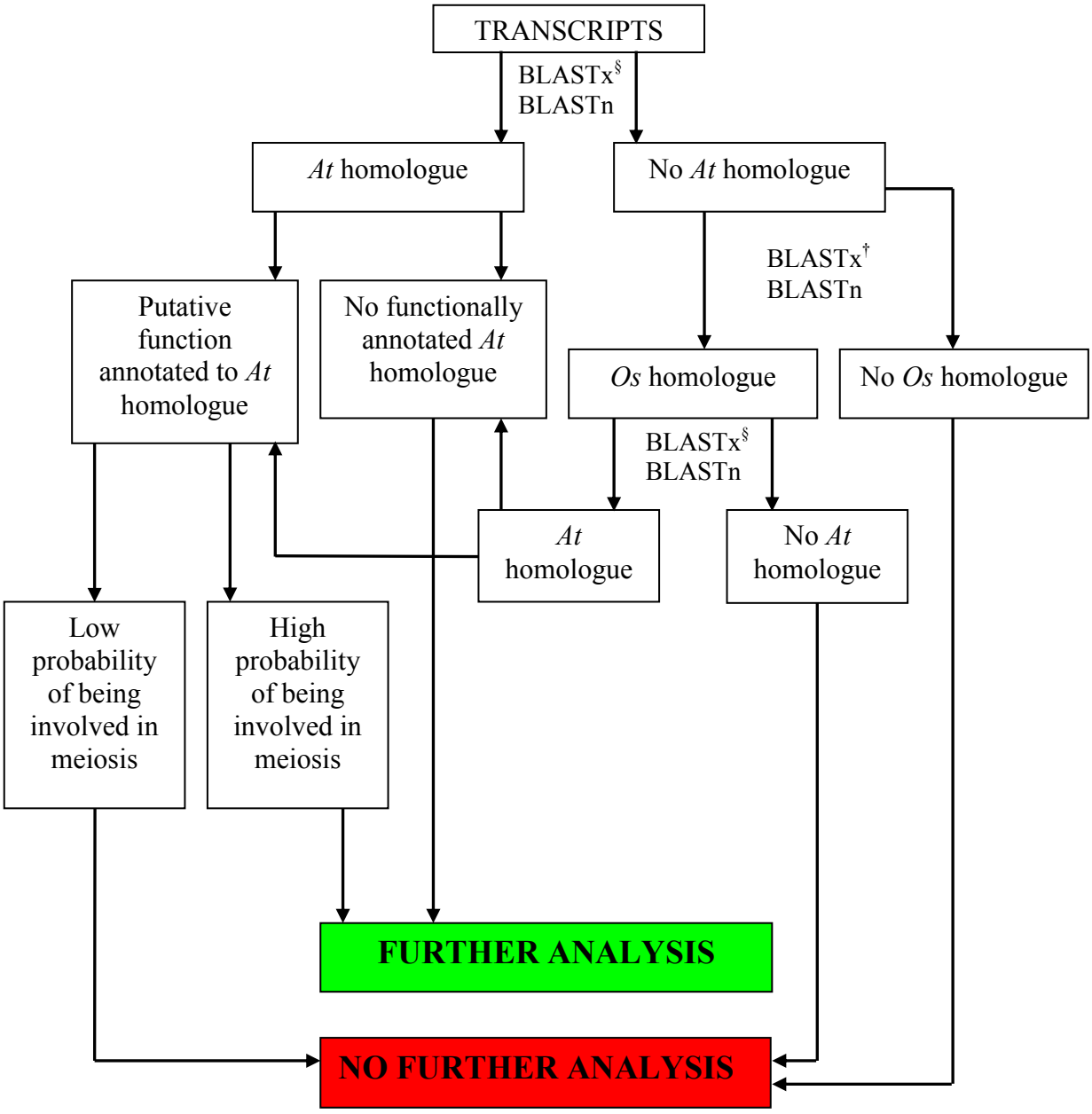


Figure 2.2 - Secondary sequence similarity filtration of 20 transcripts. Sequence similarity of the 20 unique transcripts, mapped to wheat chromosomes, was determined via the stated selection criteria, using NCBI database BLASTn and BLASTx functions. § = sequences restricted to *Arabidopsis thaliana* (*At*), † sequences restricted to *Oryza sativa* (*Os*). *E*-value $\geq E^{-20}$ used as the minimum threshold for sequence similarity.

2.3 – Results

2.3.1 – Comparative expression profiling

K-means clustering was used to determine similarity of transcript expression profiles of the 1094 non-functionally annotated genes found to be meiotically regulated (Crismani *et al.*, 2006). Using 10, 15, 20, 25 and 30 partitions, the clustering returned 3, 3, 4, 2 and 3 groups respectively; which showed higher expression occurring in the first three stages examined compared to the remaining four stages. Scrutinising these groups further, it was found that 2, 2, 3, 2 and 2 groups from 10, 15, 20, 25 and 30 partitions, respectively, displayed an average absolute expression level above 9 (log base 2, RMA normalised), over the first three stages. Whereas the remaining groups displayed only moderate expression levels (under 9) during the same stages (Appendix 2.3; yellow boxes versus the blue boxes highlighted).

The *k*-means clustering applying 20 partitions ($k = 20$) identified three transcript clusters with an early meiosis expression peak, which were followed by a downward trend in expression during the later stages of meiosis (Figure 2.3). These three transcript clusters contained a cumulative total of 167 transcripts. Based on the selection criteria (section 2.2.2), 39 of these transcripts were chosen for further analysis (examples of which are highlighted in Figure 2.4, with all transcript profiles shown in Appendix 2.4). Comparative expression profiling of these transcripts with the known meiotic genes of *ASY1* and *DMC1* showed that these genes have similar expression over the seven stages of meiosis and anther development (Figure 2.5). Comparing *ASY1* and *DMC1* against the starting 1094 transcripts revealed that 35 of the 39 candidates grouped with these known genes (correlation coefficient $r \geq 0.79$, Figure 2.6).

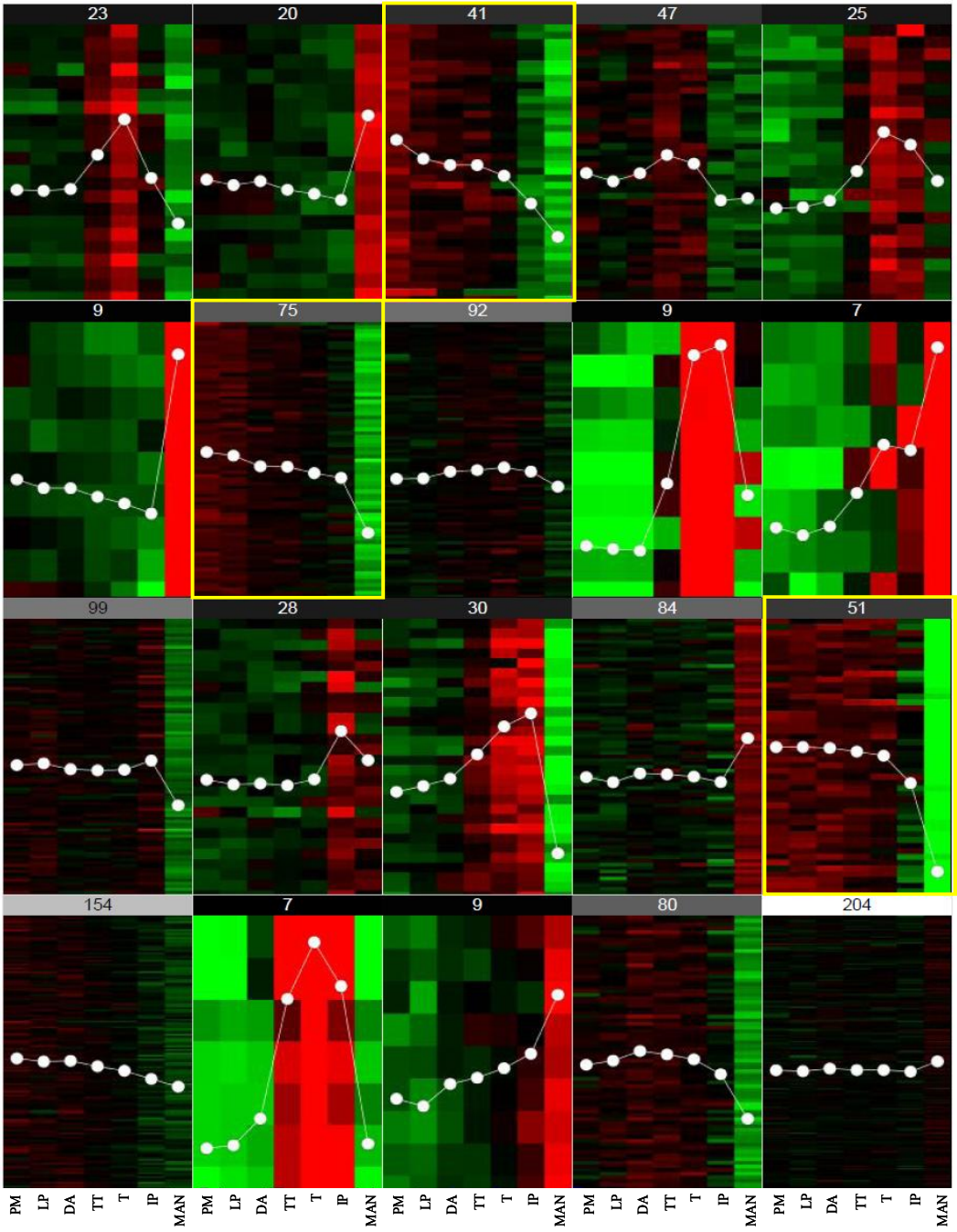


Figure 2.3 - K-means clustering of 1094 meiotically regulated novel wheat transcripts. Expression profiles of 1094 wheat transcripts were grouped via *k*-means clustering (k = 20) across seven anther development stages. Labels represent the number of transcripts within each of the clusters. The clusters of 41, 75 and 51 (boxed in *yellow*) were chosen for further analysis due to their similar expression profiles of interest, showing a higher level of expression during the earlier stages of meiosis (PM and LP) before a reduction of expression in later stages of anther development, when compared to the other clusters shown. Each column in each graph represents a stage that was tested in the microarray experiments conducted by Crismani *et al.*, (2006); column 1- pre-meiosis (PM), column 2 - leptotene to pachytene (LP), column 3 - diplotene to anaphase I (DA), column 4 - telophase I to telophase II (TT), column 5 – tetrads (T), column 6 - immature pollen (IP), and column 7 - mature anthers (MAN).

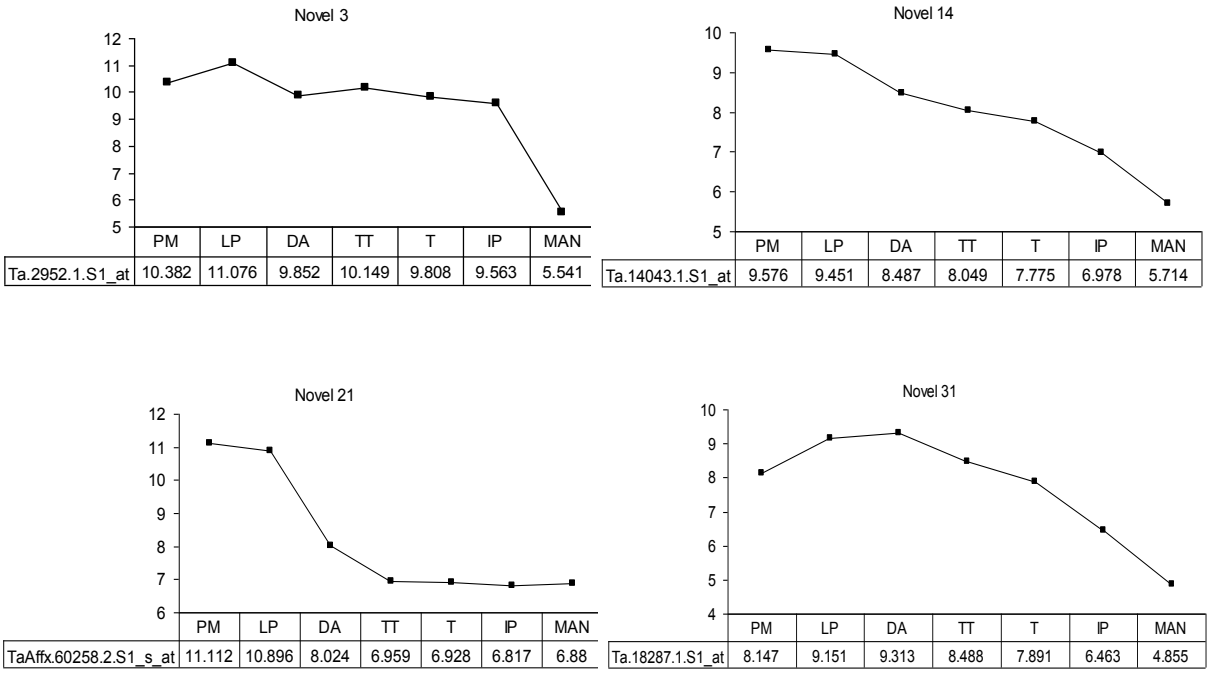
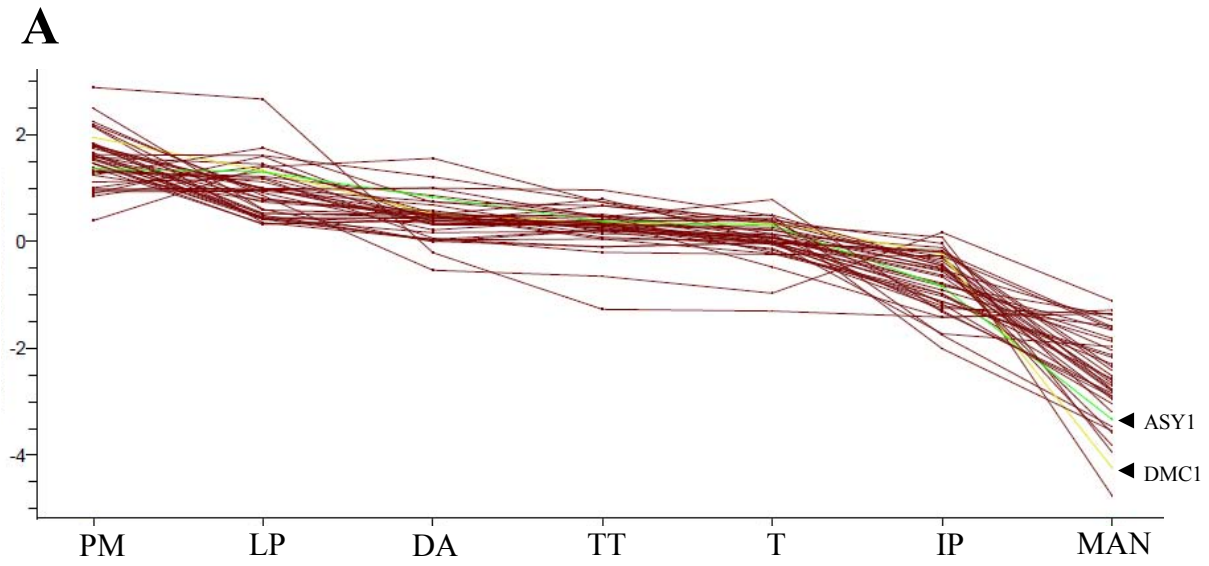


Figure 2.4 - Expression profile examples selected for further analysis. Expression values on the y axis are log base 2, RMA-normalised. Expression levels ≥ 10 are considered highly expressed, levels of 6-9 are considered moderately expressed and levels ≤ 5 are very low (basal level of expression). Stages of anther development are displayed on the x axis: pre-meiosis (PM), leptotene to pachytene (LP), diplotene to anaphase I (DA), telophase I to telophase II (TT), tetrads (T), immature pollen (IP) and mature anthers (MAN).



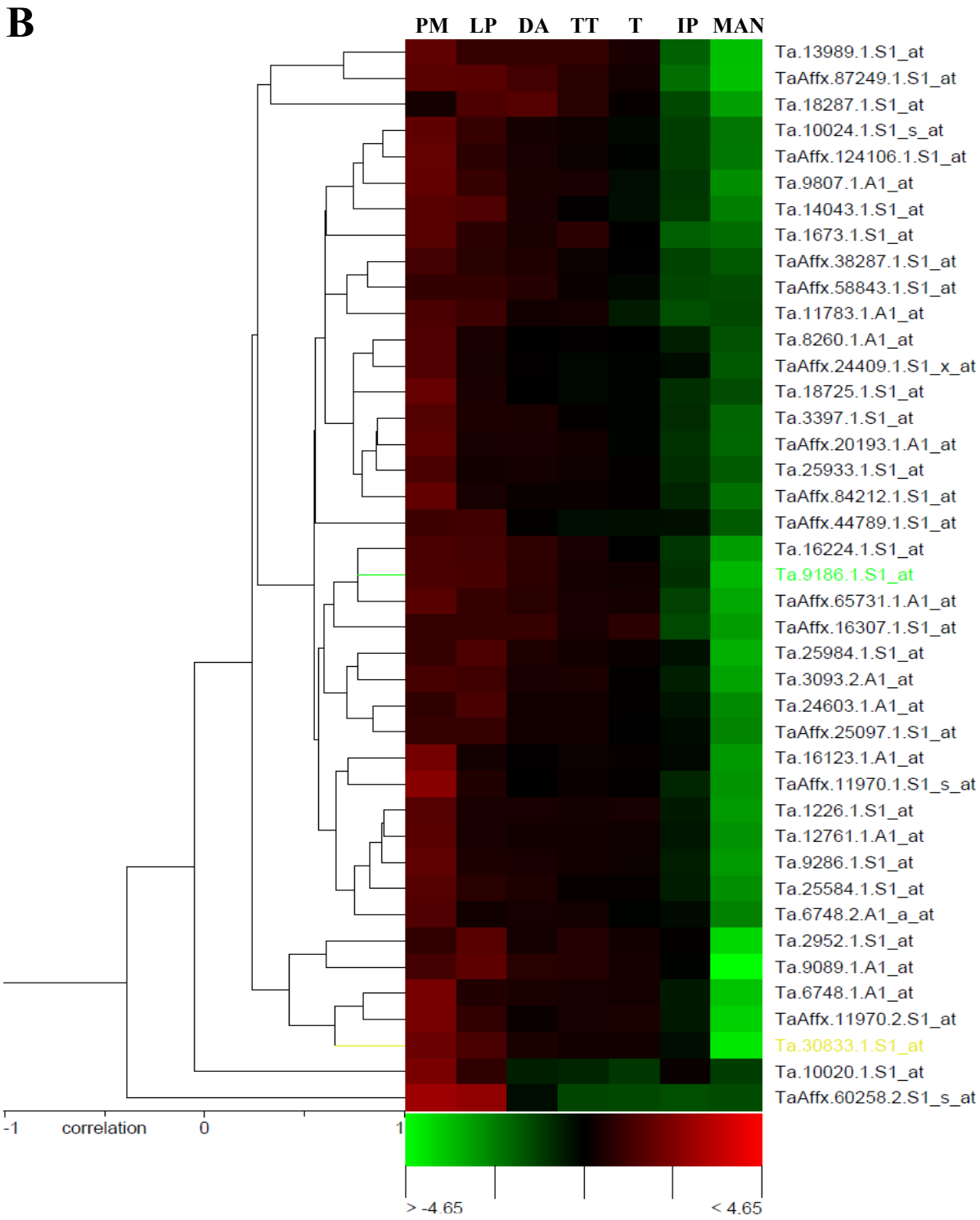
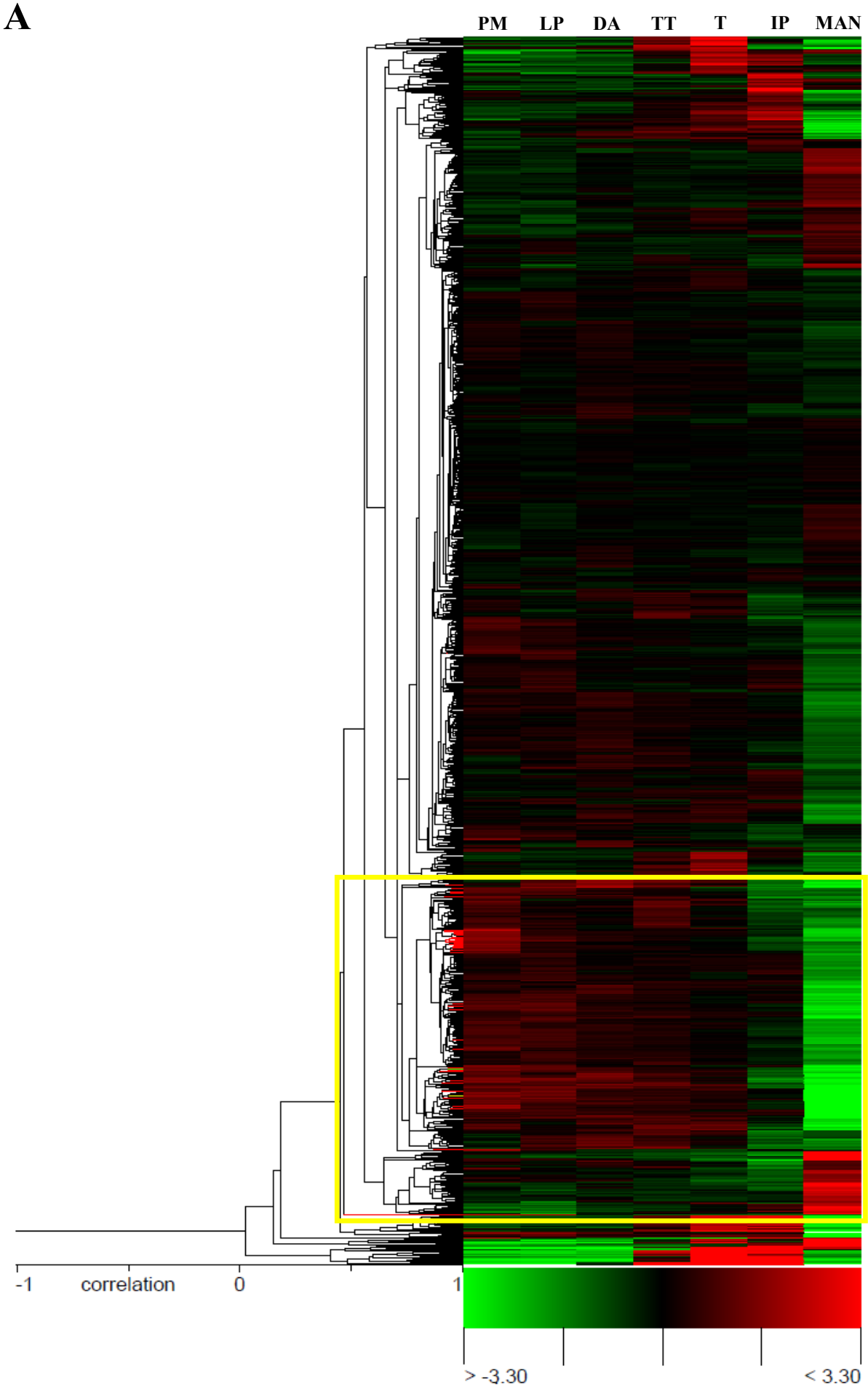


Figure 2.5 - Hierarchical clustering of novel transcripts with known meiotic genes *ASY1* and *DMC1*. Expression profiles of 39 novel transcripts chosen for further analysis were grouped via hierarchical clustering in graph form (**A**) and in heat map (**B**) across seven anther development stages (columns) and are compared to the known meiotic genes of *ASY1* (Ta.9186.1.S1_at, *green*) and *DMC1* (Ta.30833.1.S1_at, *yellow*). Pre-Meiosis (PM), leptotene – pachytene (LP), diplotene – anaphase I (DA), telophase I – telophase II (TT), tetrads (T), immature pollen (IP), mature anthers (MAN). Expression values in cluster graph (A) (y-axis) and heat map (B) (indicated by green through to red) are centred, log base 2, RMA-normalised values.


A
PM
LP
DA
TT
T
IP
MAN
-1
correlation
0
1
> -3.30
< 3.30

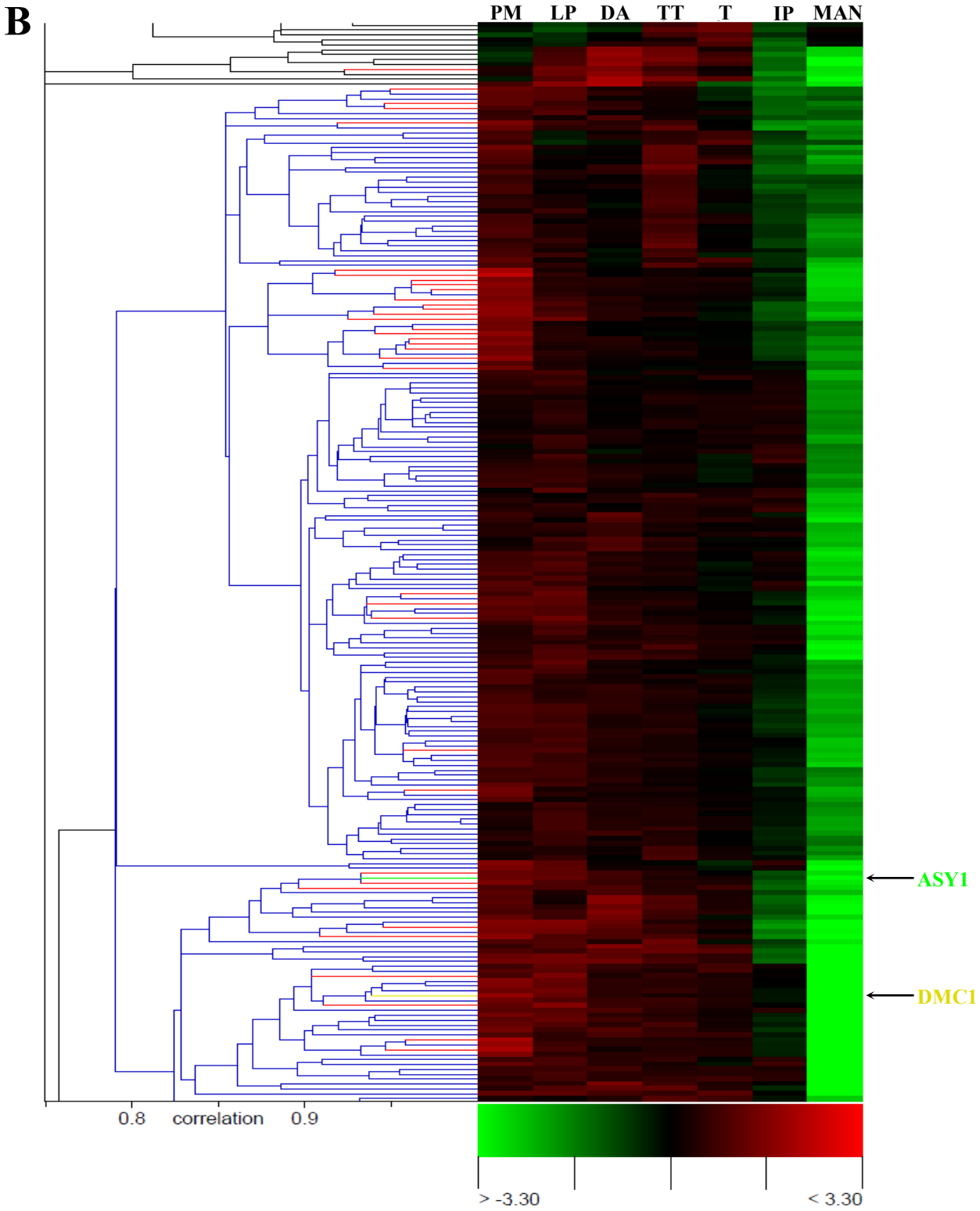


Figure 2.6 - Hierarchical clustering of 1094 novel transcripts with known meiotic genes *ASY1* and *DMC1*. The 39 transcripts (red lines on dendrogram) chosen for further analysis are shown to cluster (boxed area) with *ASY1* (*green*) and *DMC1* (*yellow*) (**A**). 35 of the 39 transcripts are observed to have a high correlation ($r \geq 0.79$) of expression with the known meiotic genes (**B**; which is a zoomed image of the boxed area shown in (**A**)). Pre-Meiosis (PM), leptotene – pachytene (LP), diplotene – anaphase I (DA), telophase I – telophase II (TT), tetrads (T), immature pollen (IP), mature anthers (MAN). Expression values (indicated by green through to red) are centred, log base 2, RMA-normalised values.

2.3.2 – Primary sequence analysis

Of the 39 transcripts selected for further investigation, BLAST sequence analysis identified eight transcripts, previously not annotated or having a predicted function, now having annotations or strong similarity to annotated genes in the public database with a low probability of these candidates being involved in meiosis (Table 2.1). The remaining 31 transcripts returned either no annotation, matches to hypothetical or predicted genes, or matched to predicted annotated genes which could have a high probability of being involved in meiosis. Ten of these transcripts returned matches to four separate annotations (Novels 2/10, 3/4/5/6, 11/17, 26/28); however all ten transcripts were advanced to the next stage due to the observation that there were varying expression levels detected (see Appendix 2.3).

2.3.3 – PCR amplification of selected candidates

Of the 31 transcripts chosen for further analysis after primary sequence analysis, 28 candidates were amplified using PCR and subsequently isolated from the agarose gels (Figure 2.7).

Table 2.1 - Annotations for transcripts of interest after primary sequence searching. BLASTx (§ = using no restriction on organism parameters) analysis was conducted and the wheat UniGene database (†) searched for sequence similarity of the 39 transcripts chosen for further analysis. Transcripts highlighted in *green* were taken for further analysis, while those in *red* were not. *Os* = *Oryza sativa*, *At* = *Arabidopsis thaliana*, *Zm* = *Zea mays*, *Ta* = *Triticum aestivum*.

Novel	Transcript ID	Annotated function
1	Ta.1226.1.S1_at	Ta.1226 Transcribed locus †
2	TaAffx.11970.1.S1_s_at	EAZ17232 *Os* Hypothetical protein §
3	Ta.2952.1.S1_at	Ta.2952 Hypothetical protein †
4	Ta.9089.1.A1_at	Ta.2952 Hypothetical protein †
5	Ta.24603.1.A1_at	Ta.2952 Hypothetical protein †
6	Ta.25984.1.S1_at	Ta.2952 Hypothetical protein†
7	Ta.3397.1.S1_at	—
8	Ta.10020.1.S1_at	Ta.27933 Transcribed locus †
9	Ta.10024.1.S1_s_at	Ta.10024 Putative homeotic protein †
10	TaAffx.11970.2.S1_at	EAZ17232 *Os* Hypothetical protein §
11	Ta.8260.1.A1_at	Ta.8260 Expressed protein †
12	Ta.13989.1.S1_at	—
13	TaAffx.16307.1.S1_at	—
14	Ta.14043.1.S1_at	—
15	TaAffx.87249.1.S1_at	—
16	TaAffx.58843.1.S1_at	NP_001050917 *Os* Hypothetical protein §
17	Ta.18725.1.S1_at	Ta.8260 Expressed protein †
18	TaAffx.25097.1.S1_at	—
19	TaAffx.24409.1.S1_x_at	Ta.6670 Hypothetical protein †
20	TaAffx.20193.1.A1_at	Ta.33981 Hypothetical protein †
21	TaAffx.60258.2.S1_s_at	—
22	TaAffx.38287.1.S1_at	Ta.45808 Transcribed locus †
23	Ta.12761.1.A1_at	Ta.12761 Transcribed locus †
24	TaAffx.84212.1.S1_at	—
25	Ta.25933.1.S1_at	Ta.25933 Hypothetical protein †
26	Ta.6748.2.A1_a_at	EAY93787 *Os* Hypothetical protein §
27	Ta.16123.1.A1_at	Ta.16123 Kinesin-related protein-like †
28	Ta.6748.1.A1_at	EAY93787 *Os* Hypothetical protein §
29	TaAffx.44789.1.S1_at	EAZ44378 *Os* Hypothetical protein §
30	Ta.16224.1.S1_at	Ta.16224 Methyltransferase-like protein †
31	Ta.18287.1.S1_at	Ta.18287 Transcribed locus †
32	Ta.9286.1.S1_at	NP_195066 *At* DEM Protein §
33	TaAffx.65731.1.A1_at	AAT78760 *Os* putative transport protein §
34	Ta.1673.1.S1_at	Ta.28849 Ubiquitin-conjugating enzyme †
35	Ta.11783.1.A1_at	Ta.11783 Putative Polygalacturonase †
36	Ta.3093.2.A1_at	Ta.3093 RNA helicase †
37	Ta255.84.1.S1_at	AAK98710 *Os* Putative ribnuclotide redutase §
38	TaAffx.124106.1.S1_at	AAL17718 *At* α-glucosidase I §
39	Ta.9807.1.A1_at	AAY45706 *Zm* RNA Polymerase IV 2nd largest subunit §

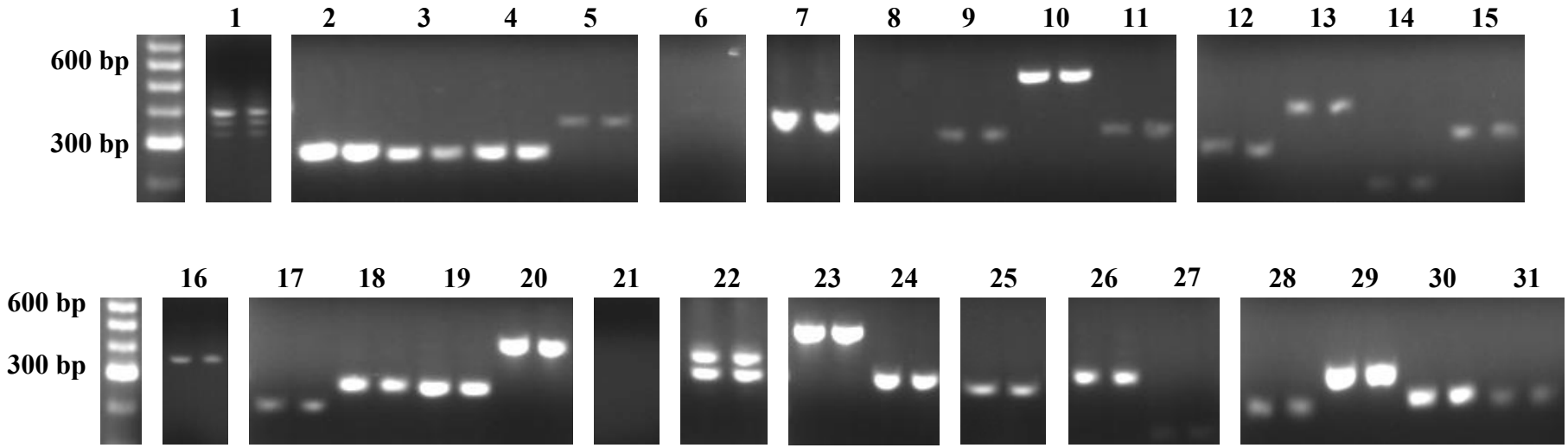
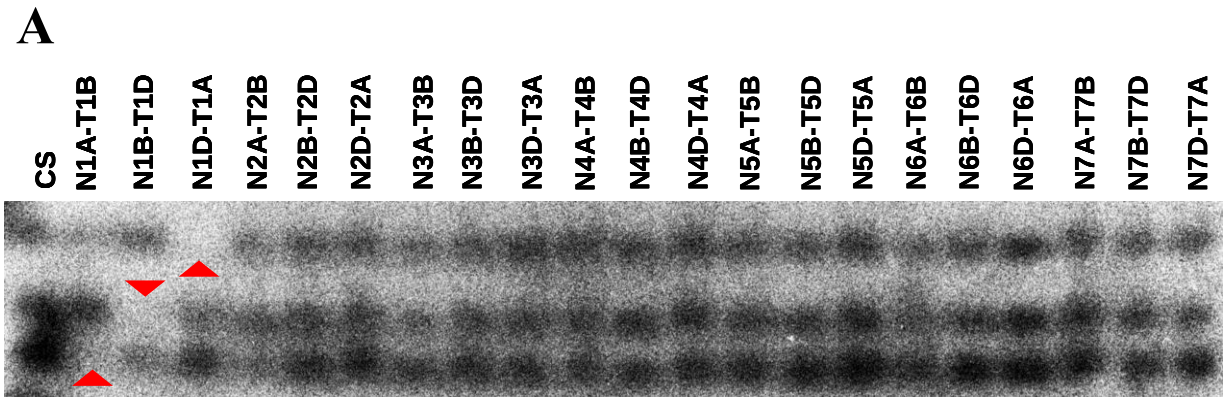


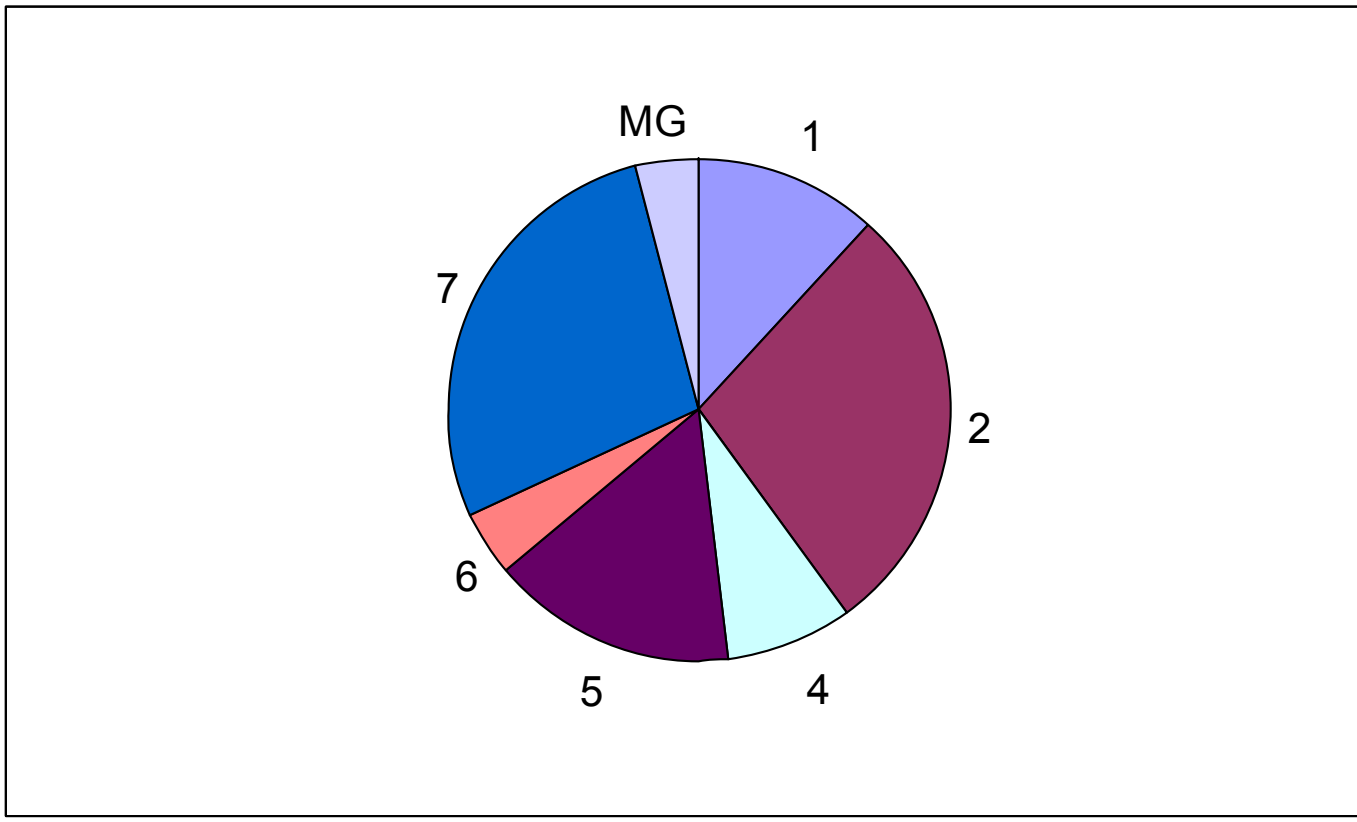
Figure 2.7 - Amplification of novel transcripts identified from the wheat microarray analysis. PCR products were analysed using 3µL that was loaded in a 2% agarose gel, and subsequently visualised using ethidium bromide and a UV transilluminator. Numerals above the gel photos correspond to the novel transcripts. Novels 1 and 22, where lanes contained multiple bands, were extracted separately straight from the gel; with the products being termed a, b and c (novel 1) and a, b (novel 22) from the highest band to the lowest band. A 1 kb DNA molecular weight marker is shown; however not all band sizes are listed.

2.3.4 – Chromosome location of novel transcripts

Of the 28 transcripts isolated via PCR amplification, 25 candidates were mapped to wheat chromosomes (Figure 2.8 A; Appendix 2.5). Over half (56%) of these candidates were mapped to either groups 2 or 7; 20% on group 5; with another 20% distributed across groups 1, 4 and 6. The remaining candidate was located on multiple chromosome groups (MG, Figure 2.8 B). Interestingly, none of the 25 meiotically regulated candidates mapped successfully were located on chromosome group 3.

Figure 2.8 - Wheat chromosome location for novel microarray candidates. An example of Southern autoradiography (**A**, and Appendix 2.5) shows the absence of a hybridisation signal from chromosome group 1 (red arrow heads). Lanes contain DNA of Chinese Spring and 21 CS-derivatives (nulli-tetra lines) (see section 2.2.5 for details). An absence of a hybridisation signal in a lane represents the loss of the target signal, and thus the chromosome location of the novel transcript can be determined. (**B**) Pie graph indicates chromosomal locations of the meiotically regulated candidates, which were found to be predominantly located on chromosome groups 2, 5 and 7; while no candidates were found to be located on chromosome group 3. Labels indicate wheat chromosome groups 1 to 7 and multiple groups (MG).





Novels 2 and 10; 3, 4 and 5; 11 and 17; as well as 26 and 28 which had differential expression (Appendix 2.4), yet matched to the same annotations (Table 2.1) were found to be located on the same chromosome groups (for example novel 2 and 10 both mapped to chromosome 4). Based on this result, only one of the matching candidates was taken for further analysis to the next stage. Multiple PCR products for a given candidate, for example novel 1 a, b and c, were also located on the same chromosome group and were therefore treated as being identical (at least genetically). All unique transcripts were then used for homologue sequence analysis searching, despite transcripts predominately being located to three chromosome groups (Table 2.2).

Table 2.2 - Chromosome locations of novel candidates. ‡, *, §, † indicate those transcripts with matching annotations. *Green* and *red* represent further analysis and no further analysis, respectively for each of the candidates listed.

Novel	Transcript ID	Location
1a	Ta.1226.1.S1_at	5
1b		5
1c		5
2‡	TaAffx.11970.1.S1_s_at	4
3*	Ta.2952.1.S1_at	7
4*	Ta.9089.1.A1_at	Not determined
5*	Ta.24603.1.A1_at	7
7	Ta.3397.1.S1_at	1
9	Ta.10024.1.S1_s_at	1
10‡	TaAffx.11970.2.S1_at	4
11§	Ta.8260.1.A1_at	2
12	Ta.13989.1.S1_at	Not determined
13	TaAffx.16307.1.S1_at	5
14	Ta.14043.1.S1_at	7
15	TaAffx.87249.1.S1_at	6
16	TaAffx.58843.1.S1_at	Multiple groups
17§	Ta.18725.1.S1_at	2
18	TaAffx.25097.1.S1_at	2
19	TaAffx.24409.1.S1_x_at	2
20	TaAffx.20193.1.A1_at	5
22a	TaAffx.38287.1.S1_at	1
22b		1
23	Ta.12761.1.A1_at	2
24	TaAffx.84212.1.S1_at	7
25	Ta.25933.1.S1_at	2
26†	Ta.6748.2.A1_a_at	7
27	Ta.16123.1.A1_at	Not determined
28†	Ta.6748.1.A1_at	7
29	TaAffx.44789.1.S1_at	5
30	Ta.16224.1.S1_at	7
31	Ta.18287.1.S1_at	2

2.3.5 – Homologue sequence analysis

BLASTx sequence searching for Arabidopsis homologue of the transcripts of interest resulted in identifying eight for further analysis, where an Arabidopsis mutant was available (Table 2.3). Analysis of the T-DNA knock-outs for these Arabidopsis loci will be reported in Chapter 5 to determine if these candidates play a role during meiosis.

Table 2.3 - Arabidopsis homologues of the meiotically-expressed transcripts of interest. Of the 20 unique transcripts mapped to wheat chromosomes, eight were found to have homologue in Arabidopsis which have previously not been characterised. These eight candidates (*green*) are to be taken for further analysis using publically available Arabidopsis T-DNA mutant stocks, while those without a homologue (*red*) were discounted from further analysis.

Novel	Transcript ID	Annotation	*At* Homologue
1	Ta.1226.1.S1_at	—	—
2	TaAffx.11970.1.S1_s_at	Putative DUF936 domain protein	At4g13370
3	Ta.2952.1.S1_at	—	—
7	Ta.3397.1.S1_at	Putative DNA Topoisomerase II	At3g23890
9	Ta.10024.1.S1_s_at	Putative Homeobox-7 Transcription factor	At5g46880
11	Ta.8260.1.A1_at	Unknown protein	At4g02800
13	TaAffx.16307.1.S1_at	—	—
14	Ta.14043.1.S1_at	Putative Gibberellin-responsive protein	At1g74670
15	TaAffx.87249.1.S1_at	60S acidic ribosomal protein P2	At3g44590
18	TaAffx.25097.1.S1_at	Unknown protein	At3g56870
19	TaAffx.20193.1.A1_at	—	—
20	TaAffx.20193.1.A1_at	—	—
22	TaAffx.38287.1.S1_at	—	—
23	Ta.12761.1.A1_at	—	—
24	TaAffx.84212.1.S1_at	Chromosome condensation protein-like	At5g37630
25	Ta.25933.1.S1_at	Condensin subunit (chromosome condensation)	At3g57060
26	Ta.6748.2.A1_a_at	Putative Kinesin Motor protein	At5g02370
29	TaAffx.44789.1.S1_at	Putative HAT B	At5g56740
30	Ta.16224.1.S1_at	—	—
31	Ta.18287.1.S1_at	—	—

2.4 – Discussion

2.4.1 – Data filtration

From microarray analysis previously conducted by Crismani and colleagues (2006), 1350 transcripts were found to be meiotically regulated over the seven stages of meiosis and anther development investigated. At the time of this study commencing, 1094 of these had either no annotation or a purely predicted annotated function. This dataset is exceedingly large, and it is unrealistic to conduct in-depth functional analysis on each of these candidates. Furthermore, it does not take into consideration that some of these candidates may not have a role that is meiosis specific, given that the whole anthers used in the microarray experiment were a combination of both mitotic and meiotic tissues. The major aim of this research

chapter was to therefore reduce the number of candidates significantly through a series of filtration stages (Figure 2.9) which were designed to enrich for meiotically-regulated transcripts that would then be used for functional analysis (see Chapter 5).

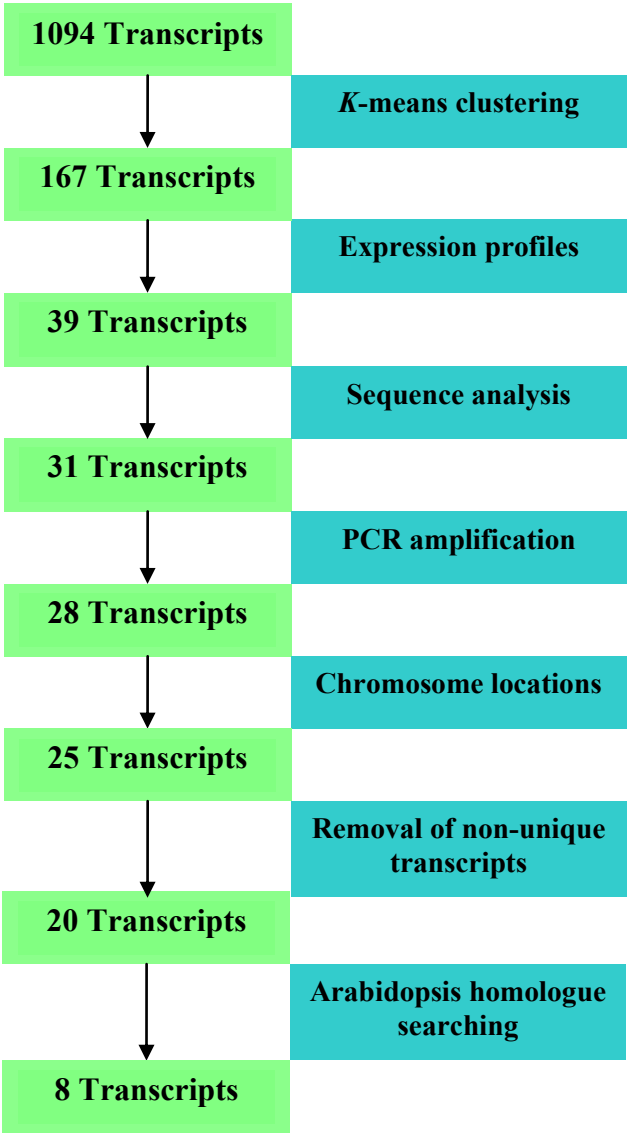


Figure 2.9 - Data filtration of novel candidates with meiotic expression. Using consecutive rounds of selection filters; the number of possible candidates for in-depth analysis was decreased from 1094 to eight.

Microarray analyses of biological processes previously reported have also found a large number of candidates of interest, and therefore have had to conduct data filtration to reach an achievable number for further analysis (Chu *et al.*, 1998; Tanaka *et al.*, 2002; Lo *et al.*, 2003; Ushizawa *et al.*, 2004). Several have filtered candidates by examining expression changes followed by expression profile clustering, often including genes known to have a role in the process; as genes with similar functions display similar expression patterns (Kaminski and Friedman, 2002).

Yeast meiotic microarray analysis previously conducted had the starting number of roughly 500 transcripts which were found to be induced during the sporulation process (Chu *et al.*, 1998). Chu and colleagues (1998) clustered these induced genes with known meiotic genes and found seven distinct expression profiles. Deletion mutants of selected novels across the seven groups exhibited similar phenotypes to those previously observed for the known meiotic genes from each of the clusters. Research in zebrafish (*Danio rerio*) (Lo *et al.*, 2003), mouse (*Mus musculus*) (Tanaka *et al.*, 2002) and bovine (*Bos taurus*) (Ushizawa *et al.*, 2004) also started with large transcript numbers regulated over their process of interest, with these studies all following a similar methodology for candidate filtration.

In the current wheat meiosis study, *k*-means clustering was conducted to isolate transcripts which displayed the expression profile of most interest to our research group; an increased level of expression during the early stages of meiosis followed by a decrease in transcript copy in the later stages of meiosis and anther development. It is during the earlier stages of meiosis that processes such as chromosome pairing, synapsis and recombination occur. Genes known to be involved in these processes exhibit this general profile and therefore novel candidates were chosen on the basis of whether a transcript followed this general trend across the seven stages examined. The novel transcripts chosen for further filtering were compared to the known meiotic genes of *ASY1* (Boden *et al.*, 2007; Boden *et al.*, 2009) and *DMC1* (Chapter 3); and as Figure 2.6 demonstrates, 89% of the transcripts chosen for further analysis showed high correlation (correlation coefficient r $\geq$0.79) to the

expression profiles of these known genes. The 39 transcripts chosen for further analysis were used to interrogate the NCBI database, resulting in 31 transcripts for further analysis. Given the frequency at which such public databases are updated, this small reduction in transcript number was expected.

Investigating transcripts which demonstrate meiotic expression and regulation with other known meiotic genes can increase the chance of discovering a novel gene which has a role in meiosis, whether it is essential or not (Eisen *et al.*, 1998). Through studying candidates that have early meiotic expression in wheat and which are similar to known genes in meiosis, it does appear that the candidates for further investigation are being enriched for meiotic genes. Of the eight candidates identified in this chapter, four match to predicted proteins with roles that can be linked to the meiotic pathway, including chromosome condensation. Even so, while enriching the dataset for meiotic genes, it is not probable that all the transcripts within the dataset will be involved in meiosis. Indeed, it has been shown that genes with meiotic regulation do not necessarily have an essential role during meiosis (Kaback and Feldberg, 1985; Primig *et al.*, 2000; Rabitsch *et al.*, 2001). Genes involved in DNA replication and metabolic pathways have also been shown to have meiotic expression and regulation in yeast (Chu *et al.*, 1998) and *C. elegans* (Reinke *et al.,* 2000).

2.4.2 – Do wheat chromosomes contain clusters of meiotic genes?

Within the last 20 years several studies have reported that genes involved in the same biochemical pathway generally display a higher degree of gene clustering than by chance; with genes being located on the same chromosome(s) and in some cases in clusters of small defined areas of a given chromosome (Lee and Sonnhammer, 1993; Hurst *et al.*, 2004; Teichmann and Veitia, 2004). Of interest in this study is the finding of candidates being predominantly located to chromosome groups 2, 5 and 7 (76% combined).

Historically it is known that a number of loci which control chromosome pairing in bread wheat are located on chromosome groups 3 and 5 (Riley and Chapman, 1958; Sears,

1977; Sears, 1982). While four candidates where found to be located on group 5; none were located on chromosome group 3. The absence from chromosome group 3 could possibly be due to the small number of transcripts mapped to the wheat chromosomes in this study; as by chance only three to four would expect to be mapped to group 3 if they were randomly located throughout the three genomes. One must also take into consideration a potential source of bias in that only one type of expression profile was chosen for analysis (highest expression in the first three stages compared to the later stages). It could be that if there were meiotic genes located on chromosome groups 3; that they might not follow this expression profile and would therefore go undetected. Another possible reason for not locating a transcript which maps to chromosome group 3 could be that the Affymetrix wheat gene chip, which at the time of production was predicted to be representative of approximately 60% of the wheat genome, was under-represented for transcripts from chromosome group 3. Recent efforts have gone into wheat sequencing, especially chromosome group 3 (Paux *et al.*, 2008) and therefore production of an updated wheat chip may lead to the identification of group 3 transcripts that are expressed in meiotic tissues.

Recently a number of meiotic genes have been characterised in wheat, with the majority of these on chromosome groups 2, 5 and 7. The known meiotic genes of *RAD51A*, *B* and *D* are located on group 7; *RAD51C* and *ZYP1* are located on group 2 (Khoo *et al.*, 2008 and Khoo *et al.*, *unpublished data*); while *ASY1* (Boden *et al.*, 2007; Boden *et al.*, 2009) and *DMC1* (Chapter 3) reside on chromosome group 5. Could this clustering of both known meiotic genes and novel genes identified here provide some evidence for the novel candidates being involved in meiosis? Supporting this is the observation that three of the candidates located on chromosome groups 2/5/7 are found to match to putative chromosome condensation proteins (groups 2 and 7) and the regulatory histone acetytransferase B protein (group 5), all of which could play important roles during meiosis.

Despite the majority of the novel candidates being located on only three chromosome groups, perhaps without providing evidence of the roles for the transcripts (for both the

predicted and unknown proteins) one cannot say that wheat meiotic genes are clustering on a particular chromosome. One transcript, mapped to chromosome group 7, was found to match to a putative Gibberellin-responsive protein and therefore deemed not likely to have a role during meiosis; as gibberellin-responsive genes with meiotic regulation have only been implicated in pollen hydration, germination and tube growth to date (Wang *et al.*, 2005). A small number of promising candidates do not locate to the predominant chromosome locations of the transcripts; for example, the putative homoeobox transcription factor 7 is located on chromosome group 1. It is also unknown whether the candidates which preferentially map to the three chromosome groups are spread randomly over the whole chromosome or whether they are clustered within Centimorgans (cM) of one another. If transcripts were found to be located within a small distance of each other, this would provide further evidence towards joint functionality, as several studies have found that related genes group together in almost operon-like fashion (Lee and Sonnhammer, 1993; Blumenthal and Spieth, 1996; Teichmann and Veitia, 2004). To determine if the transcripts were clustered together a number of techniques could be conducted. Through the use of comparative genomics it could be established if these particular transcripts are clustered on the given chromosomes in rice (Sutton *et al.*, 2003), however the synteny despite being good on a marco-level tends to break down at the micro-level, with rearrangements of the gene order and content (Sorrells *et al.*, 2003). A more time consuming but accurate method to determine whether there are clusters of meiotic genes on individual chromosomes of bread wheat is to interrogate bacterial artificial chromosome (BAC) libraries with probes for each of the transcripts of interest. Identifying the same BACs using multiple probes would suggest that clusters exist, with sequencing of the individual BACs determining the distance between the candidates.

2.4.3 – Using the model species of *Arabidopsis thaliana* for functional analysis of wheat genes

To determine whether a novel transcript has a role during meiosis, looking at your organism of interest with the gene of interest either knocked-down or knocked-out provides a definitive answer to this question. In yeast, when essential meiotic genes are not expressed, spore formation will not occur (Primig *et al.*, 2000; Rabitsch *et al.*, 2001); whereas in plants, the process of meiosis continues, despite the absence or the reduction of the essential meiotic gene, and generally results in decreased fertility (Grelon *et al.*, 2001; Lloyd *et al.*, 2007). Although this is a generalisation, studying meiosis in plant species enables a more comprehensive dissection of where any particular gene may exert its effect. With the visualisation of DNA with DAPI, one can examine where meiotic defects occur; whether it is early meiosis during zygotene as with *MND1* (Kerzendorfer *et al.*, 2006) or at later stages (post-meiosis) as is the case with *TDM1* (Ross *et al.*, 1997).

For studying the candidates identified here, and their absence during meiosis, the first choice of organism to investigate is wheat (for example, the investigation of *TaASY1* by Boden and colleagues [2009]). However, production of transgenic wheat for gene function is both time consuming and expensive, especially when taking into consideration that there is no guarantee that the candidates selected for further analysis will play a role in meiosis. The alternative of looking at the candidates in wheat would be to look at the effect during meiosis in the model species *Arabidopsis thaliana*. Arabidopsis has been considered by the plant community as a valuable model species, helped by its relatively small genome size (140 Mb) which has been fully sequenced (Arabidopsis Genome Initiative, 2000; Haas *et al.*, 2005). This plant species has also become highly valuable from a research perspective through the publically available mutant stocks with transfer DNA (T-DNA) which has been inserted into your gene of interest. An insertion can be anywhere within the promoter or the gene itself, and

therefore the gene of interest can have a choice of either being knocked out (T-DNA in gene) or just knocked-down (in promoter) (Krysan *et al.*, 1999).

These types of mutants have been extensively used for the dissection of biological processes of interest, including meiosis, and can be used via either a forward or reverse genetic approach (Caryl *et al.*, 2003). In the forward genetic approach, a mutagenised population is produced and screened for the phenotype of interest; in the case of meiosis, often defects in the proper production and completion of meiotic cells and the meiotic process (Motamayor *et al.*, 2000). The alternative approach of reverse genetics is to generate or identify mutations in known genes which either show sequence similarity to meiotic genes from other species or have expression patterns suggestive of a meiotic role (Caryl *et al.*, 2003). An example of this is the known meiotic gene of *ASY1* which was found to be a homologue of the yeast gene *HOP1*. Whether using a forward or reverse genetic approach the Arabidopsis mutant stocks provide a valuable resource, as they yield a simple and direct route for exploring functionality of your particular gene of interest (Krysan *et al.*, 1999). Significantly, and of importance for the research reported here, none of the eight candidates identified in this Chapter have been studied during meiosis in Arabidopsis. This is therefore a unique opportunity to investigate a subset of genes that may provide additional data on the mechanics of meiosis in plant species.

Chapter 3

Chapter 3 – Isolation and functional characterisation of the meiotic genes *TaDMC1*, *TaHOP2* and *TaMND1* from bread wheat

3.1 – Introduction

Through information that has resulted from genome sequencing projects and the ability for the manipulation of these genomes, many genes that have been classically shown to be essential for meiosis in yeast are now being further characterised in higher eukaryotes (Cohen *et al.*, 2004; Hamant *et al.*, 2006; Muyt *et al.*, 2009). However the majority of this research has been limited to the model diploids, such as mice and Arabidopsis (Cohen and Pollard, 2001; Armstrong and Jones, 2003). As more than 70% of all flowering plants are polyploid (Masterson, 1994; Bowers *et al.*, 2003) using model diploids could result in over simplifying meiosis in these complex species (Martínez-Pérez *et al.*, 1999). Consequently, understanding how characterised genes isolated from diploids behave in complex genetic backgrounds is important.

One particular gene of interest is the recombinase protein DMC1 (*D*isrupted *M*eiotic *c*DNA 1), which in Chapter 2 was found to be regulated during wheat meiosis. DMC1 was previously discovered and characterised in yeast (Bishop *et al.*, 1992), and through later studies was found to have associated roles with the HOP2:MND1 protein complex (Petukhova *et al.*, 2005; Enomoto *et al.*, 2006; Pezza *et al.*, 2006). All three genes were originally characterised in yeast mutant screens with observable effects in the meiotic processes of pairing, recombination and synapsis (Bishop *et al.*, 1992; Leu *et al.*, 1998; Tsubouchi and Roeder, 2002). Additional studies have since shown that these three genes are vital for meiosis across a number of diverse species, including the *Basidiomyctes*, humans, mice, Arabidopsis and rice (Li *et al.*, 1997; Doutriaux *et al.*, 1998; Nara *et al.*, 1999; Ding *et*

al., 2001; Schommer *et al.*, 2003; Petukhova *et al.*, 2005; Kerzendorfer *et al.*, 2006; Panoli *et al.*, 2006).

Expression analysis in yeast has shown that *DMC1*, *HOP2* and *MND1* are specifically expressed during meiosis (see Figure 3.1A); while in plants all three genes have both vegetative and meiotic expression. However, highest expression is recorded in meiotic tissues; except for *AtHOP2* where the highest expression is seen in seedlings (Figure 3.1 B) (Bishop *et al.*, 1992; Ding *et al.*, 2001; Schommer *et al.*, 2003; Kerzendorfer *et al.*, 2006). These proteins have all been found to localise to chromosomes during meiosis (Figure 3.1 C and D), with protein loading of either HOP2 or MND1 dependent of one another (Tarsounas *et al.*, 1999 ; Tsubouchi and Roeder, 2002).

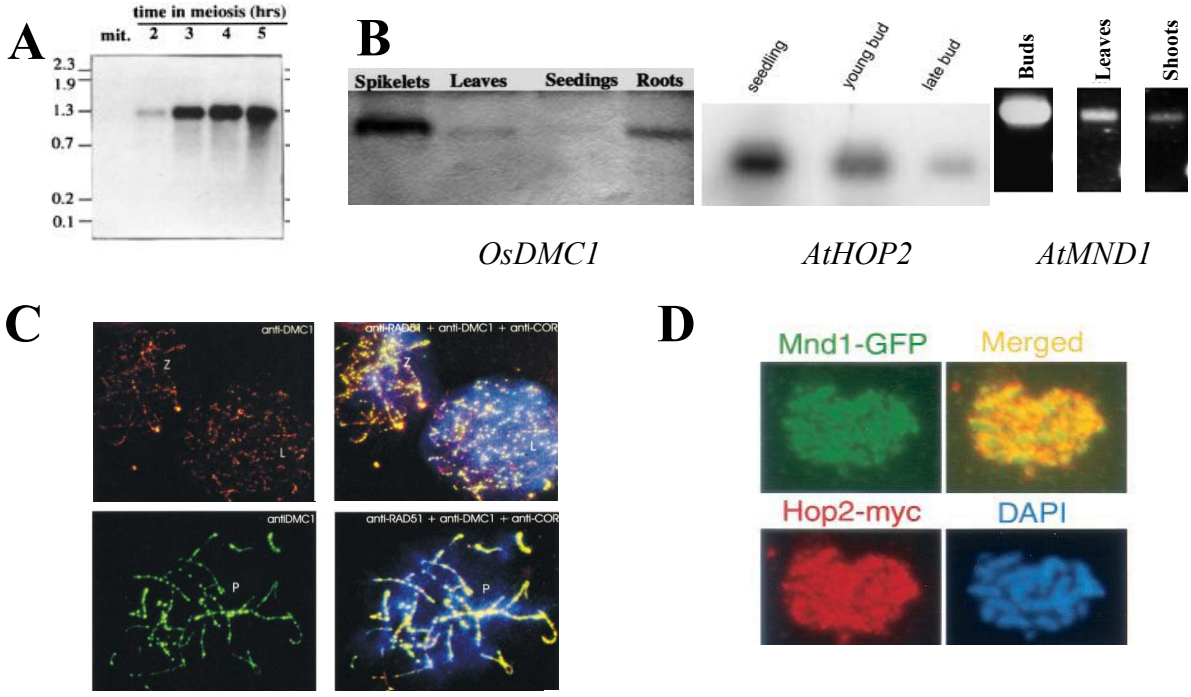


Figure 3.1 - Expression and protein localisation of DMC1, HOP2 and MND1 in diverse species. (**A**) *DMC1* transcript expression in *Saccharomyces cerevisiae* (*Sc*) mitotic cells (mit) and in sporulating cells at various time-points after the initiation of meiosis. (**B**) RT-PCR analysis of *DMC1* expression from rice (*Oryza sativa, Os*) and *HOP2*, *MND1* from *Arabidopsis thaliana* (*At*). (**C**) DMC1 protein localisation in mouse during leptotene (L) and zygotene (Z) (*red*, top left) and pachytene (P) (*green*, bottom left) and with DMC1 co-localisation with RAD51 and COR1 and DAPI (*blue*) (right). (**D**) Localisation of *Sc*HOP2 and *Sc*MND1, which co-localise to DNA (blue). Figures adapted from Bishop *et al*., 1992; Tarsounas *et al*., 1999; Ding *et al*., 2001; Tsubouchi and Roeder, 2002; Schommer *et al*., 2003; Kerendorfer *et al*., 2006.

While the function of the DMC1 protein has been determined as a recombinase, which enables it to conduct the homology dependent strand invasion process (Figure 3.2 A) (Chen *et al.*, 2004), this capability was greatly enhanced in the presence of the HOP2:MND1 (H2M1) protein complex, where in the case of the human proteins a 3-fold increase of strand invasion was observed (Figure 3.2 B). Owing to the important role(s) of these three genes, and that *TaDMC1* has already been shown to have high expression during early meiosis (Crismani *et al.*, 2006), these candidates were selected for isolation and characterisation in bread wheat.

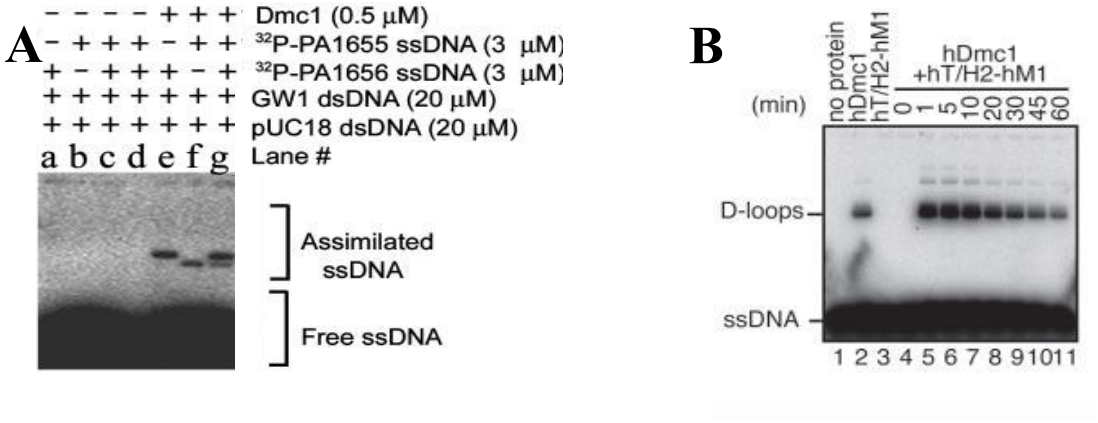


Figure 3.2 - DMC1 strand invasion ability with enhanced activity in the presence of the HOP2:MND1 complex. (**A**) The exchange of genetic material is unable to occur without the presence of the DMC1 protein (lanes a-d). In the presence of DMC1, strand invasion occurs in a homology dependent fashion (lanes e-g). (**B**) Strand invasion capability of *Hs*DMC1 (lane 2; hDmc1) is greatly enhanced in the presence of the human HOP2:MND1 (H2M1) protein complex (Lane 5-8), yet after time this effect is negated *in vitro*. Figures adapted from Chen *et al*., 2004 and Enomoto *et al*., 2006.

3.2 – Material and Methods

3.2.1 – Isolation of the wheat homologues *DMC1*, *HOP2* and *MND1*

3.2.1.1 – *In silico* sequence searching

Previously annotated sequences from the barley *DMC1* and Arabidopsis *HOP2* and *MND1* genes (AAF42940, AJ627484 and AM162278, respectively) were used to interrogate

the public domain databases for corresponding wheat expressed sequence tags (ESTs); using the Basic Local Alignment Search Tool (BLAST; http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 157.0). In the instance that no significant similar ESTs were obtained (*E*-value: $\geq E^{-20}$); the aforementioned sequences where used to obtain an intermediary rice sequence, which was subsequently used to interrogate the wheat ESTs on the public databases.

Corresponding ESTs for *TaDMC1*, *TaHOP2* and *TaMND1* were assembled into contigs for each of the genes and then converted into Vector NTI (Suite 10) sequence files, where open reading frames could be predicted. Oligonucleotides were designed from the contigs, corresponding to sequences flanking both the start and stop codons, for each of the identified *in silico* genes.

3.2.1.2 – Amplification of the wheat *DMC1*, *HOP2* and *MND1* genes

Oligonucleotides used for the isolation of the three genetic sequences were *TaDMC1*_Forward (5′-CGCGAACCTCCGCCTCTTACA-3′), *TaDMC1*_Reverse (5′-GTCC TCAGCACAAATCCTTAAGGCATTAC-3′), *TaHOP2*_Forward (5′-CGCAATGCCGCCT AAATCGC-3′), *TaHOP2*_Reverse (5′-CTCCCATGCCCCTACATCCTCAG-3′), *TaMND1*_ Forward (5′-GTGCTCCATAGCCTGCCATGTCG-3′), *TaMND1*_Reverse (5′-CATATGCA GGCTGTTGAGCCTTACTGC-3′).

Reaction mixtures (20 µL) contained 2 µL of 10x Reaction Buffer (Invitrogen), 0.8 µL of $MgCl_2$ (Invitrogen, 50 mM), 4 µL of dNTP solution (1.25 mM), 1 µL of forward oligonucleotide (10 µM), 1 µL of reverse oligonucleotide (10 µM), 2 µL of cDNA template (Chapter 2, section 2.2.4), 0.4 µL of *Taq* Polymerase (Invitrogen) and 8.8 µL of sterile deionised water. PCR was initiated with a denaturation step at 95°C for 7 minutes, followed by 35 cycles of 94°C for 30 seconds, 62°C, 63°C or 64°C for 30 seconds for *TaHOP2*, *TaDMC1* and *TaMND1* respectively, 72°C for 1 minute, with a final extension at 72°C for 10 minutes. Post PCR reactions were analysed by gel electrophoresis (1.5% agarose) and

visualised using ethidium bromide and a UV transilluminator (FirstLight™ UV illuminator, UVP, USA). PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per manufacturer"s instructions (Macherey-Nagel).

3.2.1.3 – *TaDMC1*, *TaHOP2* and *TaMND1* ligations and bacterial transformation

The purified products were ligated into the pGEM®-T Easy vector (Appendix 3.1) according to the manufacturer"s protocol (Promega manual). Reactions were incubated for 4 hours at room temperature, and subsequently transformed into competent *Escherichia coli* (*E. coli*) DH5α cells by heat shock treatment according to the manufacturer"s protocol (Promega manual).

Transformed cells were plated onto Luria Betani (LB) agar plates (0.5% Yeast extract, 1% Tryptone, 1% NaCl, 1.5% Agar, pH 7.0) containing ampicillin (100 μg mL^{-1}) as a selective agent. Additional selection reagents were added, including 5-bromo-4-chloro-3-indolyl-β-D-galactopyrano-side (X-gal, 50 μL of 2% X-gal in dimethyl formamide) and isopropyl-β-D-thiogalatopyrano-side (IPTG, 100 μL of 0.1 M solution), which allowed for blue/white selection of colonies that contained the pGEM®-T Easy vector and the respective coding sequence insert.

3.2.1.4 – Validation of fragment insertion via colony PCR

White colonies were screened via PCR for the presence of the coding sequences of the three genes (from independent transformation reactions): *TaDMC1*, *TaHOP2* and *TaMND1*. Each colony PCR mixture (20 μL) contained 2 μL of 10x ImmoBuffer (Bioline), 0.75 μL of $MgCl_2$ (Bioline, 50 mM), 2 μL of dNTP solution (1.25 mM each), 0.5 μL of the corresponding forward oligonucleotide (section 3.2.1.2, 10 μM), 0.5 μL of the corresponding reverse oligonucleotide (section 3.2.1.2, 10 μM), 0.2 μL of Immolase™ DNA Polymerase (Bioline, 5 U/μL) and 14.05 μL of sterile deionised water. Using a sterile pipette tip, individual white colonies were selected from plates, and dipped into single 20 μL PCR

mixtures before being placed into 5 mL liquid aliquots of LB/Ampicillin (100 μg mL^{-1} as a selective agent) for overnight growth and generation of glycerol stocks.

PCR was initiated with a denaturation step at 95°C for 7 minutes followed by 35 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute 30 seconds, with a final extension at 72°C for 10 minutes. Post PCR mixtures were analysed by gel electrophoresis (1.5% agarose) and visualised previously in section 3.2.1.2. Plasmid DNA was isolated from positive colony liquid cell cultures as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen, AUS) to provide a template for sequencing reactions and analysis of the coding regions of the three genes of interest.

3.2.1.5 – Sequencing of the meiotic genes *TaDMC1*, *TaHOP2* and *TaMND1*

Sequencing reactions (10 μL) contained 0.5 μL of Big Dye Terminator (version 3; Applied Biosystems), 1.5 μL of Big Dye Buffer, 1 μL of SP6 (5′-ATTTAGGTGACACTA TAG-3′) or T7 (5′-TAATACGACTCACTATAGGG-3′) oligonucleotide (10 μM), 2 μL of purified positive plasmid DNA and 5 μL of sterile deionised water. Thermal cycling was initiated with a denaturation step at 96°C for 30 seconds, followed by 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and extension at 60°C for 4 minutes.

Reaction products were precipitated in preparation for capillary electrophoresis. To the 10 μL reaction mixture, 40 μL of freshly prepared 75% isopropanol was added; and the 50 μL mixture transferred to a 1.5 mL microcentrifuge tube. Tubes were subsequently incubated at room temperature for 20 minutes, during which time the DNA precipitated out of solution. Tubes were then centrifuged at 16500 *g* for 20 minutes, followed by removal of the resulting supernatant. A further 200 μL of freshly prepared 75% isopropanol was added for DNA washing, followed by centrifugation at 16500 *g* for 10 minutes. The supernatant was then removed without dislodging the DNA pellet; with tubes subsequently air-dried for approximately 10 minutes in a 37°C heat block (with protection from the light). Samples were analysed by the Institute of Medical and Veterinary Science (IMVS), Adelaide.

Returned sequences were compiled in Contig Express (Vector NTI Suite 10), with repeated reads from plasmids assembled into contigs. Plasmid sequence within each of the contigs was identified using VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html, NCBI-GenBank Flat File Release 159.0) and removed. A Vector NTI sequence file was then generated using the confirmed sequences, where open reading frames could be identified. Resulting sequences were then compared to the *in silico* obtained sequences using Basic Local Alignment Search Tool (BLAST bl2seq option). The open reading frames for the three genes were then used to examine the public domain database Basic Local Alignment Search Tool (BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 159.0) for determining orthologous sequence similarity.

3.2.2 – Comparative protein sequence analysis

3.2.2.1 – Conserved domains

Using the predicted translation of the open reading frame for each of the three candidates, the BLAST Conserved Domain tool of the NCBI website (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi NCBI-GenBank Flat File Release 159.0) was queried in order to detect any possible conserved protein domains.

3.2.2.2 – Full length sequence alignments

ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) was used as a tool for the comparison alignment of amino acid sequences of the three candidates along with their orthologues from other species (Table 3.1).

Table 3.1 - Orthologues of candidates used in amino acid sequence alignments. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

Candidate	Species	Protein accession
DMC1	*Os*	BAB84121
	At	AAC49617
	Mm	BAA10969
	Hs	CAG30372
	Gg	XP_425477
	Sc	NP_01106
HOP2	*Os*	BAG95812
	At	CAF28783
	Mm	BAA23155
	Hs	BAA92872
	Sc	NP_011482
MND1	*Os*	NP_001062766
	At	NP_194646
	Mm	NP_084073
	Hs	AAH32142
	Sc	NP_011332

3.2.2.3 – Phylogenetic analysis of candidates

Open reading frames from orthologous sequences (Table 3.2) obtained in section 3.2.1.5 were used for the construction of phylogenetic trees, as the phylogenetic analysis program translates these nucleotide sequences into all six amino acid frames prior to multiple alignments for analysis using the in-built ClustalW function. *Os*DMC1a was chosen to be the representative sequence for rice, despite rice also having the highly similar *Os*DMC1b protein.

The evolutionary history of the three candidates was inferred using the phylogenetic Neighbour-Joining method. MEGA software (version 4.0, Molecular Evolutionary Genetics Analysis) (Tamura *et al.*, 2007) with default setting parameters used except for the following: the "pair-wise deletion" was used for the amino acid sequences, the bootstrap value was set at 10,000 re-samplings and the model setting was Amino:Poisson correction.

Table 3.2 - Orthologues of candidates used for phylogenetic analysis. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa, Sc = Saccharomyces cerevisiae.*

Candidate	Species	Nucleotide accession
DMC1	*Os*	AB046620
	At	NM_113188
	Mm	D64107
	Hs	CR456486
	Gg	XM_425477
	Sc	YER179W
HOP2	*Os*	AK102972
	At	AJ627484
	Mm	AB000121
	Hs	AB030304
	Sc	YGL033W
MND1	*Os*	NM_001069301
	At	NM_119061
	Mm	NM_029797
	Hs	BC032142
	Sc	YGL183C

3.2.2.4 – Three dimensional modelling

The CPHmodel server, provided by the Centre for Biological Sequence Analysis (CBS) (http://www.cbs.dtu.dk/services/CPHmodels), was used for computational 3D modelling of the amino acid sequences of full length *Sc*DMC1, *Mm*DMC1, *Os*DMC1 and *Ta*DMC1 proteins (corresponding accessions in Table 3.1). The subsequent output was then visualised using MacPyMol (DeLano Scientific LLC) to generate the 3D structure for each protein.

3.2.3 – Southern blot analysis of *TaDMC1*, *TaHOP2* and *TaMND1* for determining chromosome location

Chromosome location was determined via Southern blot analysis as per Chapter 2 section 2.2.5.2; with the exception of using 9 µL and 5 µL (approximately 150 ng) of the PCR products (section 3.2.1.2) of *TaDMC1* and *TaHOP2/MND1* respectively.

3.2.4 – Quantitative real-time PCR expression analysis

As the expression of the *TaDMC1* homologue has previously been reported by Crismani *et al*. (2006), quantitative real-time PCR was only conducted for *TaHOP2* and *TaMND1*. Oligonucleotides used for the expression analysis of *TaHOP2* and *TaMND1* were as follows: *TaHOP2*_QPCR_F (5′-GCAGAAAATCTACCTTGCTCGGC-′3), *TaHOP2*_ QPCR_R (5′-GCCTCCTTTGACCTAATCTCTTCCAGT-′3), *TaMND1*_QPCR_F (5′-AGCC TGGTGGATGATGATCTTGTC-′3) and *TaMND1*_QPCR_R (5′-TCAGTGTCTTCCCTGC CTCTTTTC-3′). Quantitative real-time PCR was conducted as per Crismani *et al.* (2006), using the cDNA template produced for the analysis of expression during that study, with the optimal acquisition temperatures for *TaHOP2* and *TaMND1* being 76°C and 75°C, respectively .

3.2.5 – Protein localisation

3.2.5.1 – Immunolocalisation in polyacrylamide embedded meiocytes

A primary attempt at immunolocalisation and visualisation using confocal microscopy of the wheat DMC1 and MND1 proteins was conducted as per Boden and colleagues (2009), using antibodies raised against the Arabidopsis DMC1 and MND1 proteins kindly donated by Dr Mathilde Grelon (INRA, France). Secondary attempts at immunolocalisation for only *Ta*MND1 protein were conducted, as per the primary attempt, except that anthers were fixed using 2% paraformaldehyde and meiocytes were permeabilised for 2 hours. The antibody for *Ta*ASY1 (Boden *et al*., 2009) was used as both a positive experimental control as well as a marker of meiotic chromosomes.

3.2.5.2 – Meiotic chromosome spreads

3.2.5.2.1 – Preparation of meiotic chromosome spreads

Four wheat anthers, at particular stages of meiosis, were placed into 5 µL of 2% paraformaldehyde (PFA) (in 50 mM potassium phosphate buffer, pH 7.5) fixative, where the meiocytes were extruded from the anthers and mixed in solution for maximum fixation efficiency. 5 µL of 2% PFA fixative was added, and the slide incubated for 6 minutes in a humidity chamber. This was followed by the addition of another 5 µL of 2% PFA fixative with a further 2 minutes incubation in the humidity chamber; followed by the addition of 0.5 µL 0.4% Triton X (in 1x PBS, Phosphate buffered saline [137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4]). The cell suspension on the glass slide was mixed gently by rotation for two minutes. A coverslip (22mm x 22mm) was then placed, lifted and placed again over the meiocyte mixture. Finally, the slide was snap frozen in liquid nitrogen before being placed at -80°C for 1 hour.

3.2.5.2.2 – Antibody incubation

Slides were removed from the -80°C incubation, and before thawing, the cover slide was removed. Slides were then allowed to thaw briefly before being bathed in 1x PBS for 5 minutes, 0.1 M ammonium chloride (in 1x PBS) for 5 minutes, 1x PBS for 5 minutes. Blocking buffer (2 mL of 5% skim milk powder w/v in 1x PBS) was then placed onto the slide for 30 minutes at room temperature.

Slides were rinsed in 1x PBS before incubation with the primary antibody (50 µL) in blocking buffer at a 1:100 dilution for anti-*At*MND1(Rat):anti-*Ta*ASY1(Rabbit), and a 1:20 dilution for the anti-*At*DMC1(Rabbit):anti-*At*ASY1(Rat) (1:100, Higgins *et al*., 2004). Slides were covered with a coverslip and then incubated at room temperature for 1 hour in a humidity chamber. Slides were further incubated at 4°C over 48 hours. Coverslips were removed by rinsing slides in 1x PBS, followed by incubation with 50 µL 1:100 diluted

secondary antibody in blocking buffer containing donkey anti-Rat and anti-Rabbit IgG antibodies conjugated with Alexafluor® 488 and 568, respectively (Invitrogen). Slides were incubated for 2 hours at room temperature in a humidity chamber. Following incubation, slides were washed four times in 1x PBS for 15 minutes each. Ten µL of DAPI (2.0 µg mL^{-1} in VectaShield™, Vector Laboratories, USA) was applied to wet slides; with incubation for 3 hours at room temperature protected from light. Slides were gently pressed for the removal of excess DAPI, and edges were sealed with clear nail varnish. Slides were stored until visualisation as per Boden *et al.,* (2009).

3.2.5.3 - Western Analysis

The affinity of the anti-*At*DMC1 antibody for the Arabidopsis and wheat DMC1 proteins was tested against protein extracts from both vegetative and meiotic tissues.

3.2.5.3.1 – Protein isolation using a TCA extraction method

Tissues for protein extractions were ground in 1.5 mL tubes using tube pestles, in the presence of liquid nitrogen. To tissue samples, 800 µL of protein extraction buffer (10% w/v trichloroacetic acid, 20 mM dithiothreitol, dissolved in 100% acetone, pre-chilled to -20°C) was added and mixed well, followed by overnight incubation at -20°C. Samples were then centrifuged at 16500 *g* for 15 minutes at 4°C with the resulting supernatant removed and replaced with 800 µL of wash buffer (20 mM dithiothreitol, dissolved in 100% acetone, pre-chilled to -20°C) and inverted to mix. Samples were then incubated at -20°C for 1 hour, followed by 16500 *g* for 15 minutes at 4°C with resulting supernatant removed. Another 800 µL of wash buffer was applied to samples and inverted to mix, followed by 30 minutes at -20°C. Samples were then centrifuged at 16500 *g* for 15 minutes at 4°C with resulting supernatant removed and a further 800 µL of wash buffer applied to the sample, with inversion to mix. Samples were incubated at -20°C for 30 minutes, followed by centrifugation at 16500 *g* for 15 minutes at 4°C with resulting supernatant removed. Samples were allowed

to dry in a laminar flow cabinet (Clemco, AUS) before being snap frozen in liquid nitrogen and stored at -80°C until ready for re-suspension.

To each tube, 60 µL of suspension buffer (7 M urea, 2 M thiourea, 2% w/v CHAPS [3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate], 30 mM Tris, dissolved in sterile deionised water) was then added. Gentle agitation was applied to tubes on a Bioline platform rocker (Edwards Instrument Company, Australia) for 3 hours at room temperature, followed by centrifugation at 77 *g* for 15 minutes with supernatant being transferred to new tubes. Samples were stored at -80°C.

3.2.5.3.2 – Protein sample quantitation

Bradford assays were conducted to determine protein concentration for the equal loading of protein samples for western analysis. Triplicates of a Bovine Serum Albumin (BSA) dilution series were prepared with concentrations of 13.7 µg mL^{-1}, 10.275 µg mL^{-1}, 6.85 µg mL^{-1}, 3.425 µg mL^{-1}, and 1.37 µg mL^{-1} using the BIO-RAD Protein Assay Standard II, lyophilised BSA, diluted in sterile deionised water. A 200 µL aliquot of the Bradford reagent (BIO-RAD) was added to 800 µL of each sample of the dilution series and incubated for 7 minutes at room temperature. Triplicates of 1:200 dilutions of the protein samples isolated in section 3.2.5.3.1 were prepared using sterile distilled water, with 200 µL of Bradford reagent added to 800 µL of diluted protein samples. Tubes were then incubated at room temperature for 7 minutes. All samples were then quantified using a spectrophotometer at an absorbance wavelength of 595 nm. Values obtained for both the standard and protein samples were entered into Genstat (version 8.0, Numerical Algorithms Group). The values of the BSA standards were used to generate a standard curve, with the prediction of the concentration of the protein samples based on their recorded absorbance readings.

3.2.5.3.3 – SDS-PAGE and electroblot transfer of protein samples

Protein samples were prepared for electrophoresis in reactions (15 μL) containing increasing concentrations of proteins (3 μg to 50 μg) with 3 μL of NuPAGE® LDS sample buffer (4x) (Invitrogen), 2 μL of β-Mercaptoethanol and sterile deionised water. Samples were heated at 70°C for 10 minutes before being loaded into 15-well NuPAGE® Novex® 4-12% Bis(2-hydroxyethyl)-imino-tris(hydoxymethyl)-methane (Bis-Tris) mini-gels (Invitrogen). BIO-RAD Precision Plus Dual Colour Protein Sample (10 μL) was also loaded into the gels as a protein size marker, with samples being electrophoresed using conditions as suggested by the manufacturer. Upon completion of the SDS-PAGE, proteins were transferred to a 0.22 μm NitroBind nitrocellulose membrane (GE Osmonics, USA) using electroblotting according to the Invitrogen XCELL II™ Blot Manual; with conditions as follows: 30V for 1 hour, starting current 170 mA and ending at 110 mA.

3.2.5.3.4 – Western blot analysis of DMC1

Once electroblotting was completed, the membranes were dried at 37°C for 10 minutes. Membranes were then immersed and incubated in blocking solution, 5% skim milk powder w/v in 1x Tris-buffered-saline + Tween20 (1x TBS/Tween20) (4 g NaCl, 0.1 g KCl, 1.5 g Tris[hydroxymethyl]aminomethane, 250 μL Tween20 and 500 mL deionised water, pH 7.2) overnight at 4°C, with gentle agitation on a Bioline platform rocker. The membranes were then washed 3 times with 1x TBS/Tween20 for 5 minutes with gentle agitation at room temperature.

Primary antibody incubations were conducted in 5 mL of a 1:200 dilution of anti-*At*DMC1, and a separate 1:500 dilution of anti-*Ta*ASY1, both in 1x TBS/Tween20. Membranes were sealed into plastic bags with the primary antibody and allowed to incubate overnight at 4°C with gentle agitation. Washing with 1x TBS/Tween 20 was then performed 3 times at 4°C for 15 minutes with gentle agitation. The secondary IgG antibody, a donkey anti-Rabbit-alkaline phosphatase conjugated (Invitrogen), was then added at a 1:5000 dilution (in

1x TBS/Tween 20) in a total volume of 25 µL and left to incubate at room temperature for 2 hours with gentle agitation.

Visualisation for westerns was done using approximately 3 mL of 5-Bromo-4-Chloro-3′-Indolyphosphate p-Toluidine Salt (BCIP) (BIO-RAD). Colour development was allowed to continue until desired signal intensity was achieved, with copious amounts of water then being added to quench the reaction. The membranes were then dried briefly with paper towel and then scanned using an Epson Perfection 4180 Photo Scanner.

3.2.6 – Heterologous protein expression

3.2.6.1 – Protein expression vector preparation

3.2.6.1.1 – Isolation of the coding regions of the three meiotic genes for protein production

Oligonucleotides used for the amplification of the coding regions for the purposes of producing recombinant proteins included *TaDMC1*_ATG (5′-ATGGCGCCGTCCAAGCAG TAC-3′), *TaDMC1*_Stop (5′-GTCCTCAGCACAAATCCTTAAGGCATTAC-3′), *TaHOP2*_ ATG (5′-ATGCCGCCTAAATCGGACAGC-3′), *TaHOP2*_Reverse (section 3.2.1.2), *TaMND1*_ATG (5′-ATGTCGAAGAAGAGGGGCCTTTCC-3′) and *TaMND1*_Reverse (section 3.2.1.2).

Reaction mixtures (20 µL) contained 2 µL of 10x reaction buffer (Invitrogen), 0.7 µL of $MgCl_2$ (Invitrogen, 50 mM), 3.2 µL of dNTP solution (1.25 mM), 1 µL of forward oligonucleotide (10 µM), 1 µL of reverse oligonucleotide (10 µM), 0.2 µL *Taq* Polymerase (5 U/µL) (Invitrogen) 2 µL of DNA template (isolated coding region:pGEM®-T Easy vectors, section 3.2.1.4) and 9.9 µL of sterile deionised water. PCR was initiated with a denaturation step at 95°C for 7 minutes, followed by 35 cycles of 94°C for 30 seconds, 61°C for 30 seconds, 72°C for 1 minute, with a final extension at 72°C for 10 minutes. Post PCR reactions were analysed by gel electrophoresis (1.5% agarose) and visualised using ethidium bromide

and a UV transilluminator. PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per manufacturer"s instructions (Macherey-Nagel).

3.2.6.1.2 - TOPO cloning and transformation of *TaDMC1*, *TaHOP2* and *TaMND1*

The purified coding regions from section 3.2.6.1.1 were ligated into the pCR®8/GW/TOPO® vector (Appendix 3.2) as per the pCR®8/GW/TOPO® TA Cloning® Kit user manual. Transformation of OneShot® TOP10 competent *E. coli* cells was conducted with the addition of 2 µL of the ligation reaction, following the chemical transformation protocol according to the pCR®8/GW/TOPO® TA Cloning® Kit user manual. Transformed cells were plated onto Luria Betani (LB) agar plates containing spectinomycin (50 µg mL^{-1} as a selective agent).

3.2.6.1.3 – Verification of the inserted coding region into pCR®8/GW/TOPO® by colony PCR

PCR was used to screen colonies for the validation of the coding sequences of the three genes. Colony PCR was conducted as per section 3.2.1.4, with the exception of using oligonucleotides GW1 (5′-GTTGCAACAAATTGATGAGCAATGC-3′) and GW2 (5′-GTT GCAACAAATTGATGAGCAATTA-3′) (Invitrogen), and using spectinomycin as the selective agent for overnight culturing. PCR was initiated with a denaturation step at 95°C for 7 minutes followed by 35 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute, with a final extension at 72°C for 10 minutes. Post PCR mixtures were analysed by gel electrophoresis (1.5% agarose) and visualised using ethidium bromide and a UV transilluminator. Plasmid DNA was isolated from positive colony liquid cell cultures as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen, AUS) for verification of insert via sequence analysis.

3.2.6.1.4 – Sequence verification of the coding regions for *TaDMC1*, *TaHOP2* and *TaMND1*

Sequence analysis was used to verify vectors for both insertion and orientation of the coding regions of the genes. The sequencing protocol was as per section 3.2.1.5, with the exception of using oligonucleotides M13F (5′-GTTTTCCCAGTCACGAC-3′) or M13R (5′-CAGGAAACAGCTATGACC-′3). Samples were analysed by the Institute of Medical and Veterinary Science (IMVS), Adelaide.

Returned sequence runs were compiled into Contig Express and results were compared to the confirmed coding sequences (section 3.2.1.5). Upon verification of both the insert sequence and orientation within the vector, plasmids with the correct orientation (attL1 – 5′ coding 3‟ – attL2) were then used in the Gateway® recombination reaction.

3.2.6.1.5 - Gateway® recombination reaction between coding region:pCR®8/GW/TOPO® vectors and pDEST17® protein expression vector and subsequent transformation

Gateway® recombination technology was used to transfer the coding regions of the three genes from the pCR®8/GW/TOPO® vector to the pDEST17® protein expression vector (Appendix 3.3); with 1µL of coding region:pCR®8/GW/TOPO® (section 3.2.6.1.3) used in the recombination reaction as per the manufacturer‟s protocol (Gateway® LR Clonase™ II Enzyme Mix product, Invitrogen). Transformation of OneShot® TOP10 competent *E. coli* cells with 2 µL of the recombination reaction was conducted as per the manufacturer‟s protocol (Gateway® LR Clonase™ II Enzyme Mix product, Invitrogen). Transformed cells were subsequently plated onto LB agar containing ampicillin (100 µg mL^{-1} as a selective agent).

3.2.6.1.6 – Colony PCR identification of positive colonies with coding region:pDEST17® protein expression vectors

Resulting colonies were screened using PCR to validate the presence of the coding regions, with reactions (20 µL) containing 2 µL of 10x ImmoBuffer (Bioline), 0.75 µL of

$MgCl_2$ (Bioline, 50 mM), 3.2 µL of dNTP solution (1.25 mM each), 0.5 µL of the corresponding forward oligonucleotide (section 3.2.6.1.1, 10 µM), 0.5 µL of the corresponding reverse oligonucleotide (section 3.2.6.1.1, 10 µM), 0.2 µL of Immolase™ DNA Polymerase (Bioline, 5 U/µL), and 12.85 µL of sterile deionised water. Using a sterile pipette tip, individual colonies were selected from plates and dipped into single 20 µL PCR mixtures before being placed into 5 mL liquid aliquots of LB/Ampicillin (100 µg mL^{-1} as a selective agent) for overnight growth and generation of glycerol stocks.

PCR was initiated with a denaturation step at 95°C for 7 minutes followed by 35 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute, with a final extension at 72°C for 10 minutes. Post PCR mixtures were analysed by gel electrophoresis (1.5% agarose) and visualised using ethidium bromide and a UV transilluminator. Plasmid DNA was isolated from positive colony liquid cell cultures as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen, AUS) for sequence analysis of the coding regions of the three genes of interest.

3.2.6.1.7 – Sequence verification of coding regions in the protein expression vector pDEST17®

Sequence analysis was used to verify vectors for both insertion of the coding region of the genes and the orientation of each fragment. The sequencing protocol was as per section 3.2.1.5, with the exception of using the T7 oligonucleotide exclusively. Samples were analysed by the Institute of Medical and Veterinary Science (IMVS), Adelaide.

Returned sequence runs were compiled into Contig Express with confirmed coding sequences (section 3.2.1.5) and the pDEST17® vector sequence. Using compiled consensus sequences in Vector NTI (suite 10), open reading frames from the positive colonies were investigated for correct orientation, and in-frame recombinant protein coding sequences.

3.2.6.2 – Heterologous protein expression

3.2.6.2.1 - Transformation of coding region:pDEST17® vector into protein expression optimised bacterial cells

BL21-AI protein expression optimised cells were transformed with 3 µL of confirmed coding region:pDEST17® protein expression vector as per section 3.2.1.3. Transformed cells were plated onto LB agar plates containing ampicillin (100 µg mL^{-1} as a selective agent). A single colony for each transformant was picked from the growth plate and placed into 5 mL of LB/Carbenicillin (50 µg mL^{-1}). The colony was allowed to culture overnight at 37°C with agitation in an orbital mixer incubator (Ratek); followed by the production of glycerol stocks (1 mL liquid cell culture and 1 mL of 50% glycerol, snap frozen in liquid nitrogen and stored at -80°C) for all three candidates.

3.2.6.2.2 – BL21-AI cell culturing for protein production

Starter cultures of BL21-A1 protein expression cells were commenced with the inoculation of an individual 5 mL starting culture with 200 µL of glycerol stock (section 3.2.6.2.1) followed by overnight culturing at 37°C with agitation. Inoculation of individual 5 mL cultures into 250 mL LB/Carbenicillin (50 µg mL^{-1}) in 500 mL sterile conical flasks was conducted, with duplication of 250 mL cultures for the desired final volume (for example 4x 250 mL for the final 1 L of protein production culture). These 250 mL cultures were then allowed to culture for 2 hours at 37°C with agitation.

3.2.6.2.3 – Induction of protein production and cell collection

After the aforementioned 2 hour incubation, 0.4% w/v of L-(+)-arabinose (Sigma-Aldrich) in cell culture (from section 3.2.6.3.2) was used to induce protein production; with no L-(+)-arabinose added to the sample that was used as a non-induced control. Cell cultures were left to grow further at 37°C with agitation for 4 hours, followed by cell collection. Subsequently, cells were collected in 50 mL aliquots which were spun at 3000 *g* for 15

minutes at 4°C for cell pelleting. The resulting supernatant was discarded and the process repeated until the final desired culture volume was collected. Cell pellets were snap-frozen using liquid nitrogen and stored at -80°C until protein extraction and purification.

3.2.6.3 – Heterologous protein extraction

3.2.6.3.1 – *Ta*HOP2 protein purification

3.2.6.3.1.1 – Primary *Ta*HOP2 protein isolations

For primary *Ta*HOP2 protein isolation, 1 L of both induced and non-induced cell cultures were removed from storage at -80°C and resuspended in 30 mL of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 5mM imidazole, 1% Triton-X, pH 8), and mixed vigorously. To the resuspended cells, lysozyme (1 mg mL^{-1}), RNase A (10 µg mL^{-1}) and DNase I (166 µg $µL^{-1}$) was added, gently shaken to mix, and incubated on ice for 30 minutes. The suspension mixture was then snap-frozen in liquid nitrogen, thawed and vortexed for 30 seconds; with this process repeated three times in total. Cells were then sonicated six times for 10 seconds each time, with resting on ice for 30 seconds between each sonication. The homogenised mixture was then centrifuged at 10000 *g* for 20 minutes at 4°C with a 100 µL aliquot taken for gel analysis (cell lysate).

The resulting supernatant was transferred to a new 50 mL tube for selective ammonium sulphate precipitation, where 0.24 mg mL^{-1} ammonium sulphate was added and shaken vigorously to mix. The mixture was centrifuged at 10000 *g* for 20 minutes at 4°C, with supernatant being transferred to a new 50 mL tube. Ammonium sulphate (0.13 mg mL^{-1}) was added to the suspension, shaken vigorously to mix and centrifuged at 10000 *g* for 20 minutes at 4°C with supernatant being discarded. To the cell lysate pellet, 10 mL of binding buffer (20 mM NaH_2PO_4, 500 mM NaCl, 20mM imidazole, pH 7.4) was added and the pellet resuspended. An aliquot of 100 µL was taken for gel analysis (loaded).

At 4°C the remaining sample was loaded very slowly into a previously equilibrated HisTrap^{TM}HP 5 mL column (GE Healthcare Life Sciences, UK), with a 100 µL aliquot of flow through (flow through) collected for gel analysis. The column was washed with 100 mL of binding buffer, with 100 µL aliquots taken at 2 mL (wash 2), 50 mL (wash 50) and 100 mL (wash 100) for gel analysis. The *Ta*HOP2 protein was then eluted from the column with 3x 5 mL elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 250mM imidazole, pH 7.4), with 100 µL aliquots of the resulting elution fractions taken for gel analysis (Elution 1, 2 and 3). Concentrations of the elutions were determined as per section 3.2.5.3.2.

Protein purification samples (9 µL of collected aliquots with 3 µL of NuPAGE® LDS Sample Buffer (4x) (Invitrogen)) were heated at 70°C for 10 minutes before being loaded into 15-well NuPAGE® Novex® 4-12% Bis-Tris mini gels. BIO-RAD Precision Plus Dual Colour Protein Ladder (10 µL) was also loaded onto the gels, which were electrophporesed according to the manufacturer‟s instructions. After electrophoresed, the protein gels were removed from the plastic casing and placed into 50 mL of fixing solution (15 mL ethanol, 5 mL acetic acid, 30 mL sterile deionised water) for 1 hour. After incubation, the fixing solution was then replaced with staining solution (50% methanol, 7% acetic acid, 0.125 w/v Brilliant Blue G (Sigma)) and left at room temperature overnight with gentle agitation. The protein gels were then destained using coomassie destain (50% methanol, 10% acetic acid) at room temperature. The protein gel was then scanned using an Epson Perfection 4180 Photo Scanner.

3.2.6.3.1.2 – Secondary *Ta*HOP2 protein purification

Secondary purification of *Ta*HOP2 protein was as per section 3.2.6.3.1.1, with the exception of: 1) having a total volume of 2 L induced cells; 2) not repeating an extraction for a non-induced sample; and, 3) doing four elutions instead of three. Elution 2 and 3 from this secondary purification were combined and concentrated using the Millipore® Centriprep 10

kDa Centrifugation Filter Device as per the manufacturer"s instructions to a final volume of approximately 1 mL. Protein concentration was determined as per section 3.2.5.3.2.

Upon the positive secondary protein purification results, the gel band corresponding to *Ta*HOP2 was sent to the Adelaide Proteomics Centre for protein sequencing using Tryptic digestion, Matrix Assisted Laser Desorption Ionisation mass (MALDI-MS) and tandem mass spectrometry (MS/MS).

3.2.6.3.2 – *Ta*DMC1 protein extractions

3.2.6.3.2.1 – Primary *Ta*DMC1 protein purification

Primary purification of the *Ta*DMC1 protein was conducted as per section 3.2.6.3.1 with the exception of not conducting the ammonium sulphate precipitation. Extraction elutions 2 and 3 were combined in both induced and non-induced samples and were concentrated down as per section 3.2.6.3.1.2.

3.2.6.3.2.2 – Secondary *Ta*DMC1 protein expression

The secondary attempt for *Ta*DMC1 protein purification was done using cells induced at 18°C for 4 hours instead of at 37°C as per section 3.2.6.2.3. Protein extraction and purification were completed as per section 3.2.6.3.2.1.

3.2.6.3.2.3 – *Ta*DMC1 protein expression and repression with subsequent protein extraction

The concurrent expression and repression of *Ta*DMC1 protein expression was done by inoculating 25 mL LB/Carbenicillin (50 µg mL^{-1}) with the *Ta*DMC1:pDEST17® glycerol stock. This was cultured at 37°C for 10 hours with agitation. Following this, the 25 mL cell cultures were then added to a further 225 mL of LB/Carbenicillin (50 µg mL^{-1}) and incubated at 37°C for 40 minutes with agitation, until an OD_{600} measurement of 0.4 was recorded. Upon recording an OD_{600} measurement of 0.4 in the cell culture, 0.4% w/v of L-(+)-arabinose was added to one 250 mL culture for an induced sample, while 0.5% w/v of D-glucose was added

to another 250 mL culture for a repressed sample. A third 250 mL culture with nothing added was also used as a non-induced, non-repressed sample. All cultures were then placed at 23°C for a further 10 hours with agitation. Subsequent protein purification was as per section 3.2.6.3.2.1.

3.2.6.3.3 – *Ta*MND1 protein extractions

3.2.6.3.3.1 – Primary protein purification

Primary *Ta*MND1 protein purification was conducted as per section 3.2.6.3.1.1, except that there were four elution fractions instead of three.

3.2.6.3.3.2 – Secondary protein purification

Secondary *Ta*MND1 protein purification was conducted as per section 3.2.6.3.1.1, except that ammonium sulphate precipitation was not conducted. Elution samples from the induced and non-induced were combined for concentrating as per section 3.2.6.3.1.2.

3.2.7 – DNA binding assays

3.2.7.1 – Digestion of double strand DNA

PhiX174 RF II (1 µg, New England Biolabs (NEB), USA) was linearised in 10 µL reactions containing 4 µL of Enzyme Buffer 2 (NEB), 2 µL of *Xho*I (2 Units, NEB) and 3 µL of sterile deionised water. Reactions were incubated at 37°C for 5 hours, followed by enzyme heat inactivation at 65°C for 20 minutes.

3.2.7.2 – Protein DNA binding assays

Two forms of DNA binding ability were assessed: 1) the ability of each of the candidates to bind to either single-stranded or double-stranded DNA alone (non-competitive DNA binding assay); and, 2) the ability of each candidate to bind either single-stranded or double-stranded DNA in the presence of both species of DNA (competitive DNA binding

ability). DNA binding ability assays were conducted as per (Khoo *et al.*, 2008) with the exception of the non-competitive binding assays in which either single-stranded or double-stranded was added.

3.3 – Results

3.3.1 – Isolation of the wheat homologues *DMC1*, *HOP2* and *MND1*

3.3.1.1 – *In silico* identification of *TaDMC1*, *TaHOP2* and *TaMND1*

The previously annotated sequences for *DMC1*, *HOP2* and *MND1* genes were successfully used for the identification of highly similar wheat ESTs, as shown in Table 3.3. These ESTs were assembled into contigs, with open reading frames identified; *TaDMC1* (1032 bp), *TaHOP2* (681 bp) and *TaMND1* (621 bp). Oligonucleotides were designed from these *in silico* obtained sequences for confirmation sequencing from cDNA prepared from wheat meiotic tissues.

Table 3.3 - Wheat ESTs identified with high similarity to *DMC1*, *HOP2* and *MND1*. BLASTn analysis was conducted using the previously annotated accessions of AAF42940 (*HvDMC1*), AJ627484 (*AtHOP2*) and AM162278 (*AtMND1*) for the identification of wheat ESTs which have high similarity (E-value ≥ E^{-20}). * denotes wheat Uni-genes (a collection of wheat ESTs). *At = Arabidopsis thaliana* and *Hv = Hordeum vulgare.*

Gene	Intermediary Rice Accession	Wheat ESTs Accession	Similarity to query annotation
HvDMC1	—	Ta.47427*	79%
		Ta.30833*	
AtHOP2	—	AL830241	89%
		CJ629398	89%
		CJ520686	85%
AtMND1	Os09g10850	CA700892	90%
		BQ841694	90%
		CJ505449	89%
		BQ168165	91%
		CJ517951	92%

3.3.1.2 – Isolation of *TaDMC1*, *TaHOP2* and *TaMND1* candidates from bread wheat

Isolation of the three candidate cDNAs from wild-type Chinese Spring meiotic tissues was successful, with the amplification of 1257 bp, 707 bp and 660 bp cDNA fragments corresponding to *TaDMC1*, *TaHOP2* and *TaMND1* respectively (Figure 3.3).

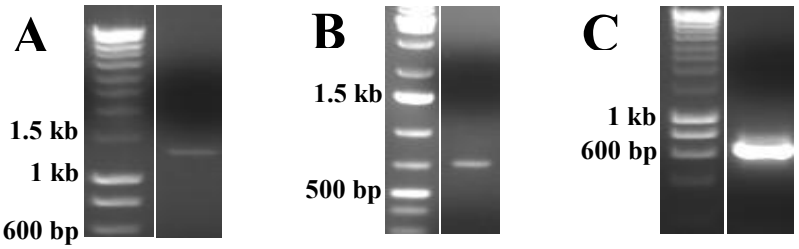


Figure 3.3 - PCR amplification of *TaDMC1*, *TaHOP2* and *TaMND1* from bread wheat. Amplification of *TaDMC1* (**A**), *TaHOP2* (**B**) and *TaMND1* (**C**) was successful, with products at the expected sizes of 1257 bp, 707 bp and 660 bp, respectively (based on the length of products expected, given the oligonucleotides designed). 1 kb DNA molecular weight marker is shown in lane 1 across all three gels; not all band sizes are labelled.

Ligation of the three coding regions with the pGEM®-T Easy vector was successful, with multiple white colonies being isolated. Figure 3.4 illustrates that of the ten white colonies screened via PCR for each candidate, all were found to contain the predicted coding regions of the genes of interest. For each candidate two putative positive plasmids (Figure 3.4, yellow boxes) were isolated for sequence conformation.

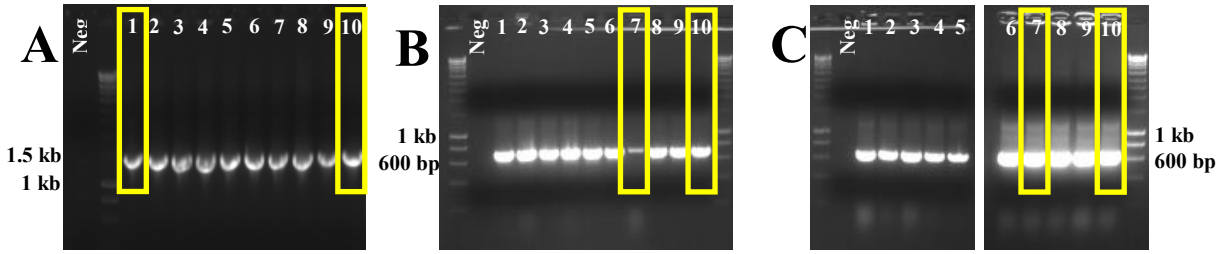


Figure 3.4 - Identification of positive colonies for the insertion of *TaDMC1*, *TaHOP2* and *TaMND1*. All ten colonies tested for *TaDMC1* **(A)**, *TaHOP2* **(B)** and *TaMND1* **(C)** were positive for the isolated product via colony PCR. Plasmids from two colonies (*yellow* boxes) for each candidate were isolated for sequencing. Neg = negative control. A 1 kb DNA molecular weight marker is shown in all gel photos; however not all band sizes are labelled.

Sequencing of the putative positive plasmids confirmed that *TaDMC1*, *TaHOP2* and *TaMND1* had been isolated and cloned. These sequences were found to be 98% (*TaDMC1*), 97% (*TaHOP2*) and 99% (*TaMND1*) similar to the *in silico* wheat EST sequences identified in section 3.3.1.1. This difference observed in sequences between those confirmed and those isolated *in silico* is most probably due to different cultivars of wheat used; with this investigation using the wheat cultivar Bobwhite 26 compared to the *in silico* cultivars including Chinese Spring, Brevor and Cheyenne. These small sequence differences are not expected effect the functionality between the Bobwhite 26 genes and thoses from other wheat cultivar, since there are larger differences exsisting between species and kingdoms yet these genes, in particular DMC1, has been found to have high functional similairy. Open reading frames of the same length as in the *in silico* sequences (section 3.3.1.1) were identified in the amplified sequences using Vector NTI (full sequences in Appendix 3.4) and used for further analysis. Table 3.4 shows orthologous sequence analysis using the public domain databases. *TaDMC1* is conserved in part with its orthologues, having 53% identity and 72% similarity to the yeast protein sequence. The sequences of the other two proteins, *TaHOP2* and *TaMND1*, show reduced levels of identity when compared to their respective orthologues. Even so, there is still a high level of similarity between the more closely related plant species (Table 3.4).

Table 3.4 - Summary of similarity between *DMC1*, *HOP2* and *MND1* orthologues. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

Candidate	Species	Nucleotide accession	E-value	Protein accession	E-value	Identities %	Positives %
*Ta*DMC1	*Os*	AB046620	0.0	BAB84121	0.0	96	98
	At	NM_113188	0.0	AAC49617	5e-163	81	91
	Mm	D64107	4e-16	BAA10969	5e-113	61	76
	Hs	CR456486	2e-19	CAG30372	4e-114	62	76
	Gg	XM_425477	2e-30	XP_425477	4e-114	61	76
	Sc	YER179W	2e-17	NP_01106	1e-96	53	72
*Ta*HOP2	*Os*	AK102972	0.0	BAG95812	1e-113	89	95
	At	AJ627484	4e-95	CAF28783	3e-90	69	87
	Mm	AB000121	7e-07	BAA23155	2e-32	40	58
	Hs	AB030304	5e-09	BAA92872	2e-29	39	57
	Sc	YGL033W	0.19	NP_011482	3e-06	21	46
*Ta*MND1	*Os*	NM_001069301	0.0	NP_001062766	4e-107	89	97
	At	NM_119061	5e-81	NP_194646	4e-87	71	85
	Mm	NM_029797	8e-07	NP_084073	8e-40	41	65
	Hs	BC032142	6e-08	AAH32142	2e-39	41	64
	Sc	YGL183C	0.19	NP_011332	1e-10	27	47

3.3.2 – Comparative protein analysis

3.3.2.1 – Sequence conservation

Conserved domain (cds) analysis revealed a number of expected conserved domains within the predicted amino acid sequences encoded by each of the wheat candidates (Table 3.5). The *Ta*DMC1 protein (344 aa) contains two overlapping protein domains, recomb_DMC1 and Rad51_DMC1_radA; along with a number of small motifs, including ATP-binding sites and the Walker A and B motifs (Figure 3.5 A). All three of the domains determined by the NCBI conserved protein domain database comprise the same sequences yet have been termed as three distinct domains due to being found in three distinct previously studied accessions. The *Ta*HOP2 (227 aa) only has one conserved domain, the TBPIP protein domain (Figure 3.5 B); while the *Ta*MND1 protein (207 aa) contains two protein domains, the MND1 domain and the COG5124 domain (Figure 3.5 C).

Table 3.5 - Summary of conserved domains within *Ta*DMC1, *Ta*HOP2 and *Ta*MND1 predicted amino acid sequences.

Protein	Conserved domain(s)	Domain Length (aa)	E-value
***Ta*DMC1**	recomb_DMC1 Rad51_DMC1_radA	313 235	2e-156 7e-96
***Ta*HOP2**	Tat binding protein 1(TBP-1)-interacting protein (TBPIP)	169	4e-33
***Ta*MND1**	Mnd1 COG5124	188 209	3e-46 6e-17

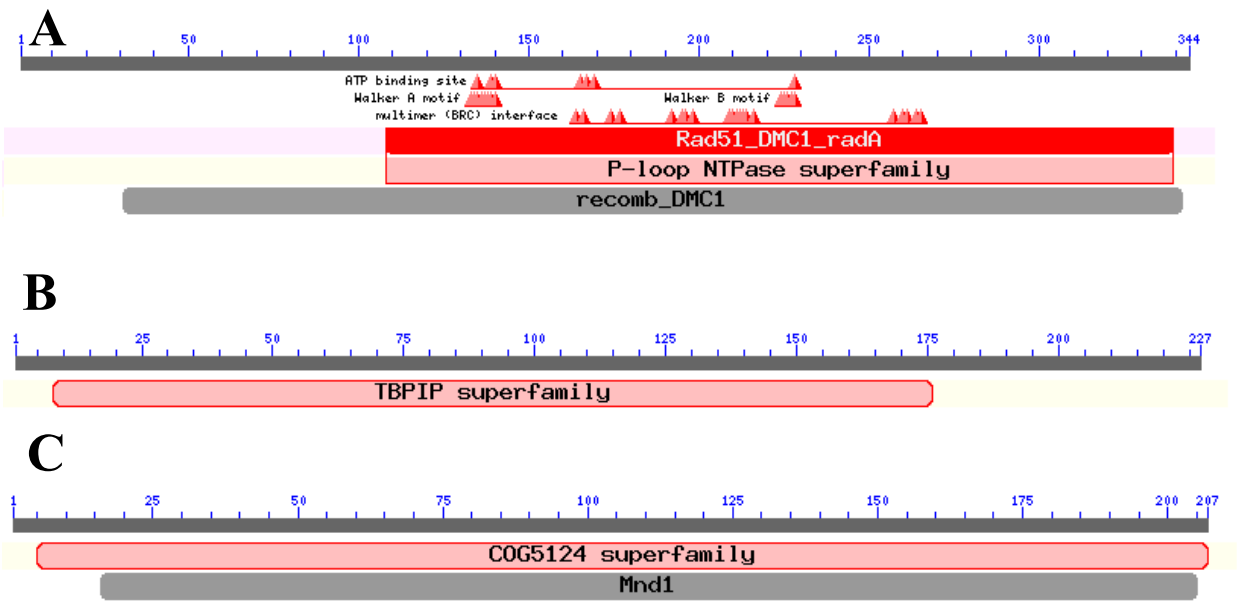


Figure 3.5 - Conserved domains of the predicted amino acid sequences of *Ta*DMC1, *Ta*HOP2 and *Ta*MND1. Within the 344 amino acids (aa) of *Ta*DMC1 (**A**) there are two overlapping protein domains, recomb_DMC1 (311 aa) and Rad51_DMC1_radA (232 aa). Also located in the protein are ATP-binding sites, Walker A and B motifs, and multimer BRC interfaces. *Ta*HOP2 (227 aa) has only one conserved domain, which belongs to the TBPIP superfamily (**B**). *Ta*MND1 contains an MND1 (189 aa) and a COG5124 (205 aa) domain (**C**). Figure generated using NCBI conserved domain predictor (**http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi**).

Aligning the orthologous proteins enabled comparison along the whole length of the sequence. *Ta*DMC1 shows a high level of amino acid conservation (Figure 3.6), with the Walker A and B motifs and the seven ATP binding sites being near identical across all eukaryotes represented in the analysis. However, in comparing both *Ta*HOP2 (Figure 3.7) and *Ta*MND1 (Figure 3.8) across other eukaryotes, the identity and similarity is reduced.

A

```
TaDMC1   MAPSKQYDEGGQLQLMEADRVEEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMMHTKKSLTGIK 70
OsDMC1   MAPSKQYDEGGQLQLMDAERIEEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMMHTKKSLTGIK 70
AtDMC1   MMASLKAEETSQMQLVEREENDEDEDLFEMIDKLIAQGINAGDVKKLQEAGIHTCNGLMMHTKKNLTGIK 70
MmDMC1   MK-------EDQVVQEESGFQDDEESLFQDIDLLQKHGINMADIKKLKSVGICTIKGIQMTTRRALCNVK 63
HsDMC1   MK-------EDQVVAEEPGFQDEEESLFQDIDLLQKHGINVADIKKLKSVGICTIKGIQMTTRRALCNVK 63
GgDMC1   MKAM-----EDQVVQEESGYHDDEESFFQDIDLLQKHGINVADIKKLKSVGICTIKGVQMTTRRALCNVK 65
ScDMC1   ------------MSVTGTEIDSDTAKNILSVDELQNYGINASDLQKLKSGGIYTVNTVLSTTRRHLCKIK 58

TaDMC1   GLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVVRITTGSQTLDELLGGGIETLCITEAFGEFRSGKT 140
OsDMC1   GLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVVRITTGSQALDELLGGGIETLCITEAFGEFRSGKT 140
AtDMC1   GLSEAKVDKICEAAEKIVNFGYMTGSDALIKRKSVVKITTGCQALDDLLGGGIETSAITEAFGEFRSGKT 140
MmDMC1   GLSEAKVEKIKEAANKLIEPGFLTAFQYSERRKMVFHITTGSQEFDKLLGGGIESMAITEAFGEFRTGKT 133
HsDMC1   GLSEAKVDKIKEAANKLIEPGFLTAFEYSEKRKMVFHITTGSQEFDKLLGGGIESMAITEAFGEFRTGKT 133
GgDMC1   GLSEVKVDKIKEAANKLIEPGFLTAFEYSEKRKMVFHITTGSQEFDKLLGGGIESMAITEAFGEFRTGKT 135
ScDMC1   GLSEVKVEKIKEAAGKIIQVGFIPATVQLDIRQRVYSLSTGSKQLDSILGGGIMTMSITEVFGEFRCGKT 128

TaDMC1   QLAHTLCVSTQLPLHMHGGNGKVAYIDTEGTFRPERIVPIAERFGMDANAVLDNIIYARAYTYEHQYNLL 210
OsDMC1   QLAHTLCVSTQLPIHMHGGNGKVAYIDTEGTFRPERIVPIAERFGMDANAVLDNIIYARAYTYEHQYNLL 210
AtDMC1   QLAHTLCVTTQLPTNMKGGNGKVAYIDTEGTFRPDRIVPIAERFGMDPGAVLDNIIYARAYTYEHQYNLL 210
MmDMC1   QLSHTLCVTAQLPGTGGYSGGKIIFIDTENTFRPDRLRDIADRFNVDHEAVLDNVLYARAYTSEHQMELL 203
HsDMC1   QLSHTLCVTAQLPGAGGYPGGKIIFIDTENTFRPDRLRDIADRFNVDHDAVLDNVLYARAYTSEHQMELL 203
GgDMC1   QLSHTLCVTAQLPGPKGYTGGKIIFIDTENTFRPDRLRDIADRFNVDHDAVLDNVLYARAYTSEHQMELL 205
ScDMC1   QMSHTLCVTTQLPREMGGGEGKVAYIDTEGTFRPERIKQIAEGYELDPESCLANVSYARALNSEHQMELV 198

TaDMC1   LGLAAKMAEEP--FRLLIVDSVIALFRVDFSGRGELAERQQKLAQMLSRLTKIAEEFNVAVYITNQVIAD 278
OsDMC1   LGLAAKMAEEP--FRPLIVDSVIALFRVDFSGRGELAERQQKLAQMLSRLTKIAEEFNVAVYITNQVIAD 278
AtDMC1   LGLAAKMSEEP--FRILIVDSIIALFRVDFTGRGELADRQQKLAQMLSRLIKIAEEFNVAVYMTNQVIAD 278
MmDMC1   DYVAAKFHEEAGIFKLLIIDSIMALFRVDFSGRGELAERQQKLAQMLSRLQKISEEYNVAVFVTNQMTAD 273
HsDMC1   DYVAAKFHEEAGIFKLLIIDSIMALFRVDFSGRGELAERQQKLAQMLSRLQKISEEYNVAVFVTNQMTAD 273
GgDMC1   DYVAAKFHEEAGIFKLLIIDSIMALFRVDFSGRGELAERQQKLAQMLSRLQKISEEYNVAVFVTNQMTAD 275
ScDMC1   EQLGEELSSGD--YRLIVVDSIMANFRVDYCGRGELSERQQKLNQHLFKLNRLAEEFNVAVFLTNQVQSD 266

TaDMC1   PGGGM-FITDP-KKPAGGHVLAHAATIRLMLRKGKGEQRVCKIFDAPNLPEGEAVFQITTGGLMDVKD 344
OsDMC1   PGGGM-FITDL-KKPAGGHVLAHAATIRLMLRKGKGEQRVCKIFDAPNLPEGEAVFQVTSGGIMDAKD 344
AtDMC1   PGGGM-FISDP-KKPAGGHVLAHAATIRLLFRKGKGDTRVCKVYDAPNLAEAEASFQITQGGIADAKD 344
MmDMC1   PGATMTFQADP-KKPIGGHILAHASTTRISLRKGRGELRIAKIYDSPEMPENEATFAITAGGIGDAKE 340
HsDMC1   PGATMTFQADP-KKPIGGHILAHASTTRISLRKGRGELRIAKIYDSPEMPENEATFAITAGGIGDAKE 340
GgDMC1   PGATMTFQADP-KKPVGGHILAHASTTRISLRKGRGELRIAKIYDSPEMPENEATFAITPGGIGDAKK 342
ScDMC1   PGASALFASADGRKPIGGHVLAHASATRILLRKGRGDERVAKLQDSPDMPEKECVYVIGEKGITDSSD 334
```

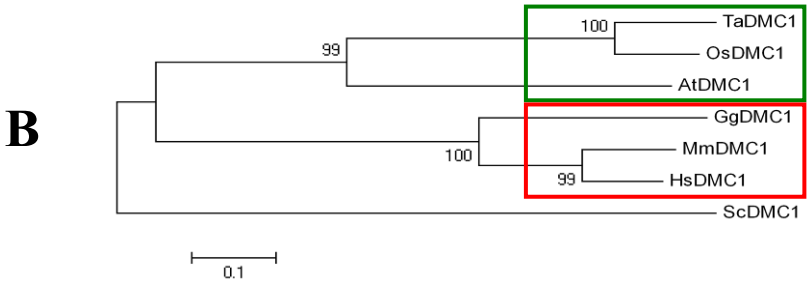


Figure 3.6 - Amino acid conservation of *Ta*DMC1. (**A**) Amino acid sequence alignment shows high levels of conservation within the recom_DMC1 (*grey* line), RAD51_DMC1_recA (*orange* line), the Walker A (*blue* line) and B (*green* line) motifs and the seven ATP binding sites (*red* lines). (**B**) Phylogenetic analysis of DMC1 shows the distinct clades between the animal (*red* box) and plant (*green* box) kingdoms. Neighbour-joining percent bootstrap values are presented for each clade while the scale bar indicates the number of amino acid substitutions per site. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

A

```
TaHOP2   -----------MPPKSDSVEGIVLNFVNEQNRPLNSQNVADALQKFS-LKKTAVQKGLDALADSGQISFK 58
OsHOP2   -----------MPPKSDSVEGIVLSFVNEQNRPLNSQNVADALQKFS-LKKTAVQKALDALADSGQISFK 58
AtHOP2   -----------MAPKSDNTEAIVLNFVNEQNKPLNTQNAADALQKFN-LKKTAVQKALDSLADAGKITFK 58
MmHOP2   -------MSKSRAEAAAGAPGIILRYLQEQNRPYSAQDVFGNLQKEHGLGKAAVVKALDQLAQEGKIKEK 63
HsHOP2   -------MSKGRAEAAAGAAGILLRYLQEQNRPYSSQDVFGNLQREHGLGKAVVVKTLEQLAQQGKIKEK 63
ScHOP2   MAPKKKSNDRAIQAKGSEAEQLIEDYLVSQYKPFSVNDIVQNLHNKV--TKTTATKALENLVNEKRIVSK 68

TaHOP2   EYGKQKIYLARQDQFDIPNGEELEEMK-KANSKLQEELAEQKKITSEVESEIKSLQSNLTLEEIRSKEAR 127
OsHOP2   EYGKQKIYLARQDQFNIPNGEELEEMK-KANIKLQEELADQKKAISEVESEVRGLQSNLTLAEIKSKEAK 127
AtHOP2   EYGKQKIYIARQDQFEIPNSEELAQMK-EDNAKLQEQLQEKKKTISDVESEIKSLQSNLTLEEIQEKDAK 127
MmHOP2   TYGKQKIYFADQNQFDTVSDADLHGLD-ASIVALTAKVQSLQQSCRHMEAELKELTSALTTPEMQKEIQE 132
HsHOP2   MYGKQKIYFADQDQFDMVSDADLQVLD-GKIVALTAKVQSLQQSCRYMEAELKELSSALTTPEMQKEIQE 132
ScHOP2   TFGKIIIYSCNEQDTALPSNIDPSQFDFETVLQLRNDLIELERDKSTAKDALDSVTKEPENEDLLTIIEN 138

TaHOP2   LQSEVQEMEEKLKKLQSGVILVKPEDKKIIEETFSEKANQWRKRKRMFKELWDNITE-NSPKDQKEFKEE 196
OsHOP2   LQREVQEMEEKLNKLRNGVILVKPEDKKIIEESFSEKVNQWRKRKRMFKELWDNITE-NSPKDQKEFKEE 196
AtHOP2   LRKEVKEMEEKLVKLREGITLVRPEDKKAVEDMYADKINQWRKRKRMFRDIWDTVTE-NSPKDVKELKEE 196
MmHOP2   LKKECAQYTERLKNIKAATNHVTPEEKEKVYRDRQKYCKEWRKRKRMTTELCDAILE-GYPKSKKQFFEE 201
HsHOP2   LKKECAGYRERLKNIKAATNHVTPEEKEQVYRERQKYCKEWRKRKRMATELSDAILE-GYPKSKKQFFEE 201
ScHOP2   EENELKKIESKLQSLQDDWDPANDEIVKRIMSEDTLLQKEITKRSKICKNLIATIKDSVCPKNMNEFLVC 208

TaHOP2   LGLEYDEDVGVNFQPYSEMLASLNKRRKVSR 227
OsHOP2   LGLEYDEDVGVNLQSYSDMLTSLSKRRKVSR 227
AtHOP2   LGIEYDEDVGLSFQAYADLIQHGKKRPRGQ- 226
MmHOP2   VGIETDEDHNVLLPDP--------------- 217
HsHOP2   VGIETDEDYNVTLPDP--------------- 217
ScHOP2   ILNIFRDLFF------------------- 218
```

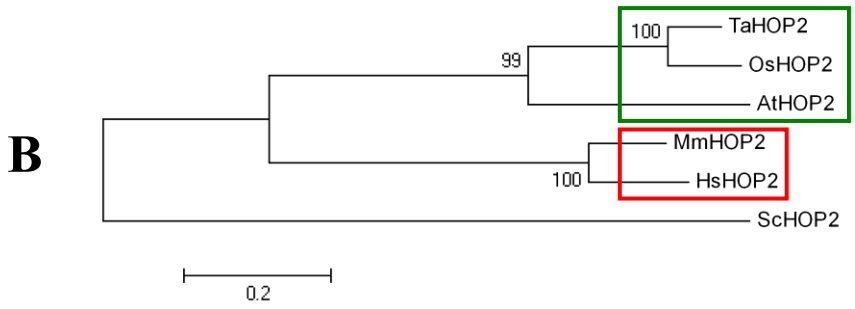


Figure 3.7 - Amino acid conservation of *Ta*HOP2. (**A**) The HOP2 orthologues show reduced sequence conservation, especially when compared to the yeast (*Sc*) protein. The purple line represents the TBPIP protein conserved domain. (**B**) Phylogenetic analysis of HOP2 shows the distinct clades between the animal (*red* box) and plant (*green* box) kingdoms. Neighbour-joining percent bootstrap values are presented for each clade while the scale bar indicates the number of amino acid substitutions per site. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

A

```
TaMND1    -MSKKRGLSLEEKWEQMLQIFYESQDFYLLKELEKMGPK-KGVISQSVKDVVQSLVD-DDLVLRDKIGTS 67
OsMND1    -MSKKRGLSLEEKREQMLQIFYDSQDFYLLKELEKLGPK-KGVISQSVKDVVQSLVD-DDLVLKDKIGTS 67
AtMND1    -MSKKRGLSLEEKREKMLQIFYESQDFFLLKELEKMGPK-KGVISQSVKDVIQSLVD-DDLVAKDKIGIS 67
MmMND1    -MSKKRGLSGEEKRTRMMEIFFETKDVFQLKDLEKLAPKEKGITAMSVKEVLQSLVD-DGMVDCERIGTS 68
HsMND1    -MSKKKGLSAEEKRTRMMEIFSETKDVFQLKDLEKIAPKEKGITAMSVKEVLQSLVD-DGMVDCERIGTS 68
ScMND1    MGPKRQTVSLQEKKNRILNFFQETYTFYNIKELEKSIPKKCGISPMIVKDLVQQMIDEDGVISVEKCGNI 70

TaMND1    VYFWSLPSCAGNQLRTTYNKLESDLSNSKKRYMELLEQRDD-LKRGR-------EDTDEREDALEELKAV 129
OsMND1    VYFWSLPSCAGNQLRTTYSKLESDLSSSKKRFIELVEQREN-LKRGR-------EDSDEREAALEELKAV 129
AtMND1    IYFWSLPSCAGNQLRSVRQKLESDLQGSNKRLAELVDQCEA-LKKGR-------EESEERTEALTQLKDI 129
MmMND1    NYYWAFPSKALHARKRKLEALNSQLSEGSQKHADLQKSIEK-ARVGR-------QETEERAMLAKELSSF 130
HsMND1    NYYWAFPSKALHARKHKLEVLESQLSEGSQKHASLQKSIEK-AKIGR-------CETEERTRLAKELSSL 130
ScMND1    NIYWCFKNQTLQKLYDSSELIKKKIQEVKCDIATYKQELDKTLATGRRKKFTVGQKSYNREALLEKRKKI 140

TaMND1    ELRHKKLKEELAAYADS--DPSALEAMKDATEVAHSAASRWTDNIFTLQQWCSTTFPQAKEQLEHMYKEV 197
OsMND1    EQHHKKLKEELAAYADS--DPAALEAMNDAIEVAHAAANRWTDNIFTLQQWCSTTFPQAKEQLEHMYREV 197
AtMND1    EKKHKDLKNEMVQFADN--DPATLEAKRNAIEVAHQSANRWTDNIFTLRQWCSNNFPQAKEQLEHLYTEA 197
MmMND1    RDQRQQLKAEVEKYREC--DPQVVEEIREANKVAKEAANRWTDNIFAIKSWAKRKFGFEESKIDKNF--- 195
HsMND1    RDQREQLKAEVEKYKDC--DPQVVEEIRQANKVAKEAANRWTDNIFAIKSWAKRKFGFEENKIDRTF--- 195
ScMND1    QDEIKKKSNSLQKIESIRWDAAKIQENKQQIRLKKVHLEKTTDNIEILIDYLYKKFFLKPEQIRKEF--- 207

TaMND1    GITEDFKYLQ---------------------- 207
OsMND1    GITEDFEYLQ---------------------- 207
AtMND1    GITEDFDYIELSSFPLSSSHEADTAKQLVQDEA 230
MmMND1    GIPEDFDYID---------------------- 205
HsMND1    GIPEDFDYID---------------------- 205
ScMND1    GIPEEFKEFTEV-------------------- 219
```

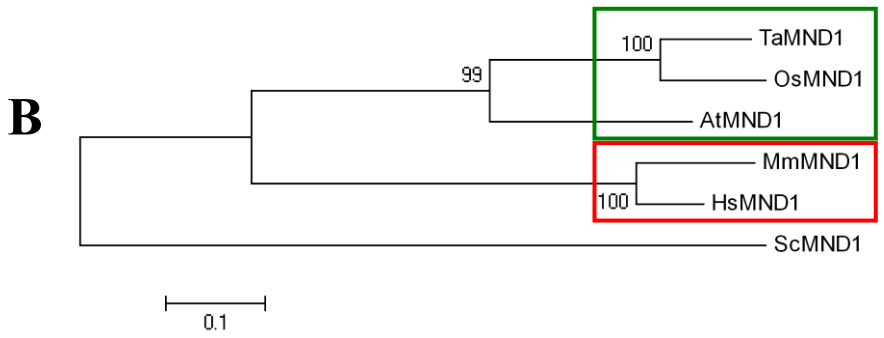


Figure 3.8 - Amino acid conservation of *Ta*MND1. (**A**) The MND1 orthologues show high conservation across the length of the MND1 protein domain. There is a noticeable amino acid divergence between the monocots (rice and wheat) and the dicot Arabidopsis, especially at the C-terminus end of the protein. The pink line represents the MND1 protein domain. (**B**) Phylogenetic analysis of MND1 shows the distinct clades between the animal (*red* box) and plant (*green* box) kingdoms. Neighbour-joining percent bootstrap values are presented for each clade while the scale bar indicates the number of amino acid substitutions per site. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

3.3.2.2 – Three dimensional modelling of DMC1 proteins

The 3D modelling of the protein sequence from *Ta*DMC1 spanning the amino acids 91-344 revealed nine α-helices and eight β-sheet secondary structures, with a chain of non-structured amino acids (of approximately 10 amino acids) residing before the last α-helix and three β-sheets (Figure 3.9 A). Superimposing the 3D protein models from rice (91-344, Figure 3.9 B), mouse (83-340, Figure 3.9 C) and yeast (16-331, Figure 3.9 D) revealed significant similarity to the wheat structure, both at the secondary and tertiary level (with the exception of the non-structured amino acid chain). The yeast DMC1 protein model revealed additional structural features towards the N-terminus, such that when superimposed over the wheat protein there was some variance at the tertiary level (Figure 3.9 D). However, no change was observed with the secondary structures (Figure 3.9 D).

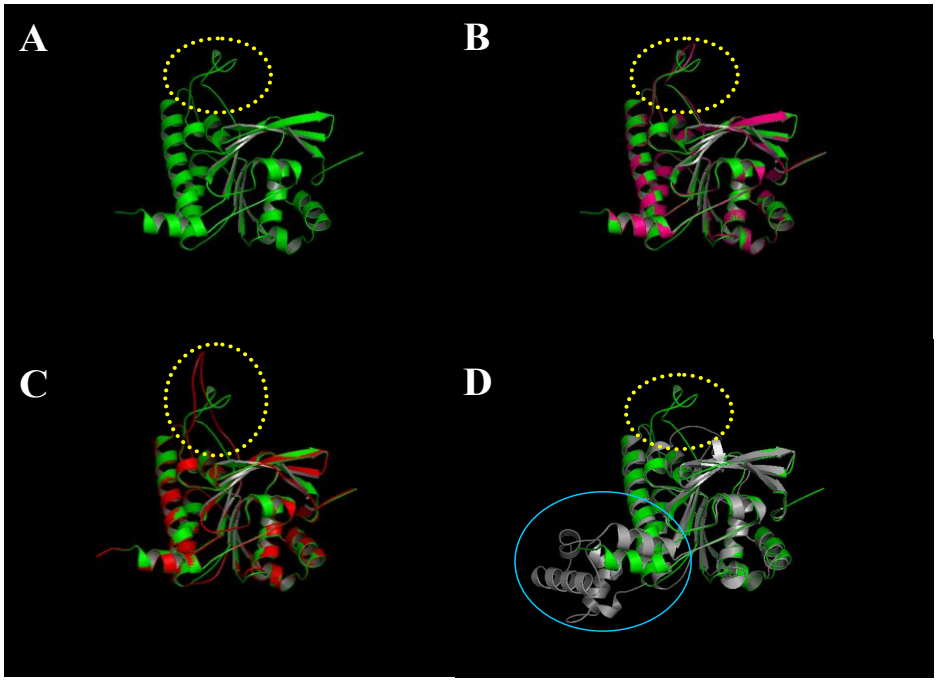


Figure 3.9 - Three dimensional structure of DMC1. (**A**) Computational modelling of the C-terminus of the wheat DMC1 protein (*green*) with α-helical and β-sheet secondary structures and a non-structured amino acid chain (*yellow* circle). Superimposing the models from rice (**B**, *pink*), mouse (**C**, *red*) and yeast (**D**, *grey*) with the wheat protein (*green* throughout the figure) confirms that there is a high level of conservation of both the secondary and tertiary structure across divergent species. The only exception is the short non-structured chain of amino acids (*yellow* circles) which differ in both length and conformation between the models, and the angular conformation (**D**, *blue* circle) of the secondary structure between the wheat and yeast proteins.

3.3.3 – Chromosome locations of the wheat homologues of *DMC1*, *HOP2* and *MND1*

The chromosome location of the three candidates was determined using Southern blot analysis with nullisomic-tetrasomic membranes (Figure 3.10). The absence of three hybridisation signals across the membranes illustrates the location of the gene candidates. Three copies of both *TaDMC1* and *TaMND1* were detected on chromosome group 5 (Figures 3.10 A and B) while three copies of *TaHOP2* were found on chromosome group 4 (Figure 3.10 C).

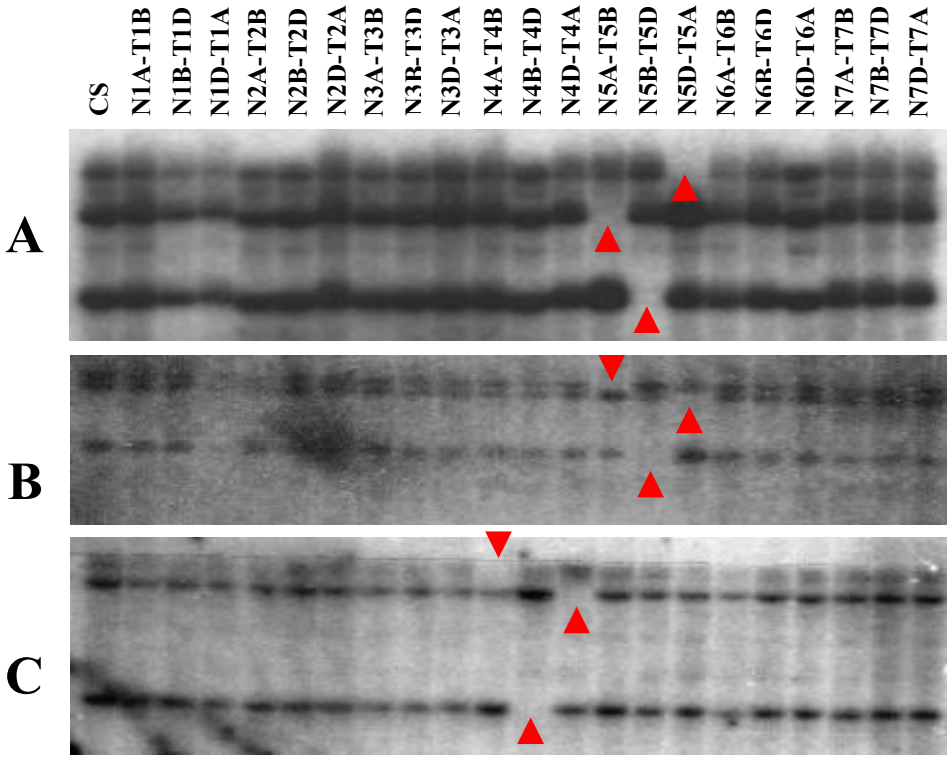


Figure 3.10 - Chromosome locations of *TaDMC1*, *TaHOP2* and *TaMND1*. Absence of hybridisation signals (*red* arrows) using Southern analysis revealed that both *TaDMC1* (**A**) and *TaMND1* (**B**) are on chromosome group 5, while *TaHOP2* is located on chromosome group 4 (**C**). Each of these genes has a copy in each of the A, B and D genomes. CS = wild-type Chinese Spring.

3.3.4 – Expression analysis of meiotic candidates

TaHOP2 expression was highest during pre-meiosis, with only a small reduction in transcript expression being observed during the later stages of meiosis and in root tissues. No expression was observed in leaf tissues (Figure 3.11). *TaMND1* was observed to have the highest expression in leaf tissue, with moderate levels of expression detected across all other tissues examined except for root tips which had significantly lower levels (Figure 3.11). As these two genes are known to form a protein complex, it was not surprising that they are also co-expressed (correlation of r = 0.98).

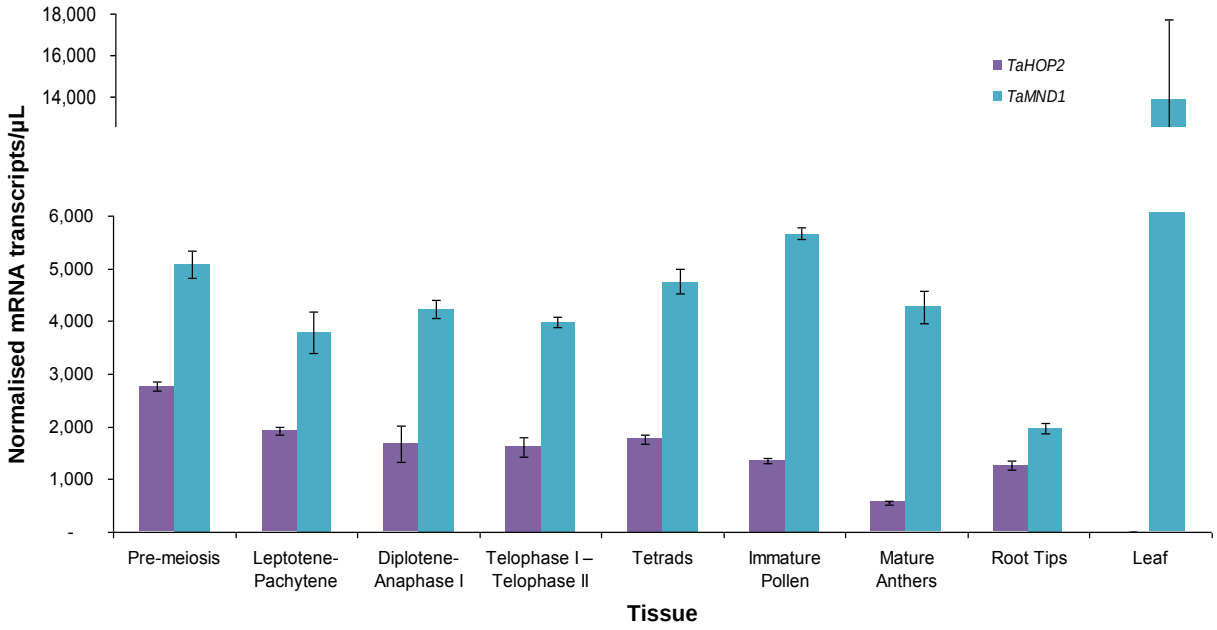


Figure 3.11 - Transcript expression of *TaHOP2* and *TaMND1*. *TaHOP2* has the highest level of expression during the earliest stages of meiosis examined, while *TaMND1* has high levels of expression in vegetative tissues (relative to the expression detected in the meiotic tissues examined). Despite this, the two genes are found to be co-expressed (r = 0.98).

As the HOP2:MND1 protein complex is known to work with DMC1, co-expression of the three genes in wheat was investigated during meiosis (Figure 3.12). *TaHOP2* and *TaMND1* are shown be expressed at much lower levels than *TaDMC1* during meiosis (stages PM-T) and into immature pollen, with the exception of mature anthers where *TaMND1* expression is highest. While not as significant as the association between *TaHOP2* and *TaMND1*, *TaDMC1* is co-expressed with *TaHOP2* (r = 0.89) and *TaMND1* (r = 0.84). Based on separate expression analysis (Able, unpublished; Khoo *et al*., 2008) the genes of *TaHOP2* and *TaMND1* are also observed to have correlated expression ($r \geq 0.95$) with the wheat genes *ASY1*, *PHS1*, *RPA*, *MSH6* and *RAD51C*, with *TaMND1* also having correlated expression with *TaRAD51D*.

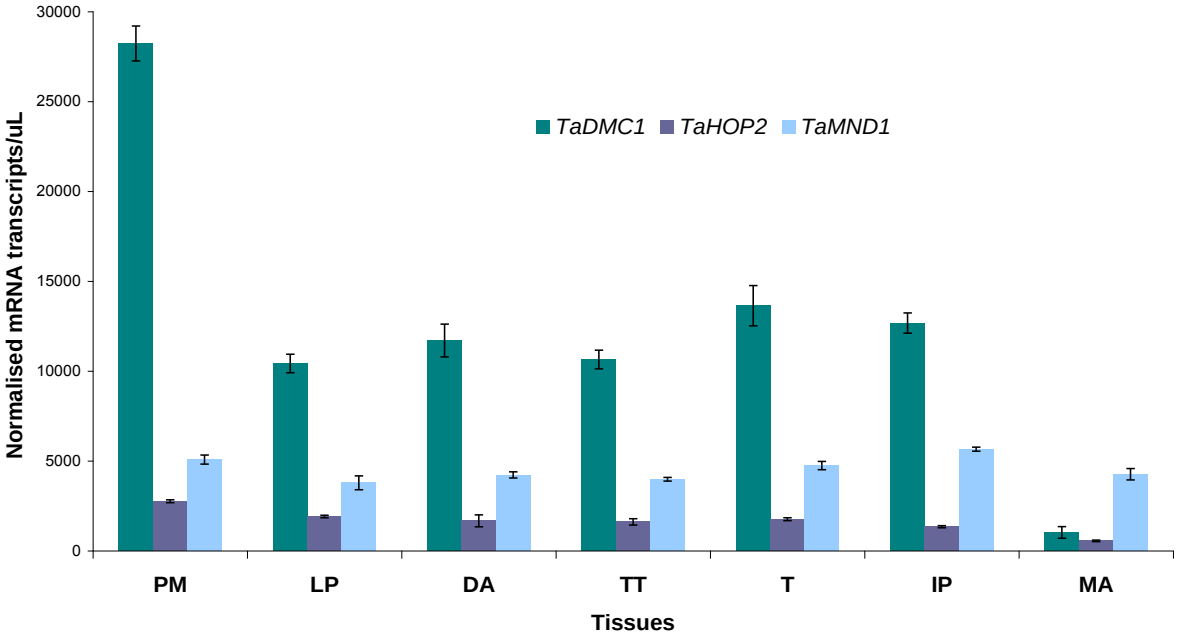


Figure 3.12 - Expression of *TaDMC1*, *TaHOP2* and *TaMND1* during seven stages of wheat reproductive development. *TaHOP2* and *TaMND1* were expressed to a lesser extent than *TaDMC1*, with a reduced correlation of expression (r = 0.89 and 0.84 for *TaHOP2* and *TaMND1* respectively, when compared to the value obtained between *TaHOP2* and *TaMND1*). PM = pre-meiosis, LP = Leptotene-Pachytene, DA = Diplotene-Anaphase I, TT = Telophase I-Telophase II, T = Tetrads, IP = Immature Pollen, MA = Mature Anthers.

3.3.5 – Protein localisation

3.3.5.1 – Immunolocalisation of *Ta*DMC1 and *Ta*MND1 using meiocytes embedded in polyacrylamide

Temporal and spatial localisation of *Ta*DMC1 and *Ta*MND1 was conducted with the antibodies raised against the Arabidopsis DMC1 and MND1 proteins. Using a 4% paraformaldehyde fixative for the meiocytes with the antibody combinations of anti-*At*DMC1(Rabbit)/anti-*At*ASY1(Rat) and anti-*At*MND1(Rat)/anti-*Ta*ASY1(Rabbit), detectable *Ta*ASY1 signal was observed while no DMC1 (above that of background) was seen (Figure 3.13). The MND1 antibody resulted in a large amount of signal, yet surprisingly this appeared to be localised only to the cytoplasm (Figure 3.13).

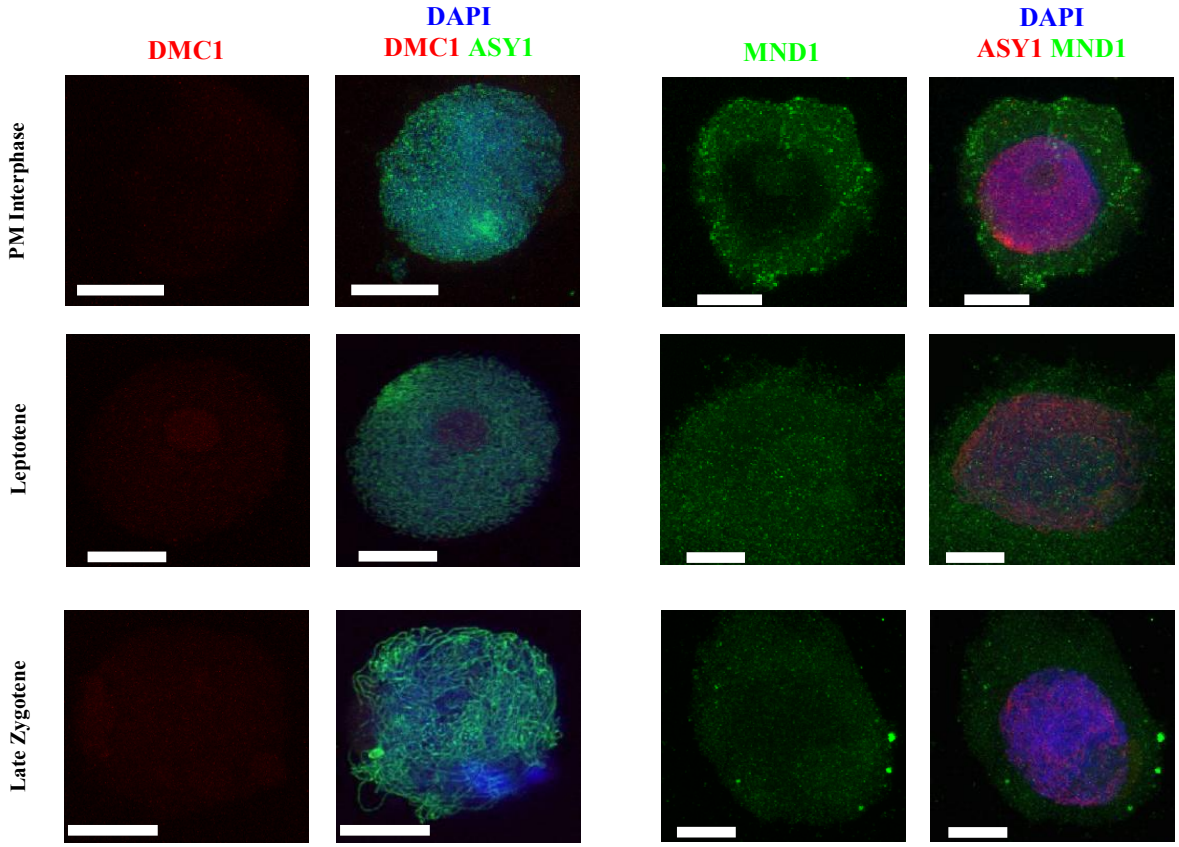


Figure 3.13 - Immunolocalisation of *Ta*DMC1 and *Ta*MND1 in 4% fixed meiocytes. *Ta*ASY1 protein signal can be visualised in green (*left*) and red (*right*), whereas the DMC signal in red (*left*) cannot be visualised above what appears to be background levels. A MND1 protein signal can be seen (*green*, *far-right panels*), however the signal appears to be only present in the cytoplasm surrounding the DNA (stained *blue*, using DAPI). Scale bars = 10 µm.

A second attempt at immunolocalisation of the *Ta*MND1 protein was conducted with fresh meiocytes embedded using 2% fixative, and with an increased permabilisation time of 2 hours. This again resulted in detectable *Ta*ASY1 signal but no observed MND1 signal above what would be classed as background (Figure 3.14). However, during pre-meiosis there does appear to be some signal present within the area of the nucleolus (Figure 3.14, arrows).

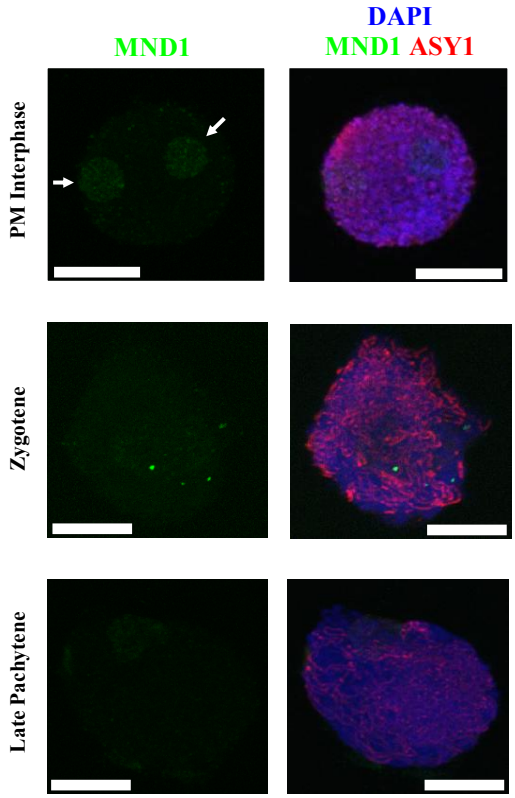


Figure 3.14 - Immunolocalisation using the *At*MND1 antibody in 2% paraformaldehyde fixed wheat meiocytes. The *Ta*ASY1 protein can be visualised in red, while the MND1 signal (*green*) is not detectable above what would be considered as background, with the exception in pre-meiosis where a signal is persisting in the area of the nucleolus (*arrows*). DNA is stained blue with DAPI. Scale bars = 10 µm.

3.3.5.2 – Immunolocalisation in meiotic chromosome spreads

Given the limited success with antibodies (other than *Ta*ASY1) using the 3-D polyacrylamide pad technique, meiotic chromosome spread immunolocalisations were attempted with the *At*DMC1 and *At*MND1 antibodies. Chromosome spreads with the *At*DMC1 antibody were also unsuccessful; with only the *Ta*ASY1 protein signal being detected (Figure 3.15).

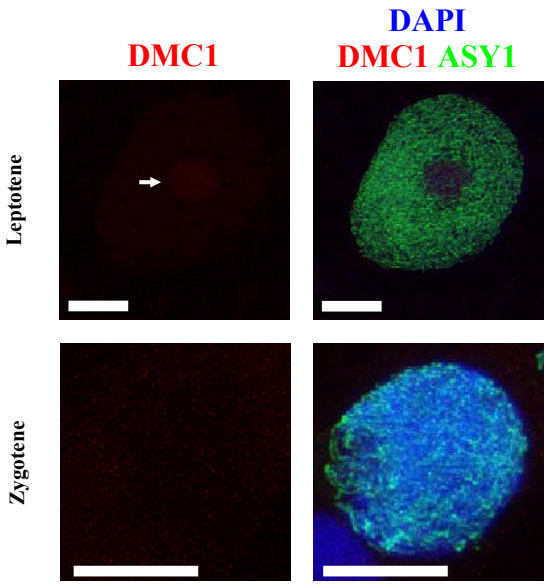


Figure 3.15 - Meiotic chromosome spread immunolocalisation with anti-*At*DMC1. *Ta*ASY1 protein signal (*green*) can be visualised, and is associated with the DNA (DAPI, *blue*). However, the DMC1 signal (*red*) cannot be seen above background levels except during leptotene where a faint signal appears to be located in the area of the nucleolus (*arrow*). Scale bars = 10 µm.

Attempts to localise the *Ta*MND1 protein using the anti-*At*MND1 antibody proved more successful. In Figure 3.17, the *Ta*MND1 protein signal could be detected during pre-meiotic (PM) interphase through to pachytene. Throughout the early meiotic time-course the MND1 signal is localised to the meiotic chromosomes, with foci more prevalent during pre-meiotic interphase. This signal diminishes and becomes more „concentrated/focussed" as the cells enter leptotene and zygotene (Figure 3.16). During the following stage of pachytene the MND1 signal was somewhat similar to that observed in pre-meiotic interphase.

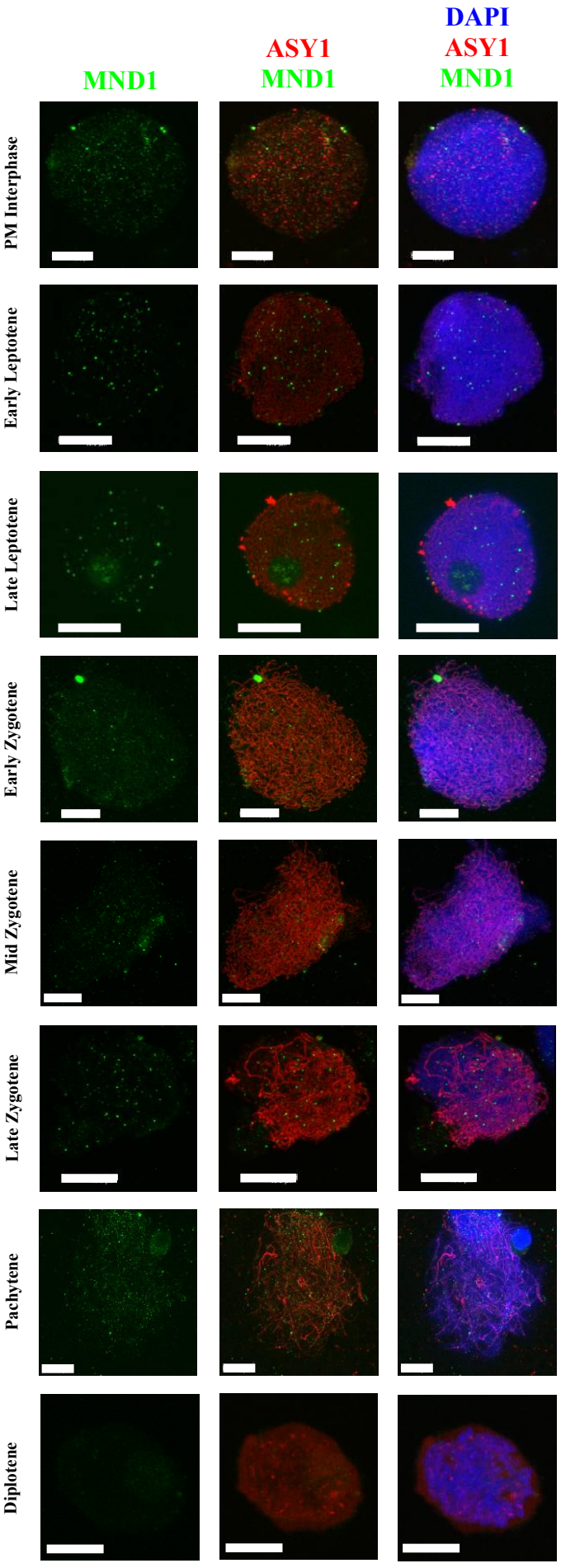
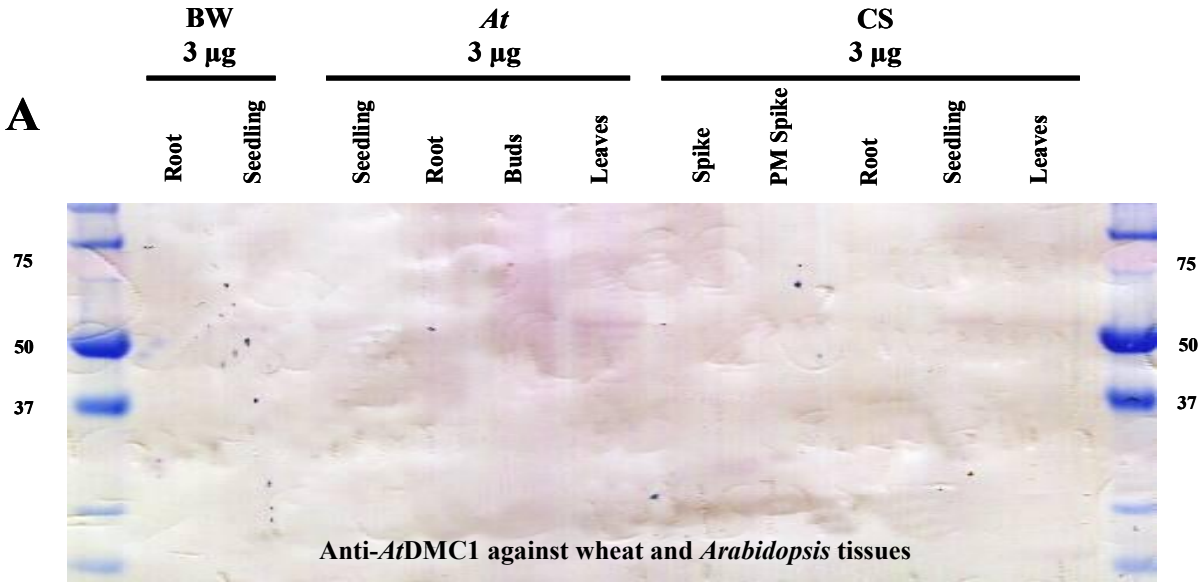
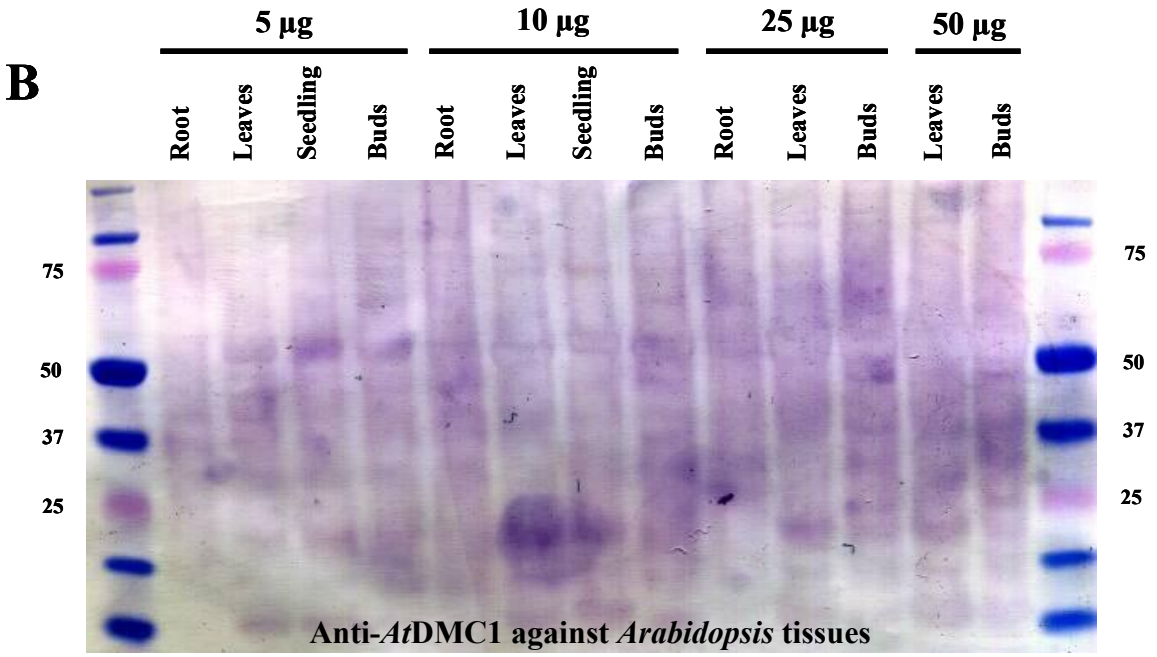

MND1
ASY1
MND1
DAPI
ASY1
MND1
PM Interphase
Early Leptotene
Late Leptotene
Early Zygotene
Mid Zygotene
Late Zygotene
Pachytene
Diplotene

Figure 3.16 - Temporal location of *Ta*MND1 during pre-meiotic interphase and prophase I sub-stages. *Ta*MND1 protein signal (*green*) localises to chromosomes (DAPI, *blue*) along with *Ta*ASY1 (*red*) during pre-meiotic (PM) interphase to pachytene. The numerous foci observed during PM interphase appear to diminish to become intense distinct foci as the meiocytes enter into leptotene and zygotene. Signal observed during the later prophase I stage of pachytene is similar to PM Interphase, followed by the reduction of „defined" signal for both *Ta*MND1 and *Ta*ASY1 by diplotene. Scale bars = 10 µm.

3.3.5.3 – Western analysis using the anti-*At*DMC1 antibody

Due to the unsuccessful immunolocalisation using the *At*DMC1 antibody the Arabidopsis antibody was tested by western analysis. Protein isolation was successfully conducted using the TCA protein extraction method. Western analysis was conducted with the anti-*At*DMC1 antibody, and also the anti-*Ta*ASY1 antibody as a positive control. This resulted in no signal at the expected size of 37 kDa where the DMC1 protein should be (Figure 3.17 A and B). The ASY1 positive control resulted in detecting the band corresponding to the ASY1 protein at the expected size of 67 kDa (Gel 1) (Figure 3.17 C).





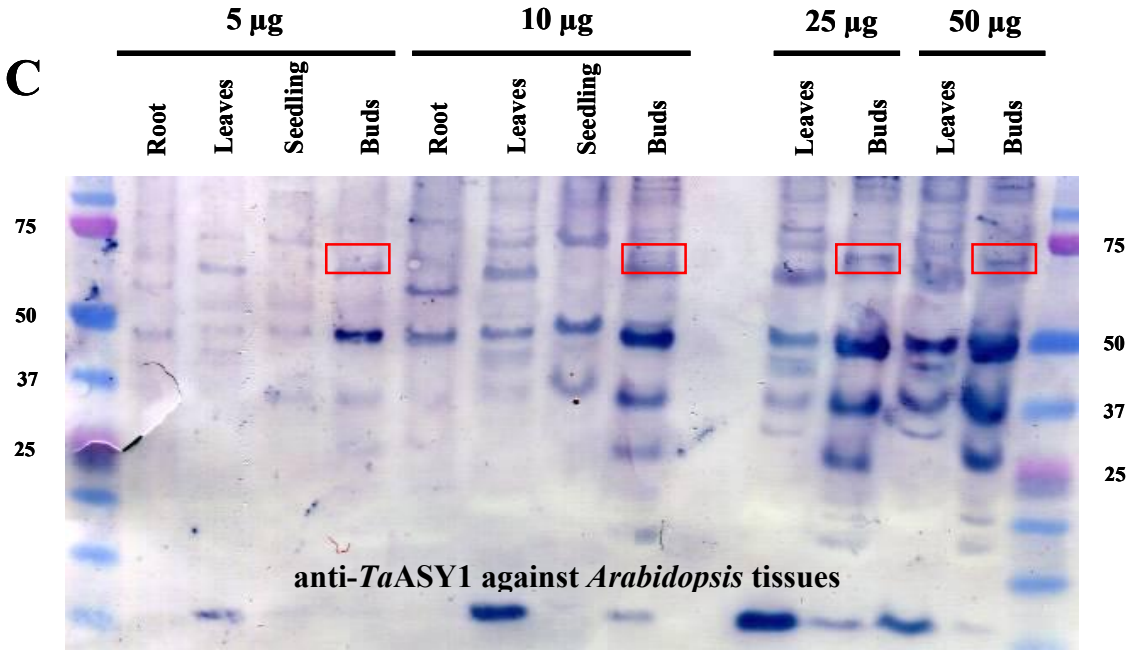


Figure 3.17 - Western analysis of the anti-*At*DMC1 antibody. Testing the anti-*At*DMC1 antibody against wheat and Arabidopsis (**A**), and Arabidopsis (**B**) tissues, resulted in no signal being observed in root, leaf, seedling and meiotic tissues even with high protein concentrations (e.g. 50 µg, **B**). The anti-*Ta*ASY1 antibody when probed against Arabidopsis tissues (**C**), resulted in detecting a band corresponding to the expected size of the *At*ASY1 protein (*red* boxes, **C**). *At* = *Arabidopsis thaliana*, BW = Bob white 26 and CS = Chinese spring. A BIO-RAD Precision Plus Dual Colour Protein Ladder was used; not all sizes are shown.

3.3.6 – Heterologous protein expression

3.3.6.1 – Protein expression vector preparation

3.3.6.1.1 – Isolation of the coding regions of the three meiotic genes for protein production

Using primers designed for the open reading frames, *TaDMC1*, *TaHOP2* and *TaMND1* were amplified (Figure 3.18). Identification of the open reading frames within the pCR®8/GW/TOPO® vector was conducted via colony PCR (Figure 3.19 A) with the correct orientation of the coding fragments verified by sequence analysis, as the M13F sequence primer corresponds to the vector region start codon for protein expression (Figure 3.19 B-D).

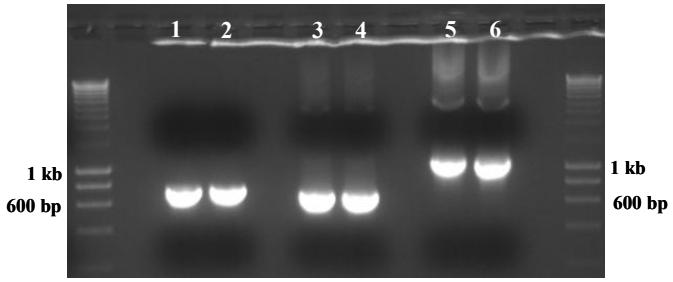
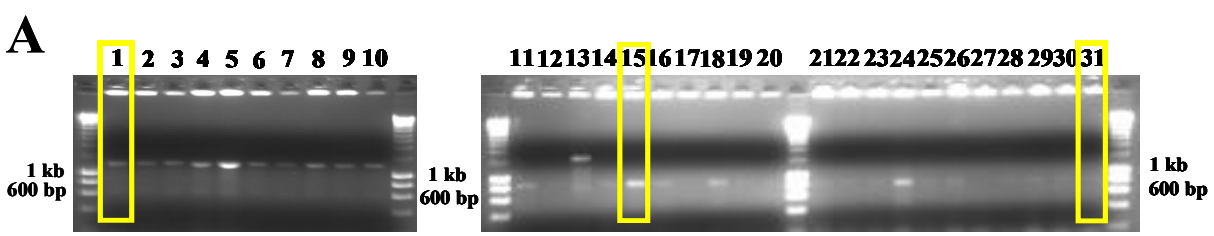


Figure 3.18 - PCR amplification of the open reading frames for *TaDMC1*, *TaHOP2* and *TaMND1*. Successful PCR duplicates of the *TaHOP2* (lanes 1 and 2), *TaMND1* (lanes 3 and 4) and *TaDMC1* (lanes 5 and 6) open reading frames were obtained. A 1 kb DNA molecular weight marker is shown at each end of the gel; however not all band sizes are listed.



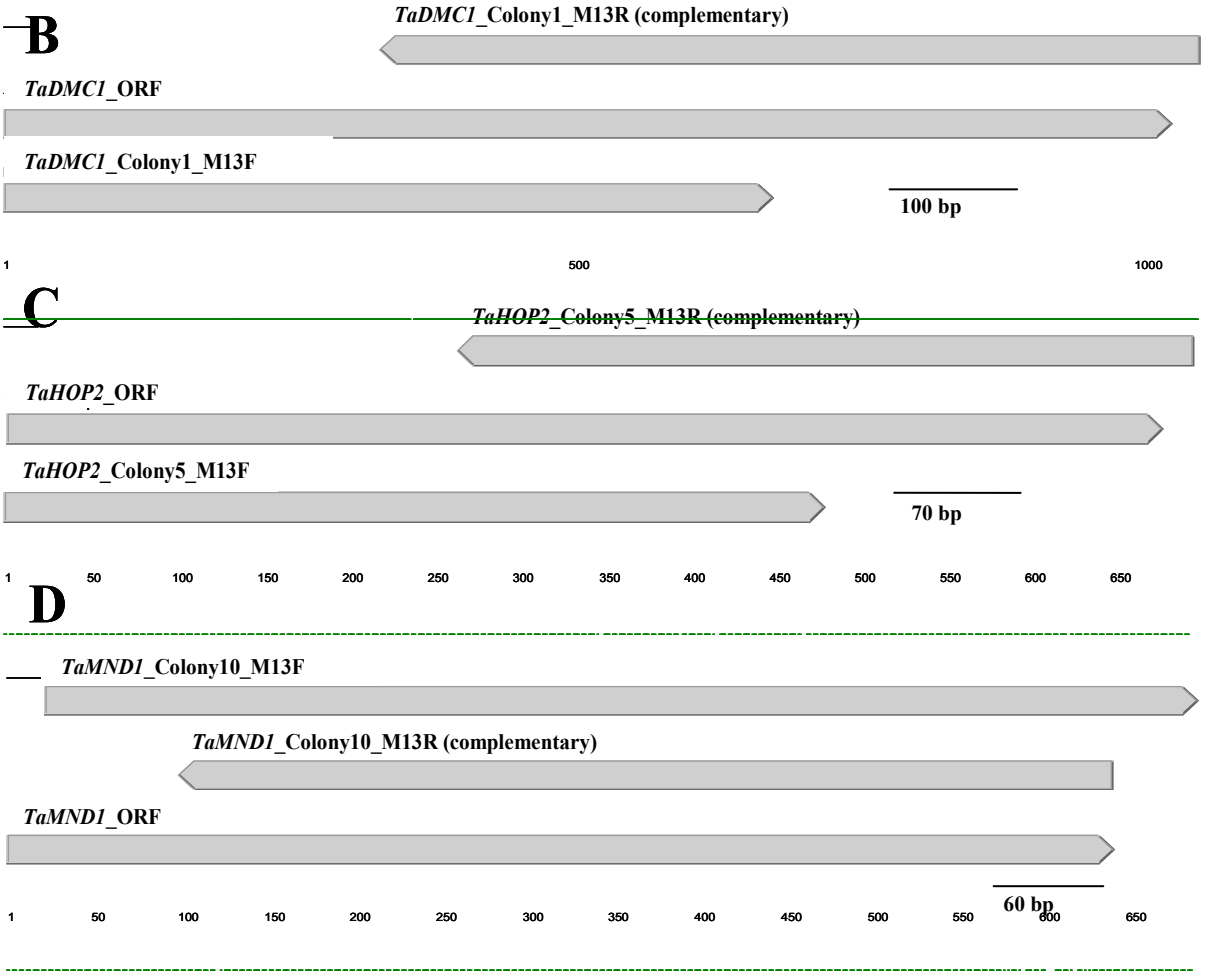


Figure 3.19 - Verification of the *TaDMC1*, *TaHOP2* and *TaMND1* open reading frames in the pCR®8/GW/TOPO® vector. (**A**) Colonies outlined in yellow boxes had plasmid DNA isolated and were sequenced. (**B-D**) M13F and M13R sequencing results are aligned with previously sequenced wheat candidate open reading frames. Alignment of sequence runs with the previously obtained open reading frames of each of the meiotic candidates *TaDMC1* (**B**), *TaHOP2* (**C**) and *TaMND1* (**D**). In (**A**), a 1 kb DNA molecular weight marker is shown, however not all band sizes are listed.

3.3.6.1.2 - Correct orientation of the open reading frames in the pDEST17® protein expression vector

Successful transfer of the coding regions from the pCR®8/GW/TOPO® amplification vector to the protein expression vector, pDEST17®, and the subsequent bacterial transformation was confirmed via colony PCR (Figure 3.20). Sequence conformation and analysis were conducted to verify that the candidates were located in the pDEST17® vector in the correct orientation and that the sequence was in-frame (Figure 3.21).

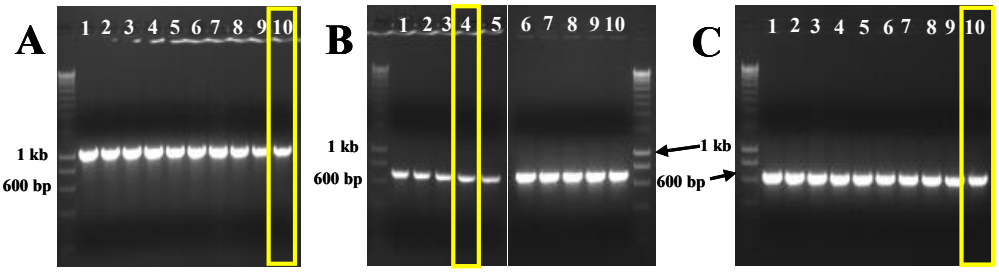


Figure 3.20 - Colony PCR of the open reading frames within the protein expression vector pDEST17®. Plasmids from colonies highlighted in yellow boxes for *TaDMC1* (A), *TaHOP2* (**B**) and *TaMND1* (**C**). In each gel a 1 kb DNA molecular weight marker is shown, however not all band sizes are listed.

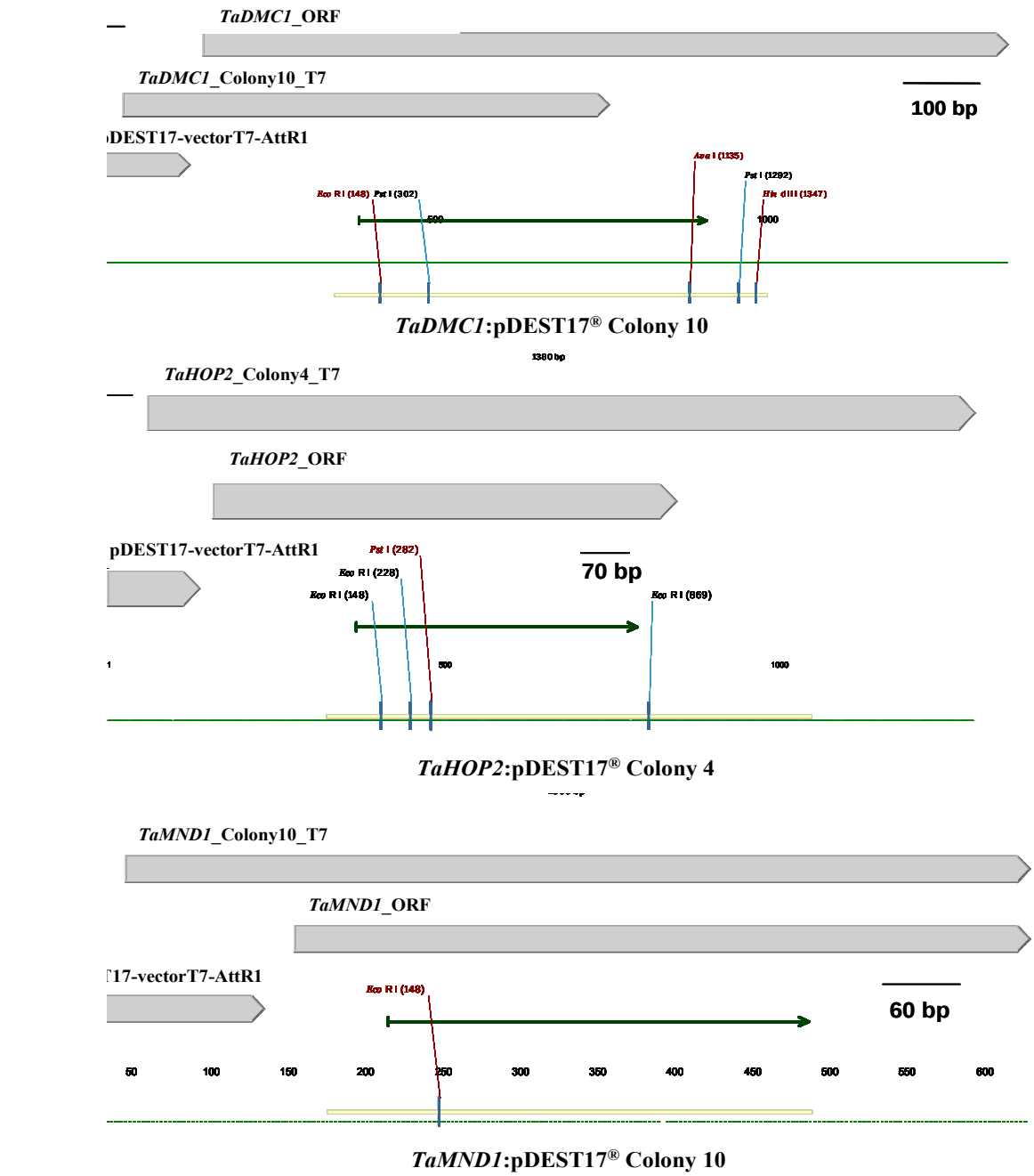


Figure 3.21 - Sequence verification of coding regions and orientation of candidates in the pDEST17® protein expression vector. The pDEST17 – vectorT7 – AttR1 fragment represents the region of the pDEST17® protein expression vector which contains the transcription and protein expression start site, in which all correctly orientated fragments will result in sequence analysis fragments in the same orientation as the pDEST17 – vectorT7 – AttR1 fragment. The candidates, *TaDMC1* **(A)**, *TaHOP2* **(B)** and *TaMND1* **(C)** are therefore in the correct orientation and have complete open reading frames. Obtained sequences (yellow lines) were confirmed to have putative open reading frames (green arrows) of the recombinant proteins.

3.3.6.2 – Heterologous protein expression of meiotic genes

3.3.6.2.1 – *Ta*HOP2 protein extractions

3.3.6.2.1.1 – Primary isolation of *Ta*HOP2

Initial protein extraction of *Ta*HOP2 was successful, with a band present at the expected size of 26 kDa (Figure 3.22). Protein concentration from 1 L of culture for the induced elution 1 sample was 0.1687 μg μL^{-1}. Expression of *Ta*HOP2 can also be observed in the non-induced samples; this is most likely due to the leaky expression from the protein producting vector.

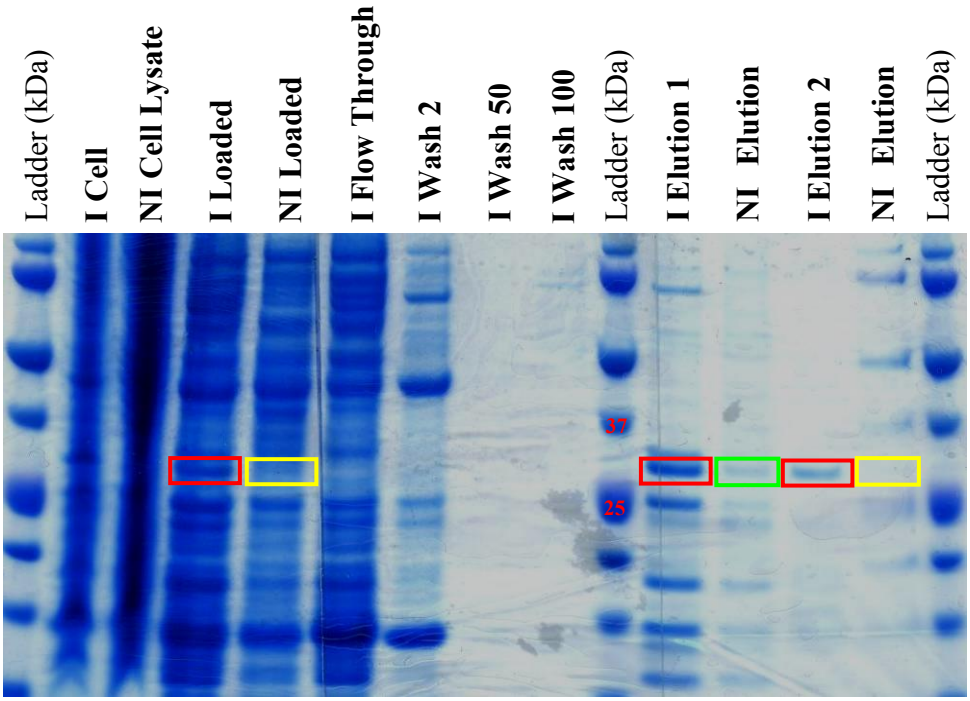


Figure 3.22 - *Ta*HOP2 protein extraction. Isolation of heterologously expressed *Ta*HOP2 protein, with bands at the expected size of 26 kDa (*red* boxes), which can be seen in the sample loaded into the purification column and also in elutions 1 and 2. There is a band at the same size, but with reduced intensity in the non-induced elution 1 (*green* box), yet no such band is present in either the loaded or elution 2 sample for the non-induced extraction (*yellow* boxes). I = Induced and NI = Non-induced. A BIO-RAD Precision Plus Dual Colour Protein Ladder was used; not all sizes are shown.

3.3.6.2.1.2 – Secondary isolation of *Ta*HOP2

Using a larger starting volume of cells which had been induced for *Ta*HOP2 protein production resulted in an increased amount of protein collected after extraction, as can be seen by the band intensity in Figure 3.23. Elution samples 2 and 3 from the secondary extraction were combined, obtaining a concentration of 0.8178 µg μL^{-1}, which is almost five times greater than the previously isolated sample. External protein sequencing services confirmed that the eluted sample was the *Ta*HOP2 protein, with 63% peptide coverage across the protein (Appendix 3.5).

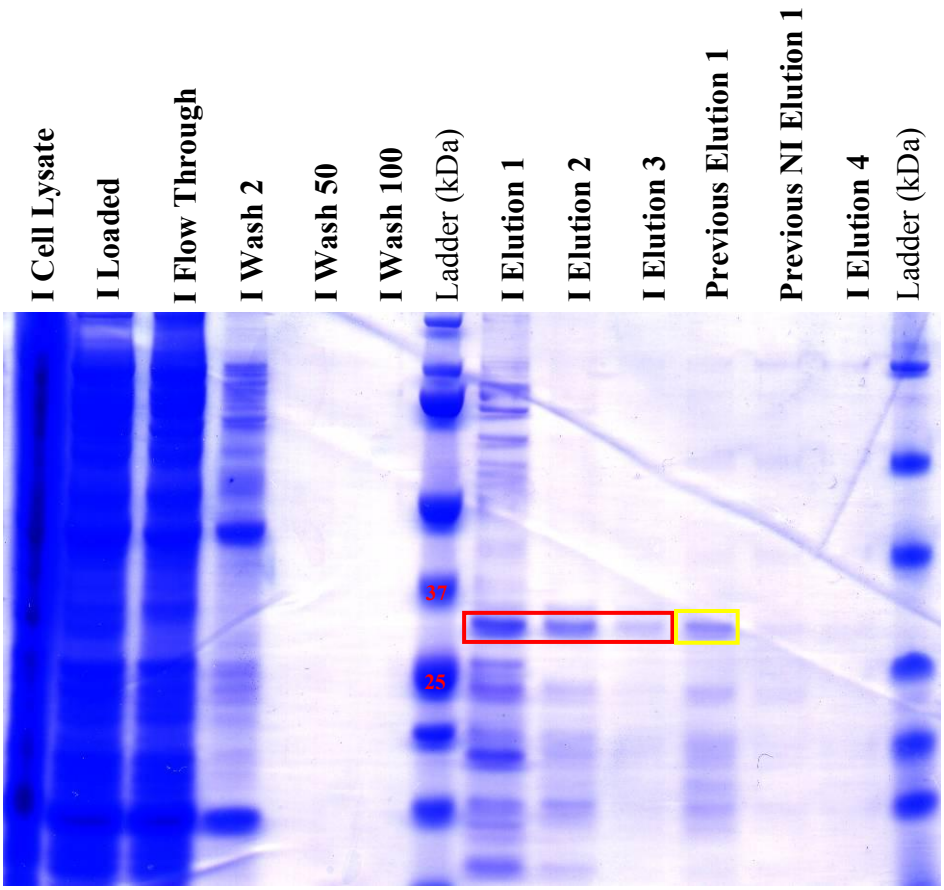


Figure 3.23 - *Ta*HOP2 protein extraction using an increased starter culture volume. Using 2 L of culture volume increased the concentration of the isolated *Ta*HOP2 protein as can be seen between elutions 1-3 (*red* box) and previous elution 1 (*yellow* box) I = Induced. A BIO-RAD Precision Plus Dual Colour Protein Ladder was used; not all sizes are shown.

3.3.6.2.2 – *Ta*DMC1 protein extractions

3.3.6.2.2.1 – *Ta*DMC1 protein extraction

A primary attempt at isolating the *Ta*DMC1 protein as per the successful extraction of the *Ta*HOP2 protein resulted in no observable band present at the expected size of 37 kDa. Using a protein induction temperature of 18°C, a band of the correct size potentially representing *Ta*DMC1 was observed. However, there also appeared to be expression in the non-induced fractions collected (Figure 3.24).

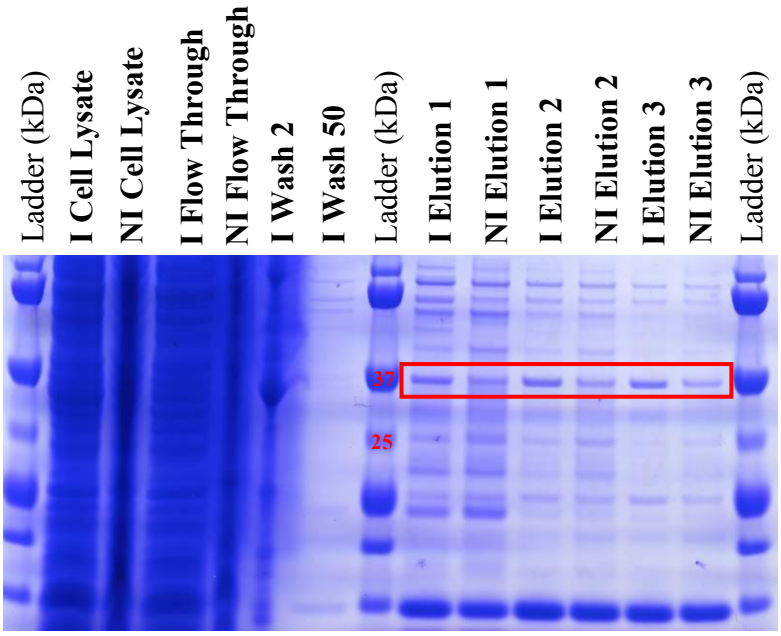


Figure 3.24 - Second *Ta*DMC1 expression and extraction. Reducing the protein expression temperature resulted in the extraction of a product at the expected size of 37 kDa, however this product was also expressed in the non-induced cells. A BIO-RAD Precision Plus Dual Colour Protein Ladder was used; not all sizes are shown.

3.3.6.2.2.2 – *Ta*DMC1 protein expression and repression and subsequent isolation

Expression and isolation of the *Ta*DMC1 protein using an 18°C induction temperature resulted in protein bands present at the expected size of 37 kDa in both the induced and non-incuced samples. Therefore a comparison of protein expression and isolation for: 1) induced heterologous expression, 2) non-induced/repression of heterologous expression, and 3) a non-induced and/no repression of heterologous expression samples was conducted. This revealed strong product bands at the expected size of 37 kDa in both the induced sample and the non-induced/non repressed samples and only a very weak band in the non-induced/repressed sample (Figure 3.25).

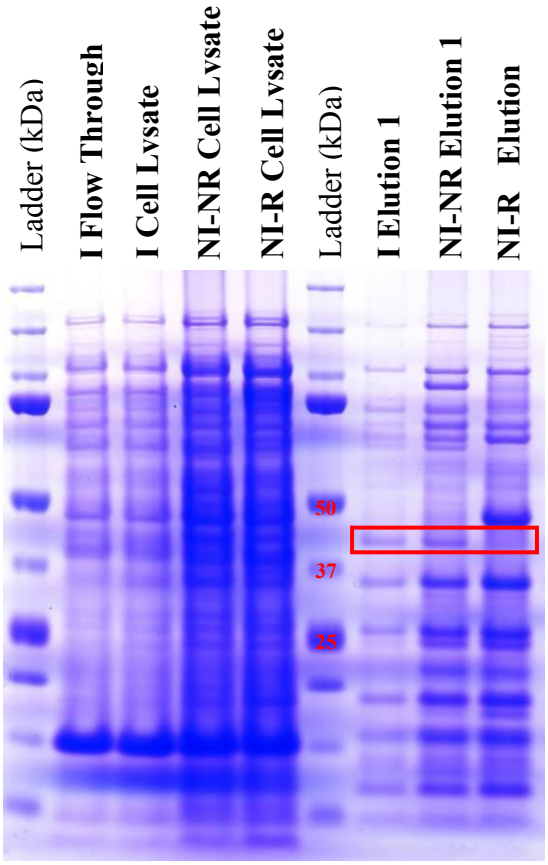


Figure 3.25 - Comparison of *Ta*DMC1 protein expression in induced, non-induced and repressed cell cultures. A strong product band at the expected size of approximately 37 kDa in the induced (I) and non-induced/non-repressed (NI-NR) samples; while in the non-induced/repressed (NI-R) sample there appears to be a weak band representing a small amount of residual expression despite active repression of expression. BIO-RAD Precision Plus Dual Colour Protein Ladder used, not all sizes shown.

3.3.6.2.3 – *Ta*MND1 protein extraction

A primary attempt of isolating the *Ta*MND1 protein as per the successful extraction of the *Ta*HOP2 protein resulted in no observable band present at the expected size of 24 kDa. Secondary protein purification of the *Ta*MND1 proteins was successful, with a product band being present within the induced sample at the expected size of 24 kDa, which is not present to the same intensity within the non-induced sample (Figure 3.26).

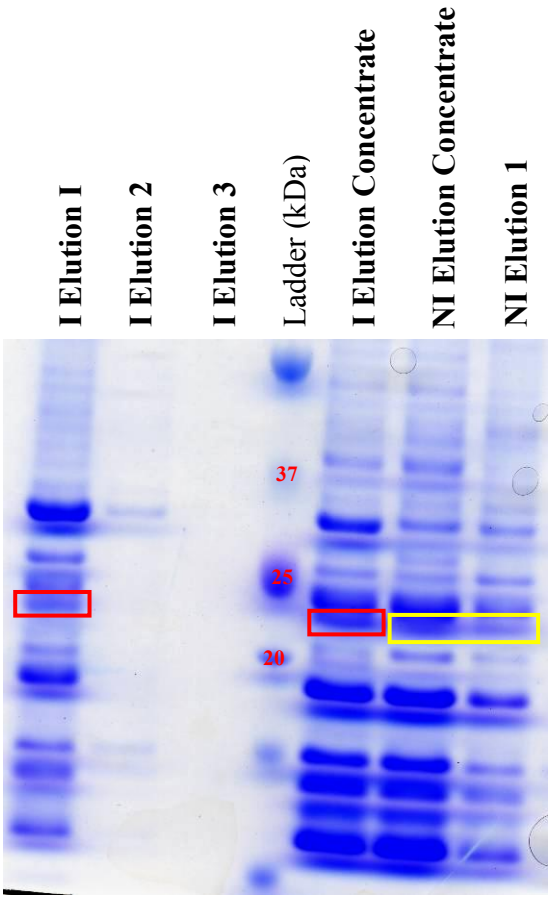


Figure 3.26 - Concentrated protein elution fractions of *Ta*MND1 isolation. A product band can be seen at the expected size of 24 kDa in both the induced (I) elution 1 and induced (I) elution concentrate samples (*red* boxes), which is not seen to the same intensity in the non-induced (NI) samples (*yellow* boxes). A BIO-RAD Precision Plus Dual Colour Protein Ladder was used; not all sizes are shown.

3.3.7 – DNA binding ability of *Ta*DMC1, *Ta*HOP and *Ta*MND1

3.3.7.1 – Linearisation of double stranded DNA

Double-stranded DNA was linearised with restriction enzymes, in preparation for the DNA binding assays (Figure 3.27).

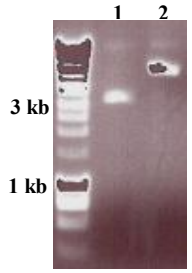


Figure 3.27 - Linearisation of double-stranded DNA for the binding assay. Restriction digestion of the double-stranded DNA (PhiX174, NEB) with the enzyme *Xho*I was successful (lane 2), when compared with the undigested DNA (lane 1). A 1 kb DNA molecular weight marker is shown, but not all band sizes are listed.

3.3.7.2 – Testing protein elution fractions for DNA binding ability

The induced and non-induced samples were tested in DNA binding assays as the protein elution samples contain not only the candidate proteins but also bacterial contaminating proteins. In Figure 3.28, DNA binding assays showed DNA binding ability in the heterologously-expressed *Ta*DMC1 (Figure 3.28 A) and *Ta*HOP2 (Figure 3.28 B). With the non-induced fractions in these gels, which contain bacterial contaminants or leaky protein expression from the protein production vector, very weak DNA binding ability was also evident. While the gel electrophoresis analysis of *Ta*HOP2 binding abilty was not continued to maximum separation, a small difference in the shifts was observed between the induced and non-induced.

In contrast, both the induced and non-induced elution fractions of *Ta*MND1 showed a high DNA binding ability (Figure 3.28 C). This binding is therefore unable to be attributed exclusively to the *Ta*MND1 protein as it is not able to be determined whether the DNA binding ability of the non-induced sample is contributed to soley by either bacterial contaminating proteins or leaky expression of *Ta*MND1. This bacterial contamination or leaky expression of the protein production vector appears to be less within the DMC1 and HOP2 protein non-induced elution samples as the DNA binding ability of these samples does not have the same degree of shift of the DNA molecules, as mentioned above. In the case of HOP2, this is most probably due to the ammonium sulphate selection conducted; however,

the higher degree of DNA binding was still not seen in the DMC1 non-induced samples which had not undergone ammonium sulphate selection and this is most probably due to protein induction occurring at different temperatures (18°C for DMC1 versus 37°C for MND1).

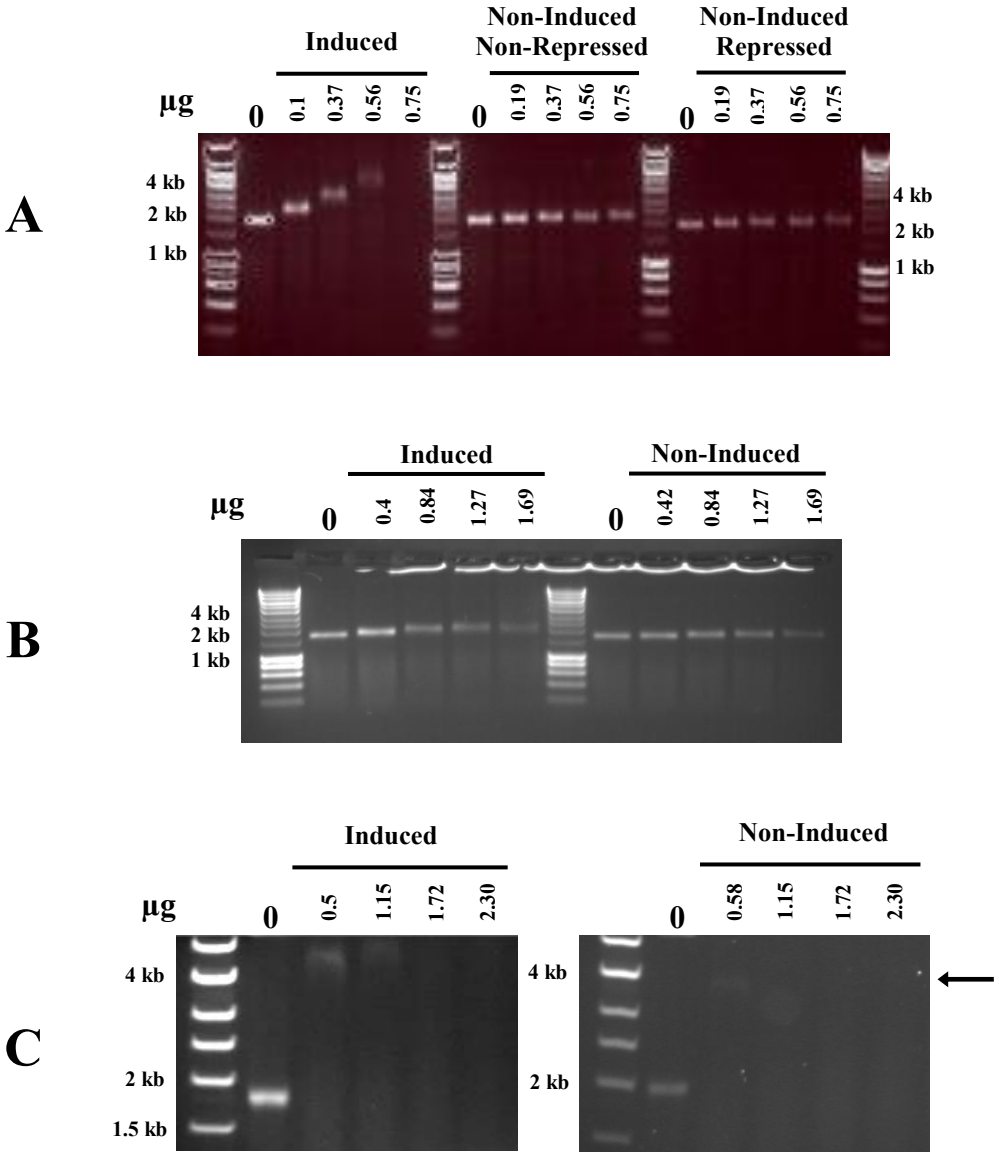


Figure 3.28 - Testing DNA binding ability of induced and non-induced protein elutions. (**A**) The *Ta*DMC1 protein is contributing to the elution fraction"s ability to bind single stranded (ss) DNA, with the non-induced elution fractions having little ability to bind to the ssDNA. (**B**) The *Ta*HOP2 protein is contributing to the elution fraction"s ability to bind ssDNA, with the non-induced elution fraction having little ability to bind to the ssDNA. (**C**) DNA binding ability is observed in both the induced (left) and the non-induced (right, arrow). A 1 kb DNA molecular weight marker is shown in each gel, however not all band sizes are listed.

3.3.7.3 – DNA binding ability of *Ta*DMC1 and *Ta*HOP2

Both *Ta*DMC1 and *Ta*HOP2 proteins were found to bind to single-stranded and double-stranded DNA (Figure 3.29) and both have more affinity to single-stranded DNA. The single-stranded DNA assay for HOP2 shows that the lane with 0.57 µg of protein appears to be less retarded when compared to the DNA in lanes either side of this sample (Figure 3.29 B). It is likely that human error led to this result given that all other lanes with increasing concentrations have increased DNA retardation. With single-stranded and double-stranded present in the *in vitro* reaction, both proteins again have a higher affinity for the single-stranded when compared to the double-stranded DNA (Figure 3.30).

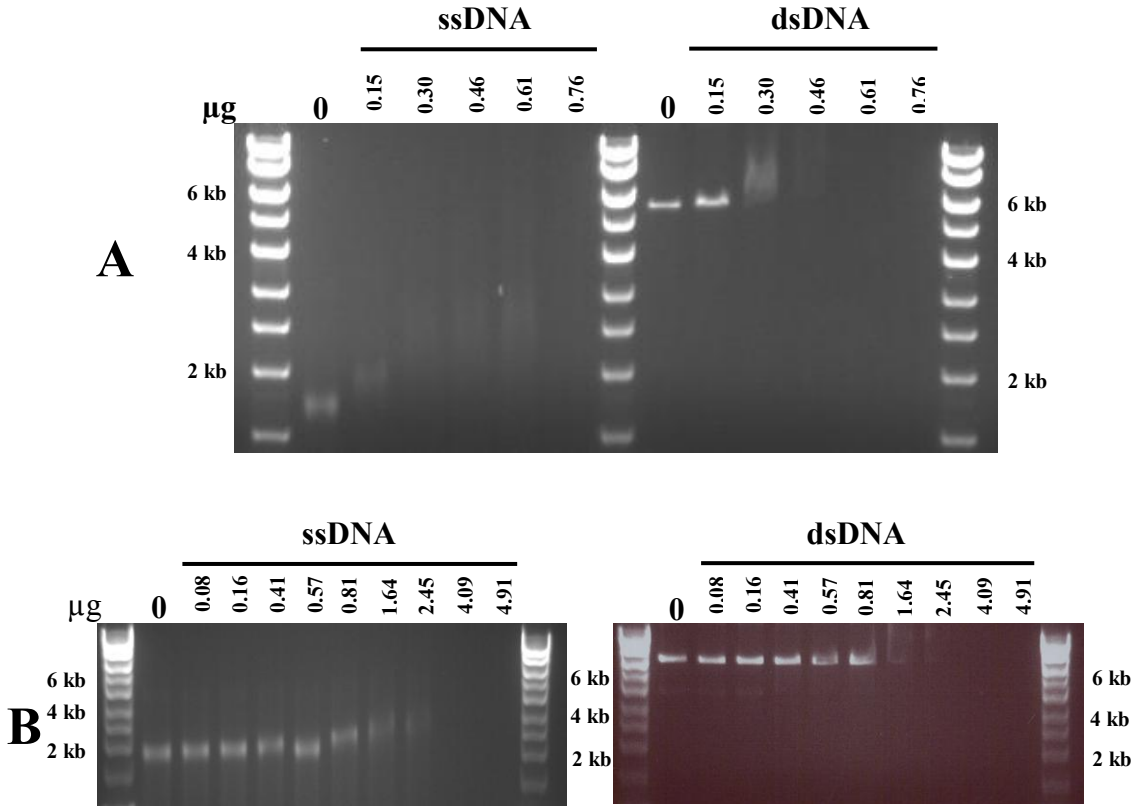


Figure 3.29 - *Ta*DMC1 and *Ta*HOP2 DNA binding ability. *Ta*DMC1 (**A**) and *Ta*HOP2 (**B**) have the ability to bind to single-stranded (ss) and double-stranded (ds) DNA, with both having a higher affinity to ssDNA. A 1 kb DNA molecular weight marker is shown; however, not all band sizes are listed.

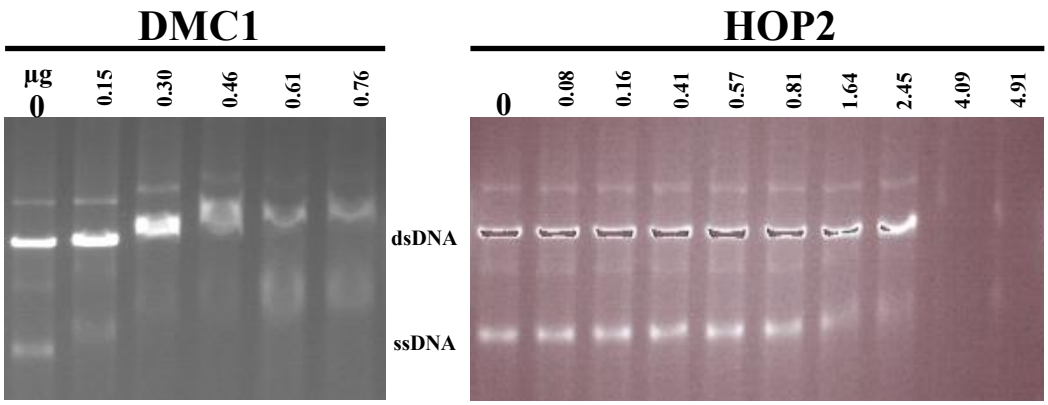


Figure 3.30 - Competitive binding ability of *Ta*DMC1 and *Ta*HOP2. Both *Ta*DMC1 and *Ta*HOP2 proteins have a higher affinity for single-stranded (ss) DNA over double-stranded (ds) DNA when both species are present.

3.4 – Discussion

3.4.1 – The meiotic genes of *DMC1*, *HOP2* and *MND1* are conserved in wheat

TaDMC1, *TaHOP2* and *TaMND1* were found to be present in wheat with their coding regions and putative translations conserved. This was particularly pronounced for *Ta*DMC1 which contained 53% of the same amino acids as yeast DMC1. *Ta*HOP2 and *Ta*MND1 are conserved to a lesser extent, with approximately 40% similarity to the more divergent human and mouse proteins. This reduced sequence conservation is not necessarily unexpected. Importantly, it has been shown that although sequence divergence can occur, functional conservation can still be retained; as is the case with the *ASY1* gene in bread wheat (Boden *et al.*, 2007; Boden *et al.*, 2009). The fact that *Ta*DMC1 has been highly conserved throughout evolution could point to the importance of this gene for a viable meiosis to occur in any higher eukaryote.

Further evidence for the importance of these candidates is the 3D modelling which was conducted with the *Ta*DMC1 putative amino acid sequence. This also revealed a high level of secondary and tertiary structure conservation. Conservation at the protein structural

level has been proposed to provide evidence for conserved functionality (Zarembinski *et al.*, 1998; Khoo *et al.*, 2008). Unfortunately since the 3D modelling computations are based on previously crystallised and resolved protein structures being available, for example the *Hs*DMC1 protein (Kinebuchi *et al.*, 2004; Hikiba *et al.*, 2008), the structures for *Ta*HOP2 and *Ta*MND1 were unable to be modelled as neither protein structure has been physically determined in any species.

Due to the low concentrations of the heterologously expressed proteins extracted and the impurity of the elution samples obtained, determination of the protein structures of *Ta*HOP2 and *Ta*MND1 was unable to be conducted as part of this project. Even so, based on the nucleotide and amino acid conservation across the species shown, it is likely that these candidate genes will have similar roles across divergent species. Therefore investigation of the bread wheat genes of DMC1, HOP2 and MND1 was carried out to confirm the conservation of these genes.

3.4.2 – Chromosome locations of *TaDMC1*, *TaHOP2* and *TaMND1*

Using nullisomic-tetrasomic wheat samples and Southern analysis, *TaDMC1*, *TaHOP2* and *TaMND1* were each located to a particular chromosome group in wheat, with one copy on each of the wheat genomes, A, B and D. *TaHOP2* was found to reside on wheat chromosome group 4, while *TaDMC1* and *TaMND1* were on group 5. Historically these two chromosome groups have been noted as containing regions of interest in relation to chromosome pairing (Driscoll, 1972; Sears, 1976). Chromosome group 5 has a major locus that affects pairing. This locus is specifically located on the long arm of 5B, and is termed the *P*airing *H*omoeologous 1 (*Ph1*) locus (Moore, 2000). The *Ph1* region is currently undergoing major sequencing initiatives and further functional characterisation, with Moore and colleagues determining that the major gene/s contributing to this locus are a group of eight *CDK*-like genes (Al-Kaff *et al.*, 2008). However, neither *TaDMC1* nor *TaMND1* have been mapped to this region (Griffiths *et al.*, 2007; Al-Kaff *et al.*, 2008), and thus reside elsewhere

on this chromosome group. Chromosome group 4 has also been found to contain pairing loci that contribute to homologous/homoeologous control (Driscoll, 1973). However, until further sequencing and research of the wheat genome (of such areas) is obtained, it will remain unknown as to whether *TaHOP2* maps to such pairing loci.

3.4.3 – Correlated expression of meiotic genes

TaHOP2 and *TaMND1*, which have been found to act in concert as a heterodimeric protein complex (Enomoto *et al.*, 2006; Pezza *et al.*, 2006), were observed to have correlated expression. This indicates that these two genes may be co-regulated. However, these two genes did not have high correlated expression with *TaDMC1*, despite having a known associated role (r = 0.89 and 0.84 for *Ta*HOP2 and *Ta*MND1, respectively). Crismani and colleagues (2006) have previously investigated correlated expression, and thus likely co-regulation, of a number of meiotic genes. Interestingly a number of the genes investigated by Crismani and colleagues (2006) were also found to have correlated expression with *TaHOP2* and *TaMND1*: the known meiotic genes of *TaASY1* which is associated with homoeologous chromosome pairing and synapsis (Boden *et al.*, 2007; Boden *et al.*, 2009), the mis-match repair protein *TaMSH6* (Wu *et al.*, 2003), *TaPHS1* (Pawlowski *et al.*, 2004); Khoo *et al.*, *unpublished data*), the recombination helper *TaRPA* (Neale and Keeney, 2006) and the recombinase proteins *TaRAD51C* and *D* (Khoo *et al.*, 2008). Interestingly the correlation of *TaMND1* with *TaRAD51D*, which has a proposed role in regulating pathogenesis-related (PR) genes during the systemic acquired defence response in plants (Durrant *et al.*, 2007), may explain the high expression of *TaMND1* seen in the vegetative leaf tissues (Figure 3.12). There was no recorded expression of the *TaHOP2* transcript in leaf tissue, despite conducting the assay in triplicate. This is surprising since many meiotic genes are found to have at least some vegetative tissue expression; as has been seen in *TaMND1* and *TaDMC1* (Figure 3.12) and also *TaASY1* (Boden *et al.*, 2007; Boden *et al.*, 2009). It could be that this transcript is

present at a very low copy number and that Q-PCR is not sufficiently sensitive to detect its expression.

3.4.4 – *Ta*MND1 localises to meiotic chromosomes during early prophase I

Localisation of *Ta*DMC1 and *Ta*MND1 was attempted in this project with anti-*At*DMC1 (Chelysheva *et al.*, 2007) and anti-*At*MND1 (Vignard *et al.*, 2007). However, HOP2 localisation was not undertaken as access to an antibody raised against this protein could not be obtained. Immunolocalisation using the anti-*At*DMC1 antibody was attempted several times without success, even after having used various techniques and modifications of the procedures within those techniques. Previously the anti-*At*DMC1 antibody has been used in meiotic chromosome spreads in Arabidopsis with success (Vignard *et al.*, 2007).

A number of reasons could be contributing to a lack of signal, including that the localisation was attempted in wheat meiocytes with an antibody raised against an Arabidopsis protein. The *At*DMC1 antibody was raised against a peptide corresponding to the first 18 amino acids of DMC1, so that it would not have cross-reactivity with the closely-related RAD51 protein (Chelysheva *et al.*, 2007). When compared to the Arabidopsis amino acid sequence, the first 18 amino acids of the wheat sequence only shares seven identical matches with an additional seven having similar properties (Figure 3.31). Consequently, the *At*DMC1 antibody may not cross-react with the wheat DMC1 protein, producing no signal in the immunolocalisation experiments.

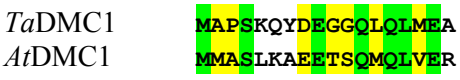


Figure 3.31 - Amino acid alignment of the anti-*At*DMC1 antigen with the corresponding wheat sequence. The Arabidopsis (*At*) antibody for the DMC1 protein was raised against a peptide corresponding to the first 18 amino acids of the protein. In green are the identical amino acids shared between the Arabidopsis and wheat (*Ta*) proteins, while the amino acids with similar properties are highlighted yellow.

However the sequence differences between the Arabidopsis and wheat proteins does not explain why the *At*DMC1 antibody failed to recognise the Arabidopsis protein within the western analysis, where the anti-*Ta*ASY1 antibody was used as a positive control for both Arabidopsis protein quality and the western technique (Figure 3.18). Unfortunately only immunolocalisation data with this particular antibody have been reported (Chelysheva *et al.*, 2007; Muyt *et al.*, 2007; Chelysheva *et al.*, 2008) and therefore it is difficult to say whether the antibody is only functional in immunolocalisations and is unsuitable for the western technique or if the antibody has not worked due to damage in transit (when sent from France to Australia).

A reason why the *At*DMC1 antibody may not recognise the Arabidopsis proteins in the western analysis could be due to the antibody not recognising a denatured form of the DMC1 protein. However, this seems unlikely as the antibody was raised against only an 18 amino acid peptide; which is not part of any of identified DMC1 domains. What seems more likely is that the antibody is not working in the western analysis due to the „environment" in which the western is conducted. For example, the pH and buffers in which the western is being conducted could have had affected the result. However the literature about the variability of antibodies between techniques, especially in the case of westerns verses immunolocalisations, is lacking. Therefore it can only be postulated that this technique variability may be contributing to the cause of the result obtained with the *At*DMC1 antibody.

The localisation of the *Ta*MND1 protein also produced variable results, with immunolocalisation of this protein only being successful using the meiotic chromosome spread technique (Figure 3.17). This was the case even after several attempts were made at optimising the polyacrylamide embedment of meiocytes by lowering the fixative percentage (4% to 2%) and increasing the time of cell permeabilisation. While the cytoplasm signal, which was detected in the primary attempt at this technique (Figure 3.14), was able to be decreased through modifying the procedures, there was no observable signal to be discerned

(above background) in the secondary attempt (Figure 3.15). However when using the same antibody with the meiotic chromosome spread technique, foci were detected.

Could the variations seen between the different techniques result from the different classes of antibodies in this investigation? The *At*DMC1 antibody is a monoclonal antibody, which as mentioned before has only seven amino acids (out of 18) that are the same as the wheat protein, while the polyclonal *At*MND1 antibody was raised against the entire protein length and has 71% of the same amino acids as the *Ta*MND1 protein. One reason why the polyclonal *At*MND1 antibody has worked against the wheat protein, yet the monoclonal *At*DMC1 antibody has not, might be due to the polyclonal nature of the antibody itself. Despite the decreased affinity of the antibodies for the wheat proteins, the polyclonal antibody in this case will be able to recognise more of the epitopes in the *Ta*MND1 protein, which in turn will allow for the creation of multimeric structures to occur between the *Ta*MND1 protein and the antibody, thus creating a larger target for the secondary antibodies (Harlow and Lane, 1988).

Foci representing the *Ta*MND1 protein were observed during pre-meiotic interphase, yet these foci were less intense then those seen in the following stage of leptotene, which despite the increased intensity were lower in number. Interestingly, resurgence in the intensity of the foci was observed during late zygotene, to be followed by a decrease in intensity during pachytene, with little signal being detected during diplotene. These results are somewhat similar to those seen for MND1 localisation in Arabidopsis spreads, as Vignard and colleagues (2007) found high foci during the early stage of leptotene which persisted until pachytene.

The foci emergence during the earliest stages of meiosis could be corresponding to when MND1 and HOP2 would be proposed to be most active. This would be the time in which the Spo11 protein induces double stranded breaks, resulting in the required repair of the chromosomes through homologous recombination and in which *Ta*DMC1 has an important role (Keeny, 2001; Vignard *et al.*, 2007). The reappearance of the *Ta*MND1 foci during late-

zygotene, when the recombination of homologous chromosomes is almost completed, could perhaps be for the removal of the *Ta*DMC1 protein from the chromosomes, whose role is also completed. Evidence supporting this is the accumulation of *At*DMC1 foci in the Arabidopsis *mnd1* mutant background, compared to the comparative wild-type stage (Vignard *et al.*, 2007).

3.4.5 – Protein expression and isolation

Protein expression and extraction was successful for *Ta*HOP2 and *Ta*DMC1, with the isolation of the *Ta*DMC1 protein occurring through changes in the conditions of cell culture and protein expression induction. These optimised conditions were based on experimental data from Khoo *et al.* (*unpublished data*) which was completed for the isolation of the full-length denatured form of the *Ta*DMC1 protein (for the production of a wheat DMC1 polyclonal antibody). Unfortunately the *Ta*MND1 protein was unable to be isolated within a usable protein elution fraction, as the non-induced sample was also able to bind DNA (Figure 3.29). This DNA binding ability gives rise to the reasoning that one of the bacterial contaminating proteins, which are also located within the *Ta*MND1 induced protein elution fraction, could be providing DNA binding ability which was seen in any binding assay conducted with the MND1 protein elution fraction.

The product bands isolated from *Ta*DMC1 and *Ta*MND1 were not sequenced, as per the *Ta*HOP2 band; as the conformational *Ta*DMC1 protein sequencing had already been conducted by Khoo *et al.* (*unpublished data*) during the isolation of the *Ta*DMC1 protein (denatured) for antibody production. As the cells used for this were from the same stock and underwent the same culturing parameters (although not the same extraction method), it was deemed unnecessary to repeat the protein sequencing. Since both the induced and non-induced elution fractions of the *Ta*MND1 protein extract showed DNA binding ability, these fractions were deemed unable to be used for further functional analysis and therefore were not sent for amino acid sequencing.

Future optimisation of the protein expression and isolation of the heterologously expressed protein could also be helpful in isolating a more pure sample of each of the proteins for further functional analysis. This optimisation could include different salt concentrations during the extraction process and also different tagging molecules for other, perhaps more specific, extraction methods. Also a useful future control for the expression of proteins of interest will be to conduct an „empty" protein production vector cell culturing instead of a non-induced protein production vector cell culture, which would eliminate the „leaky" expression of the proteins of interest which appeared to be observed during this investigation.

3.4.6 – Conserved functionality of *Ta*DMC1 and *Ta*HOP2

The conserved nature of these genes, both at the nucleotide, amino acid and the structural level (for *Ta*DMC1), provides good evidence that these wheat proteins also have conserved functions when compared to their respective orthologues. DNA binding ability assays were conducted to test whether the wheat DMC1 and HOP2 proteins possessed DNA binding function, as has been seen in all species reported to date (Chen *et al.*, 2004; Kant *et al.*, 2005; Enomoto *et al.*, 2006; Pezza *et al.*, 2006). Due to the impure elution samples, with their contaminating bacterial bands, the molarities of the protein samples were unable to be calculated to give a definitive value; as any calculation would be a misrepresentation of the protein samples. Thus, when conducting the DNA binding assays the micrograms (µg) of protein elution added to the reaction was recorded. Due to this, making comparisons to the previously published DNA binding assay in other species would be difficult; therefore the affinity of the binding ability is only loosely commented upon.

However, what can be commented upon is the preference which each of the proteins has for the two species of DNA (linear single-stranded versus double-stranded). It was found that with increasing amounts of protein elution samples, both *Ta*DMC1 and *Ta*HOP2 proteins were found to have a preference for single-stranded DNA over the double-stranded DNA, either when only one or both species of DNA were present. The *Ta*DMC1 protein was found

to bind preferentially to single-stranded DNA as per the yeast and human proteins (Masson *et al.*, 1999; Haruta *et al.*, 2006), and the DNA binding ability of *Ta*HOP2 is also consistent with the mouse protein; having a preference for single-stranded DNA (Pezza *et al.*, 2006). The binding ability of the HOP2:MND1 protein complex was unable to be tested due to the inability to isolate the *Ta*MND1 protein, however the protein complex‟s DNA binding ability has been observed in mouse, humans and yeast (Chen *et al.*, 2004; Enomoto *et al.*, 2006; Pezza *et al.*, 2006). Both the yeast and human complexes bind preferentially to linear double-stranded DNA (Chen *et al.*, 2004; Enomoto *et al.*, 2006), yet the mouse complex also shows the *Mm*HOP2 protein‟s preference for single-stranded DNA (Pezza *et al.*, 2006). Since the individual DNA preferences of the yeast and human equivalents have not yet been published, it is unknown if the DNA preference of the single protein is the protein complex‟s DNA preference.

The conservation of DNA preference is critically important as the DMC1 protein needs to bind to the processed ssDNA end for the ability to conduct strand invasion into the partnering homologous chromosome. While the actual roles of HOP2 and MND1 are still being debated; the current model is that the protein complex of HOP2 and MND1 stabilises the chromosome pair before and during strand invasion (Petukhova *et al.*, 2005; Henry *et al.*, 2006). It is postulated that they would be binding to the double-stranded DNA surrounding the break within the chromosome for recombination (for the aforementioned stabilisation) (Petukhova *et al.*, 2005; Henry *et al.*, 2006), which fits with the DNA binding ability of the yeast and human proteins. However, what of the mouse HOP2:MND1 complex and its preference for single-stranded DNA? Could this particular protein complex have a modified role within the proposed model? Could the protein complexes with single-stranded DNA preference be stabilising the single strands during recombination instead of the double-stranded DNA? Or could the protein complex be binding to the free double-stranded DNA, with the single-stranded DNA most likely occupied with other proteins including DMC1, to stabilise the strands for the recombination process?

In the isolation and functional characterisation of *Ta*DMC1, *Ta*HOP2 and *Ta*MND1, the meiotic genes were found to be highly conserved across divergent species. Importantly, confirming the conservation of these meiotic candidates in bread wheat, and investigating theses genes in conjunction with novel candidates (previously isolated in Chapter 2 and further investigated in Chapter 5), has allowed further understanding of how meiosis works in higher eukaryotes.

Chapter 4

Chapter 4 – Investigating protein interacting partners of *Ta*DMC1 and *Ta*ASY1 meiotic proteins

4.1 – Introduction

Protein-protein interactions are essential for virtually all cellular processes, and the identification of these interactions is a substantial part in understanding an individual protein‟s function (Guan and Kiss-Toth, 2008). The identification of interacting proteins during a particular biological process allows for the generation of protein interaction maps (PIMs), which when encompassed with the ever increasing genomic sequences available furthers our understanding of systems biology (Mukherjee *et al.*, 2001).

Protein-protein interactions can be analysed by various *in vitro* and *in vivo* methods, including affinity and molecular-sized based chromatography, plasmon resonance and fluorescence techniques (for example fluorescence resonance energy transfer (FRET) and split-GFP assays (Lalonde *et al.*, 2008)). However many of the interaction techniques fail to analyse large-scale protein networks effectively as they are based on interactions between two known and isolated proteins (Mukherjee *et al.*, 2001). One technique which has become popular for the identification and characterisation of large-scale protein networks is the eukaryotic *in vivo* yeast two-hybrid approach, which can be used to screen a prey cDNA library with a single bait protein allowing for the identification of multiple protein interactions at the one time (Figure 4.1) (Guan and Kiss-Toth, 2008). One of the most appealing features of the system is that the process of confirming an interacting protein results in the isolation and cloning of at least part of the corresponding gene (Mukherjee *et al.*, 2001).

Figure 4.1 - Yeast two-hybrid assay system. Using the yeast two-hybrid assay system a single protein, located within the BD-bait fusion expression plasmid, can be screened for interactions with a cDNA library located within the AD-prey fusion expression plasmid. BD = binding domain; AD = activation domain. Figure from Mukherjee *et al.*, 2001.

Protein interactions have been widely studied in the process of meiosis, although investigations have mainly centred upon studying the relationship between two known genes (Siaud *et al.*, 2004; Petukhova *et al.*, 2005; Sarai *et al.*, 2006). The meiotic protein DMC1, investigated in Chapters 2 and 3 of this dissertation, has been found to physically interact with numerous proteins, including the HOP2:MND1 protein complex (Petukhova *et al.*, 2005; Vignard *et al.*, 2007), Rad51 and Brca2 (Siaud *et al.*, 2004), MSH4 (Neyton *et al.*, 2004), p53 (Habu *et al.*, 2004), TopoII (Iwabata *et al.*, 2005) and Rad54B (Sarai *et al.*, 2006).

In contrast, the meiotic protein ASY1 (investigated in Chapter 2) has only been found to interact physically with one other protein besides self-interaction: with the yeast ASY1 homologue, HOP1, found to interact with the yeast-specific protein RED1 (Niu *et al.*, 2005). The lack of proteins interacting with ASY1 is surprising, since this particular protein has been found to be essential for meiosis in a number of species (Caryl *et al.*, 2000; Nonomura *et al.*, 2004; Boden *et al.*, 2007; Boden *et al.*, 2009). Consequently, the research presented in this

chapter is focused on furthering the understanding of the proteins that potentially interact with the two meiotic proteins, *Ta*DMC1 and *Ta*ASY1.

4.2 – Materials and Methods

The identification of interacting proteins of the two meiotic proteins of *Ta*ASY1 and *Ta*DMC1 was conducted using a yeast two-hybrid approach, with screening for interacting partners done against an early meiotic cDNA library; and then subsequently confirmed as interacting partners, isolated and characterised (Figure 4.2). The initial screen for *Ta*ASY1 was conducted with the isolated HORMA domain of the *Ta*ASY1 protein which allowed for a „broad" screen of identifying possible interacting partners, which were later confirmed/refuted to then interact with the full length *Ta*ASY1 protein. In contrast, the full length *Ta*DMC1 protein was used in screens to identify protein partners specific for *Ta*DMC1 only. This reduced probability of identifying false positives between proteins which may partner the closely related meiotic protein of *Ta*RAD51.

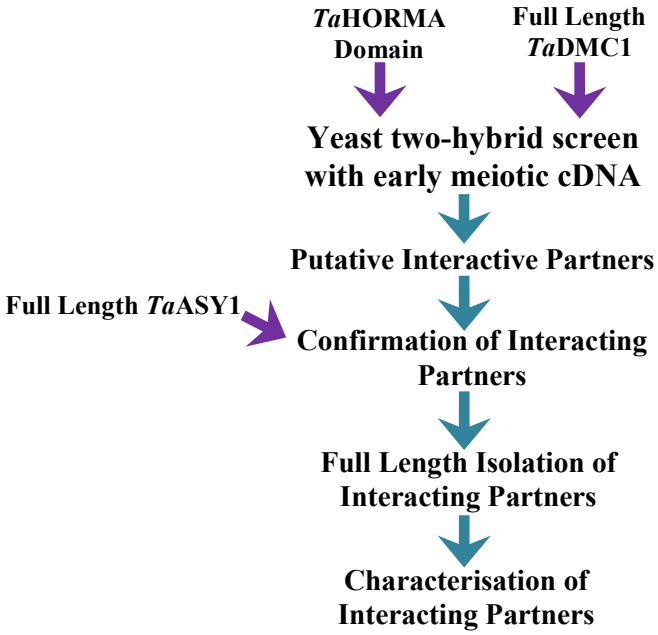


Figure 4.2 - Research outline of the identification of interacting partners of *Ta*ASY1 and *Ta*DMC1. The full length *Ta*DMC1 protein was used in the initial screen; with the *Ta*HORMA domain initially used to conduct „broad" screen for „ASY1 interactors. This was followed later by confirmation of interaction with the full length *Ta*ASY1 protein.

4.2.1 – PCR amplification of *TaDMC1*, *TaASY1* and *TaHORMA* for insertion into the yeast two-hybrid bait vector

Isolation of *TaDMC1*, *TaASY1* and *TaHORMA* (sequences in Appendix 4.1) was conducted using oligonucleotides with restriction sites which results in the placement of these sites either side of the coding region (Table 4.1); allowing for the digestion of these coding fragments from a primary vector and the insertion into the bait vector pGBKT7 (Clontech).

Table 4.1 - Oligonucleotides for the PCR amplification of *TaDMC1*, *TaASY1* and *TaHORMA*.

Coding region	Oligonucleotide 5′-3′		Restriction site
TaDMC1	Forward	AATTCATGGCGCCGTCCAAG	*Eco*RI
	Reverse	GGATCCTCAGTCTTTCACATCCATCAATC	*Bam*HI
TaASY1	Forward	CTCATATGGTGATGGCTCAGAAGACG	*Nde*I
	Reverse	GACTGCAGGGACTTCTGGCGCTTGAC	*Pst*I
TaHORMA	Forward	CTCATATGGAGATCACGGAGCAGGAC	*Nde*I
	Reverse	CTGGATCCTGTTAGCATCATTAGCATCAC	*Pst*I

PCR mixtures (20 µL) contained 2 µL of 10x PCR buffer (Invitrogen), 0.7 µL of $MgCl_2$ (50 mM, Invitrogen), 3.2 µL of dNTPs (1.25 mM), 1 µL of each oligonucleotide (10 µM), 1 µL of meiotic spike cDNA as a template (Chapter 2, section 2.2.4), 0.2 µL of *Taq* Polymerase (5 U/µL, Invitrogen) and 10.9 µL of sterile deionised water.

PCR cycles were initiated at 95°C for 7 minutes, followed by 35 cycles of 94°C for 30 seconds, 61°C, 62°C and 61°C for 30 seconds for *TaDMC1*, *TaASY1* and *TaHORMA*, respectively, 72°C for 1 minute; with a final extension cycle at 72°C for 10 minutes. PCR products were analysed via gel electrophoresis (1.5% agarose) and visualised using ethidium bromide and a UV transilluminator (FirstLight™ UV illuminator, UVP, USA). PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per the manufacturer's instructions (Macherey-Nagel).

4.2.2 – Ligation and confirmation of *TaDMC1*, *TaASY1* and *TaHORMA* in primary vectors

Ligations of the PCR fragments of *TaDMC1* into pCR®8/GW/TOPO® vector (Invitrogen), and *TaASY1* and *TaHORMA* into pGEM®-T Easy vector were conducted as per Chapter 3, section 3.2.6.1.2 and Chapter 3, section 3.2.1.3 for the pCR®8/GW/TOPO® and pGEM®-T Easy vectors, respectively (Appendix 4.2). Recombinant plasmids were transformed into competent bacterial cells as per manufacturer's instructions.

PCR was used to screen for positive colonies for insertions of fragments. Colony PCR mixtures (20 µL) containing 2 µL of 10x PCR Buffer (Invitrogen), 0.75 µL of $MgCl_2$ (50 mM, Invitrogen), 3.2 µL of dNTPs (1.25 mM), 0.5 µL of oligonucleotide GW1 (10 µM, section 3.2.6.1.3 for pCR®8) or T7 oligonucleotide (10 µM, Chapter 2 section 3.2.1.5 for pGEM®-T), 0.5 µL of oligonucleotide GW2 (10 µM, section 3.2.6.1.3 for pCR®8) or SP6 oligonucleotide (10 µM, Chapter 2, section 3.2.1.5 for pGEM®-T), 0.2 µL of *Taq* Polymerase (5U/µL, Invitogen) and 12.85 µL sterile deionised water. Single colonies were isolated using a sterile pipette tip, swirled into individual PCR mixtures and then inoculated and cultured overnight in 5 mL of LB liquid medium with spectinomycin or ampicillin (50 or 100 µg mL^{-1}, respectively) as a selective agent for the pCR®8/GW/TOPO® and pGEM®-T Easy vectors, respectively.

PCR cycling was initiated at 95°C for 7 minutes, followed by 35 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute; with a final extension cycle of 10 minutes at 72°C. PCR products were analysed using gel electrophoresis (1.5% agarose) and visualised as per section 4.2.1. Plasmid DNA was isolated from the positive colony liquid cell cultures as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen, AUS) for verification of the insert via sequence analysis. Sequence analysis, as per section 3.2.1.5 but using GW1 and GW2 oligonucleotides for *TaDMC1*:pCR®8/GW/TOPO® vector, confirmed the *TaDMC1*, *TaASY1* and *TaHORMA* within their respective primary vectors.

4.2.3 – Restriction digest of *TaDMC1*, *TaASY1* and *TaHORMA* coding regions from the primary vectors

A restriction digest of *TaDMC1, TaASY1* and *TaHORMA* fragments from the primary vectors of pCR®8/GW/TOPO® and pGEM®-T Easy was conducted on the isolated plasmid from section 4.2.2. Primary reactions for the digestion of *TaDMC1*:pCR®8/GW/TOPO® contained 500 ng of plasmid, 2 µL Buffer 2 (NEB), 1 µL *Eco*RI (NEB) and sterile deionised water to a final volume of 20 µL. Primary reactions (20 µL) for *TaASY1* or *TaHORMA*:pGEM®-T Easy contained 500 ng of plasmid, 2 µL of *Pst*I enzyme (NEB), 2 µL of Buffer 3 (NEB), 0.2 µL of BSA (10 mg mL^{-1}, NEB) and sterile deionised water to a final volume of 20 µL. Reactions were incubated for 5 hours at 37°C, followed by 65°C for 20 minutes for heat inactivation of the restriction enzyme. Restriction products were analysed using gel electrophoresis (1.5% agarose) and visualised as per section 4.2.1; and bands were excised and extracted following the Nucleospin® Extract II gel extraction protocol (Macherey-Nagel), except that products were eluted with only 10 µL.

A secondary digest was carried out in reactions containing 5 µL and 10 µL of the purified primary digestion reaction products for *TaDMC1* and *TaASY1*/*TaHORMA*, respectively, 2 µL Buffer 2 or Buffer 4 (NEB) for *TaDMC1* and *TaASY1*/*TaHORMA*, respectively, 1 µL *Bam*HI or *Nde*I (NEB) for *TaDMC1* and *TaASY1*/*TaHORMA*, respectively, 0.2 µL of BSA for *TaDMC1* (NEB) and water to a final volume of 20 µL. Reactions were incubated for 5 hours at 37°C before gel electrophoresis (1.5% agarose) and visualisation as described previously in section 4.2.1, followed by band excision and nucleic acid extraction using the Nucleospin® Extract II gel extraction protocol as per the manufacturer's instructions (Macherey-Nagel).

4.2.4 – Preparation of pGBKT7 bait vector

Linearisation of the bait vector pGBKT7 (Appendix 4.3) was conducted as per digestions conducted in section 4.2.3; with restriction with *Eco*RI and *Bam*HI for *TaDMC1*:

pCR®8/GW/TOPO®, and *Nde*I and *Pst*I for *TaASY1* or *TaHORMA*:pGEM®-T Easy. Dephosphorylation of the pGBKTK7 bait vector was carried out in a reaction (8.9 µL) containing 50 ng of linearised pGBKT7 vector, 0.9 µL of 10x dephosphorylation buffer (Roche), 1 µL of shrimp alkaline phosphatase (1U, Roche), with sterile deionised water to final volume.

4.2.5 – *TaDMC1*, *TaASY1* and *TaHORMA* ligation into the pGBKT7 yeast two-hybrid bait vector

To 50 ng of the linearised-dephosphorylated pGBKT7 vector (section 4.2.4), 150 ng of digested fragments (section 4.2.3), 10 µL of 2x T4 Buffer (Roche) and 1 µL of T4 DNA Ligase (Roche) was added and incubated for 3 hours at room temperature, followed by 5 hours at 4°C. Transformation of the recombinant pGBKT7 bait vectors (Appendix 4.4) was as per Chapter 3 section 3.2.6.1.2. Resulting colonies were cultured in 5 mL LB/Spectinomycin (50 µg mL^{-1} as a selective agent) overnight, and plasmids were isolated as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen).

Positive confirmation of the recombinant pGBKT7 vector was conducted via colony PCR, as per section 4.2.2 except that gene specific primers were used. Isolated positive colonies were confirmed via sequence analysis as per section 3.2.1.5, except that oligonucleotides T7 (Chapter 3, section 3.2.1.5) or 3′ DB sequence primer (5′-TTTTCG TTTTAAAACCTAAGAGTC-′3) were used.

4.2.6 – Yeast two-hybrid bait vector toxicity and self activation test

Yeast strains Y187 and AH109 (Clontech) were cultured and transformed with the *TaDMC1*:pGBKT7 bait vector as per manufacturer‟s protocol (Matchmaker™ library construction and screening kits user manual, Clontech). Transformed cells were plated onto yeast selective dropout media, with an absence of specific amino acids (Table 4.2)

Table 4.2 - Selective dropout media. Yeast transformations with the *TaDMC1*:pGBKT7 bait vector were plated out onto selective dropout media.

Plate name	Amino acids absent
SD-1	Tryptophan
SD-4	Adenine Histidine Leucine Tryptophan

4.2.7 – *Ta*HORMA yeast-two-hybrid screen

The cDNA libraries, oligo-dT leptotene to pachytene and random primer pre-meiosis to pachytene, were made as per the manufacturer‟s protocol (Matchmaker™ library construction and screening kits user manual) (Boden and Able *et al. unpublished data*). Library vector, pGADT7, can be viewed in Appendix 4.5.

4.2.7.1 – *TaHORMA*:pGBKT7 yeast-two-hybrid mating

The separate matings between the yeast Y187 strain containing the *TaHORMA*:pGBKT7 bait vector and the yeast AH109 strains with a meiotic cDNA library, either a mix of the oligo-dT and random primer libraries or an oligo-dT library only, was conducted as per the Matchmaker™ library construction and screening kit user manual (Clontech).

4.2.7.2 – Amplification of HORMA interactors

DNA was isolated from the blue positive yeast colony growths using the Yeast DNA extraction kit as per the manufacturer‟s protocol (Pierce, USA). The *Ta*HORMA interactors were then amplified from the yeast DNA in reactions (25 µL) containing 1µL of yeast DNA, 12.5 µL of Failsafe buffer G (Epicenter, USA), 0.5 µL of ADLD F oligonucleotide (50 µM, 5′-CTATTCGATGATGAAGATACCCCACCAAACC-′3), 0.5 µL of ADLD R oligonucleotide (50 µM, 5′-GTGAACTTGCGGGGTTTTTCAGTATCTACGATT-′3), 0.5 µL Platinum® *Taq* Polymerase (5 U/µL, Invitrogen) and 10 µL of sterile deionised water.

PCR cycles were initiated at 94°C for 2 minutes, followed by 35 cycles of 94°C for 30 seconds, 68°C for 6 minutes; with a final extension cycle at 72°C for 10 minutes. PCR reactions (17 µL) were analysed via gel electrophoresis (1.5% agarose) and visualised as described in section 4.2.1. PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per the manufacturer's instructions (Macherey-Nagel).

4.2.7.3 – Restriction digestion of HORMA interactor products

Restriction analysis was conducted on amplified products for identifying unique interactors. Digestion reactions (20 µL) contained 8 µL of the remaining individual PCR mixture (section 4.2.2.2.1), 1 µL of *Hae*III (1U, NEB), 2 µL of buffer 2 (NEB) and 9 µL of sterile deionised water. Reactions were incubated at 37°C for 4 hours followed by enzyme heat inactivation at 80°C for 20 minutes. Reactions were then analysed via gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1.

4.2.7.4 – Ligation and confirmation of *Ta*HORMA interactor fragments in pGEM®-T Easy vector

Ligation of the isolated *TaHORMA* interacting fragments (section 4.2.7.2) into the pGEM®-T Easy vector and transformation was conducted as per Chapter 3, section 3.2.1.3. White colonies were screened via PCR for the validation of the presence of the HORMA interacting fragments. Each reaction (20 µL) contained 2 µL of 10x ImmoBuffer (Bioline), 0.8 µL of $MgCl_2$ (Bioline, 50 mM), 3.2 µL of dNTP solution (1.2 mM each), 1.0 µL of M13F oligonucleotide (10 µM; Chapter 3, section 3.2.6.1.4), 1.0 µL of M13R oligonucleotide (10 µM; Chapter 3, section 3.2.6.1.4), 0.2 µL of Immolase™ DNA Polymerase (Bioline, 5 U/µL) and 11.8 µL of sterile deionised water. Individual white colonies were isolated with a sterile pipette tip and dipped into a single PCR mixture (20 µL) before being placed into 5 mL liquid aliquots of LB/Ampicillin (100 µg mL^{-1} as a selective agent) for overnight growth and generation of glycerol stocks. PCR cycling and visualisation was conducted as per Chapter 3

section 3.2.1.4, except that an annealing temperature of 50°C and an extension of 2 minutes were used.

Positive plasmids were isolated as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen) for verification of insert via sequence analysis. Sequence confirmation of the *Ta*HORMA interacting fragments within the pGEM®-T Easy vector was conducted as per Chapter 3, section 3.2.1.5 using the oligonucleotides SP6 (Chapter 3, section 3.2.1.5) and M13F (Chapter 3, section 3.2.6.1.4). Obtained sequences were used to interrogate the public domain databases for sequence homologues using the Basic Local Alignment Search Tool (tBLASTx; http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 162.0)

4.2.7.5 – Restriction of *TaHORMA* interactors from the pGEM®-T Easy vector for the production of *TaHORMA* interactor prey vectors

The *TaHORMA* interactors were digested from the pGEM®-T Easy vector for ligating into the pGADT7 library vector for confirmation of the interactions between the *TaHORMA* domain and/or *TaASY1*. Primary digestion reactions (20 µL) contained 14 µL of the individual *TaHORMA* interactor:pGEM®-T Easy vector (section 4.2.2.2.3), 4 µL of 10x buffer 4 (NEB) and 2 µL of *Eco*RI enzyme (2U, NEB). Reactions were incubated for 5 hours at 37°C, followed by enzyme heat inactivation at 65°C for 20 minutes.

Secondary digests (20 µL) contained 10 µL of the primary reaction, 4 µL of 10x buffer 4 (NEB), 0.2 µL of BSA (10 mg mL^{-1}, NEB), 2 µL of *Cla*I enzyme (2U, NEB) and 3.8 µL of sterile deionised water. Reactions were incubated for 5 hours at 37°C followed by heat inactivation of the enzyme at 65°C for 20 minutes. Reactions were then analysed via gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per the manufacturer's instructions (Macherey-Nagel).

4.2.7.6 – Preparation of pGADT7 prey vector for ligation of the *TaHORMA* interators

The pGADT7 library vector was linearised via two digestions steps; with the primary digestion reaction (20 µL) containing 5 µL of the individual *TaHORMA* interactor:pGEM®-T Easy vector (section 4.2.7.5), 4 µL of 10x buffer 4 (NEB) and 2 µL of *Eco*RI enzyme (2U, NEB) and sterile deionised water to final volume. Reactions were incubated for 5 hours at 37°C, followed by enzyme heat inactivation at 65°C for 20 minutes. Secondary digests (20 µL) contained 10 µL of the primary reaction, 4 µL of buffer 4 (NEB), 0.2 µL of BSA (10 mg mL^{-1}, NEB), 2 µL of *Cla*I enzyme (NEB) and 3.8 µL of sterile deionised water. Reactions were incubated for 5 hours at 37°C followed by heat inactivation of the enzyme at 65°C for 20 minutes. Reactions were then analysed via gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per the manufacturer‟s instructions (Macherey-Nagel). The purified linearised vector was dephosphorylated as per section 4.2.4.

4.2.7.7 – Ligation, transformation and subsequent verification of the *TaHORMA* interactor:pGADT7 vectors

The ligation, transformation and verification of the *TaHORMA* interactors in the pGADT7 prey vector was conducted as per section 4.2.1.1.3, except that oligonucleotides ADLD F (section 4.2.7.2) and ADLD R (section 4.2.7.2) were used for the colony PCR, and ampicillin (100 µg mL^{-1}) used as the selective agent.

4.2.7.8 – Co-transformation of yeast for interaction confirmation

Co-transformation of the yeast strain AH109 with the *TaHORMA* interactor:pGADT7 prey and either *TaHORMA*:pGBKT7 or *TaASY1*:pGBKT7 bait vectors was conducted as per the Matchmaker™ library construction and screening kits user manual (Clontech). A negative

control transformation of the AH109 strain was conducted with only the *TaHORMA* interactor:pGADT7 vector for detecting false positive *Ta*HORMA interactors.

From positive confirmed interactions, DNA was extracted as per section 4.2.7.2. The *Ta*HORMA interactor proteins (*Ta*HIPs) were amplified in PCR reactions (25 µL) containing 1 µL of isolated DNA, 12.5 µL of Failsafe buffer G (Epicentre, USA), 2.5 µL of ADLD F oligonucleotide (10 µM, section 4.2.7.2.), 2.5 µL of ADLD R oligonucleotide (10 µM, section 4.2.7.2), 0.5 µL of Platinum® *Taq* Polymerase (5 U/µL) and 6 µL of sterile deionised water. PCR was initiated at 94°C for 2 minutes followed by 35 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 68°C for 1 minute with a final extension cycle at 68°C for 10 minutes. Reactions were analysed via gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. PCR products were excised and purified using the Nucleospin® Extract II gel extraction protocol as per the manufacturer's instructions (Macherey-Nagel).

Extracted *TaHIP* products were ligated into the pCR8®/GW/TOPO® vector (Invitrogen) and transformed as per the manufacturer's protocol. Positive colonies were determined via colony PCR in reactions (20 µL) containing 2 µL of 10x PCR buffer (Invitrogen), 0.75 µL $MgCl_2$ (50 mM, Invitrogen), 3.2 µL of dNTP solution (2.5 mM), 0.5 µL of ADLD F oligonucleotide (10 µM, section 4.2.7.2), 0.5 µL of ADLD R oligonucleotide (10 µM, section 4.2.7.2), 0.2 µL of Platinum® *Taq* Polymerase (5 U/µL) and 12.85 µL of sterile deionised water. Individual colonies were isolated using a sterile pipette tip, swirled in the PCR mixture and cultured overnight in LB/Spectinomycin (50 µg mL^{-1} as a selective agent).

PCR was initiated at 94°C for 2 minutes followed by 35 cycles of 94°C for 30 seconds, 50°C for 30 sections and 68°C for 1 minute with a final extension cycle at 68°C for 10 minutes. Reactions were analysed via gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. Plasmids were isolated as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen). The *Ta*HIP fragments were further confirmed within the pCR8®/GW/TOPO® vector via sequence analysis as per section 4.2.2.

4.2.8 – Isolation of full length HORMA interacting proteins

4.2.8.1 – *In silico* sequence searching

Confirmed sequences obtained for *TaHIPs* were used to interrogate the public domain database for corresponding wheat ESTs. Using the Basic Local Alignment Search Tool (tBLASTx; http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 172.0), a homologous rice sequence was obtained where possible (*E*-value ≥ -20) for determining the comparative rice open reading frame (ORF) size. The rice ORF was then used in tBLASTx searches for finding corresponding wheat ESTs over the whole ORF (*E*-value ≥ -20). The wheat ESTs were then compiled in Contig Express, where the wheat ORFs were predicted. ORFs were confirmed using the tBLASTx program (http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 172.0). Oligonucleotides for the isolation of *TaHIPs* were designed off the *in silico* wheat sequences.

4.2.8.2 – RACE isolation of flanking regions

In instances where complete rice and wheat sequences were not identified from *in silico* sequence searching (section 4.2.8.1); the rapid amplification of cDNA ends (RACE, GeneRacer™ Kit, Invitrogen) technique was utilised for isolating the full length sequences (*TaHIP1*). Oligonucleotides for the isolation of *TaHIP1* were designed from the *TaHIP1* sequence confirmed from initial yeast mating screens (section 4.2.7.4). RACE reverse transcriptase (RT) template cDNA libraries for 5′ and 3′ isolation were made from RNA (Chapter 2, section 2.2.4) as per the RACE library construction protocol (GeneRacer™ Kit user manual, Invitrogen).

RACE primary amplification of the 5′ and 3′ ends of *TaHIP1* was attempted in reactions (50 µL) containing 3 µL of GeneRacer 5′ oligonucleotide or GeneRacer 3′ oligonucleotide (Invitrogen), 1 µL of either *TaHIP1*_5′ (10 µM, 5′-GTTGCGCTGAGGCT CCTTGCGATTCCGT-3′) or *TaHIP1*_3′ (10 µM, 5′-AGCGACAGAGCACCCCCCAGGCA

GATGTAG-3′) oligonucleotides for each GeneRacer oligonucleotide, 1 µL of RT template, 5 µL of 10x PCR buffer (Invitrogen), 4 µL of dNTP solution (2.5 mM each), 2 µL $MgSO_4$ (50 mM, Invitrogen), 0.5 µL of Platinum® *Taq* (5 U/µL, Invitrogen), and 33.5 µL of sterile deionised water. Primary PCR was initiated for 2 minutes at 94°C; followed by 5 cycles of 94C° for 30 seconds and 70°C for 2 minutes, 5 cycles of 94°C for 30 seconds and 68°C for 2 minutes, and 30 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 68°C for 2 minutes with a final extension cycle of 10 minutes at 72°C.

Secondary RACE PCR was conducted with reactions (50 µL) containing 1 µL of GeneRacer 5′_nested oligonucleotide (Invitrogen), 1 µL of either *TaHIP1*_5′_nested (10 µM, 5′-ACCATGCTGGTCTTCCTGGTCCGCTTGTCC-3′) or *TaHIP1*_3′_nested (10 µM, 5′-GT CCCCCCTGAGACGTGTTGTTGTTCATGATG-3′) oligonucleotides, 1 µL of the Primary PCR reaction, 5 µL of 10x PCR buffer (Invitrogen), 4 µL of dNTP solution (2.5 mM each), 2 µL $MgSO_4$ (50 mM, Invitrogen), 0.5 µL of Platinum *Taq* (5 U/µL, Invitrogen), and 35.5 µL of sterile deionised water. Secondary PCR was initiated for 2 minutes at 94°C followed by 35 cycles of 94C° for 30 seconds, 60°C for 30 seconds and 68°C for 2 minutes with a final extension cycle of 10 minutes at 72°C. Primary and secondary reactions were analysed via gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. Band products from either the primary or secondary RACE PCR were excised using the Extract II gel extraction protocol as per the manufacturer‟s instructions (Macherey-Nagel).

4.2.8.3 – Ligation, transformation and subsequent confirmation of *TaHIP1* RACE products

The isolated RACE fragments of *TaHIP1* were ligated into the pCR8®/GW/TOPO® vector and transformed as per Chapter 3, section 3.2.6.1.2. Colony PCR and sequencing confirmation was conducted as per section 4.2.1.1.1.3. Analysis of isolated sequences was conducted using Contig Express (Invitrogen), and the BLASTn program for sequence alignments with the previously obtained *TaHIP1* sequence.

4.2.8.4 – Amplification of full length *TaHIP3 and TaHIP4*

The full length genes of *TaHIP3* and *TaHIP4* were isolated from wheat in PCR reactions (25 µL) containing 2.5 µL of 10x PCR Buffer (Invitrogen), 0.75 µL of $MgCl_2$ (50 mM), 1.6 µL of dNTP solution (2.5 mM each), 1 µL of cDNA template (Chapter 2, section 2.2.4), 1 µL of either *TaHIP3*_F (10 µM, 5′-GACTGCTCTCACCAAGGCGGAG-3′) or *TaHIP4*_F (10µM, 5′-CGAGCGAGCGAGGCTCCATC-3′), 1 µL of either *TaHIP3*_R (10 µM, 5′-GTGCATACAGGATCATCAGACGCA-3′) or *TaHIP4*_R (10 µM, 5′-TCGATAAC ACGCACGAGAGATCCA-3′), 0.5 µL of Platinum® *Taq* Polymerase (5 U/µL, Invitrogen) and 16.65 µL of sterile deionised water. PCR was initiated at 94°C for 2 minutes, followed by 35 cycles of 30 seconds at 96°C, 30 seconds at 60°C, 1 minute at 68°C, with a final extension cycle of 10 minutes at 68°C. Post PCR reactions were analysed by gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. PCR products were excised and purified using the Invitrogen PureLink™ Gel Extraction Kit as per the manufacturer's protocol.

4.2.8.5 – Ligation, transformation and confirmation of *TaHIP3 and TaHIP4* pCR8®/GW/TOPO® vectors

The amplified and isolated PCR fragments of *TaHIP3* and *TaHIP4* were ligated into the pCR8®/GW/TOPO® vector, with transformation of competent cells conducted as per Chapter 3, section 3.2.6.1.2. Colony PCR and sequencing confirmation were conducted as per section 4.2.2.

4.2.9 – Comparative sequence analysis

4.2.9.1 – Homologous sequence search analysis

The obtained open reading frames of *TaHIP3* and *TaHIP4* were used to interrogate the public database (BLASTx, http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 174.0) for homologous sequences.

4.2.9.2 – Conserved domains

Using the predicted translations of the ORFs of *TaHIP3* and *TaHIP4*, the BLAST Conserved Domain tool of the NCBI website (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi NCBI-GenBank Flat File Release 174.0) was queried to detect any possible conserved protein domains.

4.2.9.3 – Full length sequence alignments

ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) was used as a tool for the comparison alignment of amino acid sequences of *Ta*HIP3 along with the orthologues from other species.

4.2.9.4 – Phylogenetic analysis of *TaHIP3* and *TaHIP4*

The evolutionary history of *TaHIP3* and *TaHIP4* was inferred using the phylogenetic Neighbour-Joining method. MEGA software (version 4.0, Molecular Evolutionary Genetics Analysis) (Tamura *et al.*, 2007) with default setting parameters was used except for the following: the "pair-wise deletion" was used for the amino acid sequences, the bootstrap value was set at 10,000 re-samplings and the model setting was Amino: p-distance.

ORFs from orthologous sequences of *TaHIP3* and *TaHIP4* (Table 4.3) obtained in section 4.2.5.1 were used for the construction of phylogenetic trees, with the program directed to translate the nucleotide sequence into all six frames prior to multiple alignments for analysis using the in-built ClustalW function.

Table 4.3 - *TaHIP3* and *TaHIP4* orthologues identified and used for phylogenetic analysis. *At = Arabidopsis thaliana, Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sb = Sorghum bicolour*, *Sc = Saccharomyces cerevisiae, Zm = Zea mays*.

Candidate	Species	Nucleotide accession
TaHIP3	*Os*	AB111915
	Mm	NM_026632
	Hs	NM_002947
	Gg	XM_418679
	Sc	NC_001136
TaHIP4	*Os*	NM_001066487
	Sb	XM_002462961
	Zm	BT038780
	At	NM_101011

4.2.9.5 – Three dimensional modelling of *Ta*HIP3 and *Ta*HIP4 protein structures

The CPHmodel server, provided by the Centre for Biological Sequence Analysis (CBS) (http://www.cbs.dtu.dk/services/CPHmodels), was used for computational 3D modelling of the amino acid sequences in Table 4.4. The subsequent output was then visualised using MacPyMol (DeLano Scientific LLC) to generate the 3D structure for each protein.

Table 4.4 - Protein accessions used for three dimensional modelling of *Ta*HIP3 and *Ta*HIP4. *At = Arabidopsis thaliana, Hs = Homo sapiens, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

Candidate	Species	Protein accession
*Ta*HIP3	*Os*	BAD06873
	Hs	NP_002938
	Sc	NP_010440
TaHIP4	*Os*	NP_001059952
	At	NP_563889

4.2.10 – Expression analysis of *TaHIP3* and *TaHIP4*

The expression of *TaHIP3* and *TaHIP4* during meiosis was investigated using the expression data set previously obtained by Crismani *et al*. (2006); along with the interacting gene *TaASY1*. Expression of *TaHIP3* and *TaHIP4* was further examined in *Taasy1* knock down mutants (Boden *et al.*, 2009) and wild-type wheat (Bob white 26 cultivar) anthers via

the Invitrogen SuperScript™ One-Step RT-PCR with Platinum® *Taq* kit. RNA used for the expression analysis RT-PCR reactions was provided by S. Boden (Boden, 2008).

Oligonucleotides used for the amplification of the transcribed products were as follows: *TaHIP3*_RT-PCR_F (5′-GACTGCTCTCACCAAGGCGGAG-3′), *TaHIP3*_RT-PCR_R (5′-GTGCATACAGGATCATCAGACGCA-3′), *TaHIP4*_ RT-PCR_F (5′-CGAGCG AGCGAGGCTCCATC-3′), *TaHIP4*_RT-PCR_R (5′-TCGATAACACGCACGAGAGATCC A-3′); and control primers *TaGapDH*_RT-PCR_F (5′-TTCAACATC ATTCCAAGCAGCA-3′) and *TaGapDH*_RT-PCR_R (5′-CGTAACCCAAAATGCCCTTG-3′) (kindly provided by Neil Shirley, ACPFG, Adelaide, AUS).

The one-step RT-PCR reactions (20 µL) contained 350 ng of RNA, 12.5 µL of 2x Reaction Mix (Invitrogen), 0.5 µL of corresponding forward oligonucleotide (10 µM), 0.5 µL of corresponding reverse oligonucleotide (10 µM), 0.5 µL of RT/Platinum® *Taq* Mix (Invitrogen) and sterile deionised water to the final volume. Conditions for the cDNA synthesis and subsequent PCR amplification were as follows: 30 minutes at 50°C, 2 minutes at 94°C, 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 68°C for 30 seconds, with a final extension cycle of 10 minutes at 72°C. Reactions were analysed by gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1.

4.2.11 – Heterologous *Ta*HIP3 and *Ta*HIP4 protein expression

4.2.11.1 – Protein expression vector preparation

4.2.11.1.1 – Amplification of *TaHIP3* and *TaHIP4* ORFs

The ORFs of *TaHIP3* and *TaHIP4* were amplified in PCR reactions (20 µL) containing 2 µL of 10x PCR Buffer (Invitrogen), 0.75 µL of $MgCl_2$ (50 mM, Invitrogen), 1.6 µL of dNTP solution (2.5 mM each), 1 µL of *TaHIP3*_ATG oligonucloetide (10 µM, 5′-ATGGATACTTCAGCTCCTTCACCATTTG-3′) or *TaHIP4*_ATG oligonucloetide (10 µM, 5′-ATGGCCGCGGCGGCGGAGGA-3′), 1 µL of *TaHIP3*_STOP oligonucleotide (10 µM, 5′-

CTACAGGAACAGGTACTTGTACTTGTCATTCG-3′) or *TaHIP4*_STOP oligonucleotide (10 µM, 5′-TTACTCCTGCTCTGGTTGCGCTGA-3′), 1 µL of *TaHIP3*:pCR8®/GW/TOPO® or *TaHIP4*:pCR8®/GW/ TOPO® vectors (section 4.2.4.4) as template, 0.2 µL of Platinum® *Taq* Polymerase (5 U/µL, Invitrogen) and 12.45 µL of sterile deionised water. PCR was initiated at 94°C for 2 minutes followed by 30 cycles of 30 seconds at 96°C, 30 seconds at 60°C and 1 minute at 68°C; with a final extension cycle at 68°C for 10 minutes. Post-PCR reactions were analysed by gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1.

4.2.11.1.2 – Ligation, transformation and subsequent confirmation of ORF pCR8®/GW/TOPO® vectors

The PCR reaction mixture of the *TaHIP3* and *TaHIP4* ORFs (section 4.2.11.1.1) were used in a ligation reaction with the pCR8®/GW/TOPO® vector and subsequently transformed into competent cells as per Chapter 3, section 3.2.6.1.2. Colony PCR and sequencing confirmation were conducted as per section 4.2.2.

4.2.11.1.3 – Transfer of ORFs into pDEST17 protein expression vector and subsequent confirmation

The transfer of the*TaHIP3* and *TaHIP4* ORFs from the pCR8®/GW/ TOPO® vector into the protein expression vector pDEST17 was conducted as per Chapter 3, section 3.2.6.1.5. Confirmation of the recombinant vectors was conducted via PCR and sequence analysis. Colony PCR reactions (20 µL) contained 2 µL of 10x PCR Buffer (Invitrogen), 0.75 µL of $MgCl_2$ (50 mM, Invitrogen), 1.6 µL of dNTP solution (2.5 mM each), 1 µL of T7 oligonucleotide (10 µM; Chapter 3, section 3.2.1.5), 1 µL of *TaHIP3*_STOP or *TaHIP4*_STOP (10 µM, section 4.2.7.1.1), 0.2 µL of Platinum® *Taq* Polymerase (5 U/µL, Invitrogen) and 13.45 µL of sterile deionised water. PCR was initiated at 94°C for 2 minutes followed by 35 cycles of 30 seconds at 96°C, 30 seconds at 50°C and 1 minute at 68°C; with

a final extension cycle at 72°C for 10 minutes. Reactions were analysed by gel electrophoresis (1.5% agarose) and visualised as described previously in section 4.2.1. Positive plasmids were isolated as per the Invitrogen PureLink™ Quick Plasmid Miniprep Kit handbook (Invitrogen). Confirmation of the *TaHIP3* and *TaHIP4* ORFs being in-frame in the pDEST17® vector was conducted via sequence analysis (Chapter 3, section 3.2.6.1.7).

4.2.11.2 – Heterologous protein expression

The *TaHIP3*_ORF:pDEST17® and *TaHIP4*_ORF:pDEST17® protein expression vectors were transformed into BL21-AI protein expression optimised cells as per Chapter 3, section 3.2.6.2.1. Four overnight starter cultures of the BL21-AI protein expression cells were commenced by inoculating 200 µL of cells into 8 mL of LB/Carbenicillin (50 µg mL^{-1}) at 37°C with agitation. The following day, four 200 mL LB/Carbenicillin (50 µg mL^{-1}) cultures were each inoculated with a 8 mL starter culture, and were subsequently incubated at 37°C with agitation until an OD_{600} measurement of 0.4 was reached. Upon recording an OD_{600} measurement of 0.4, L-(+)-arabinose was added to 0.4% w/v to two of the cultures for induced samples; with D-glucose added to 0.5% w/v to the other two cultures for repressed samples. Cultures were then incubated at either 23°C or 37°C for 5 hours with agitation, followed by cell collection at 3000 *g* for 15 minutes at 4°C for cell pelleting. The resulting supernatant was discarded and the cell pellets were snap-frozen using liquid nitrogen and stored at -80°C until protein extraction.

4.2.11.3 – Heterologous protein extraction

Both *Ta*HIP3 and *Ta*HIP4 proteins were natively extracted from the cell pellets as per Chapter 3, section 3.2.6.3.1.1 except that: 1) resuspension of the cell pellet was in 20 mL of lysis buffer; 2) no selective ammonium sulphate precipitation was conducted; 3) upon completion of protein gel electrophoresis, the gels were placed in a fixing solution of 25% isopropanol and 10% acetic acid for 1 hour with gentle agitation, which was followed by

staining in 10% acetic acid and 0.006% Brilliant Blue R250 (Sigma) at room temperature overnight with gentle agitation. The protein gel was then destained in 10% acetic acid at room temperature until staining was optimal.

4.3 – Results

4.3.1 – Primary *TaDMC1*, *TaASY1* and *TaHORMA* vector preparations

Isolation of the *TaDMC1*, *TaASY1* and *TaHORMA* coding regions with oligonucleotides containing the restriction sites was successful (Figure 4.3). Ligation of the isolated products into the pCR®8/GW/TOPO® and pGEM®-T Easy vectors was verified through colony PCR (Figure 4.4) and sequence analysis; with positive colonies 5, 4 and 1 for *TaDMC1*, *TaASY1* and *TaHORMA*, respectively, being used for further experiments.

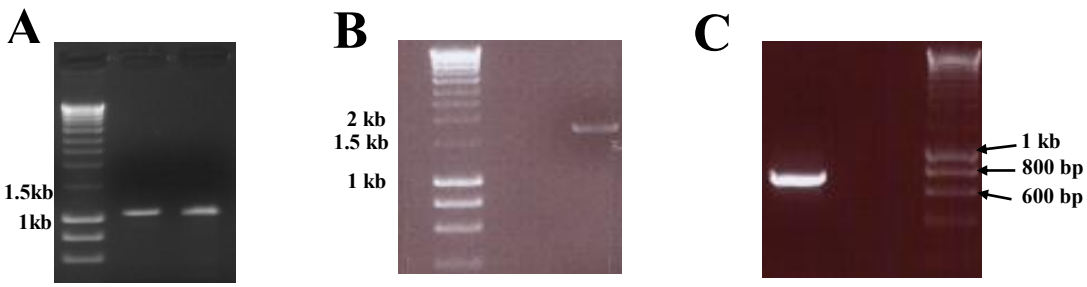


Figure 4.3 - Amplification of wheat *DMC1*, *ASY1* and *HORMA* domain coding regions. Using oligonucleotides with restriction sites, the coding regions for *TaDMC1* (**A**), *TaASY1* (**B**) and *TaHORMA* (**C**) was isolated. A 1 kb DNA molecular weight marker is shown in each, but not all sizes are shown.

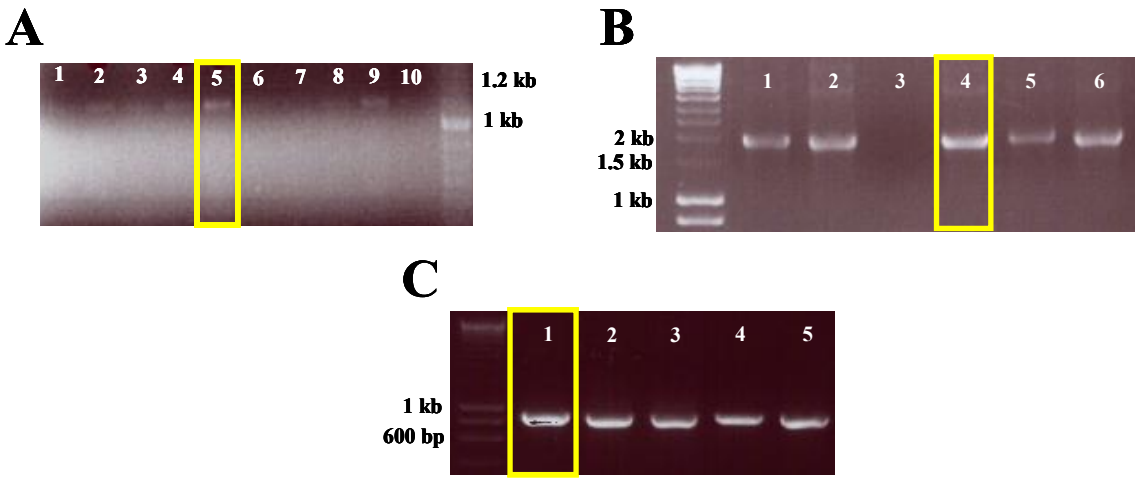


Figure 4.4 - Colony PCR of *TaDMC1*:pCR®8/GW/TOPO® and *TaASY1*/*TaHORMA*:pGEM®-T Easy colonies. Putative positive colonies were screened for the *TaDMC1* (**A**), *TaASY1* (**B**) and *TaHORMA* (**C**) inserts, with positive colonies being confirmed via sequence analysis (*yellow* boxes) and used for further experimentation. A 1 kb DNA molecular weight marker is present in each image, but not all sizes are shown.

4.3.2 – Linerisation of pGBKT7 bait vector and restriction of *TaDMC1*, *TaASY1* and *TaHORMA* from primary vectors

Restriction digestion of the *TaDMC1*:pCR®8/GW/TOPO® and *TaASY1*/*TaHORMA*:pGEM®-T Easy vectors with enzymes *Eco*RI/*Bam*HI and *Pst*I/*Nde*I, respectively, resulted in the excision of the *TaDMC1*, *TaASY1* and *TaHORMA* fragments (Figure 4.5) enabling directional ligation into similarly digested pGBKT7 bait vectors (Figure 4.6).

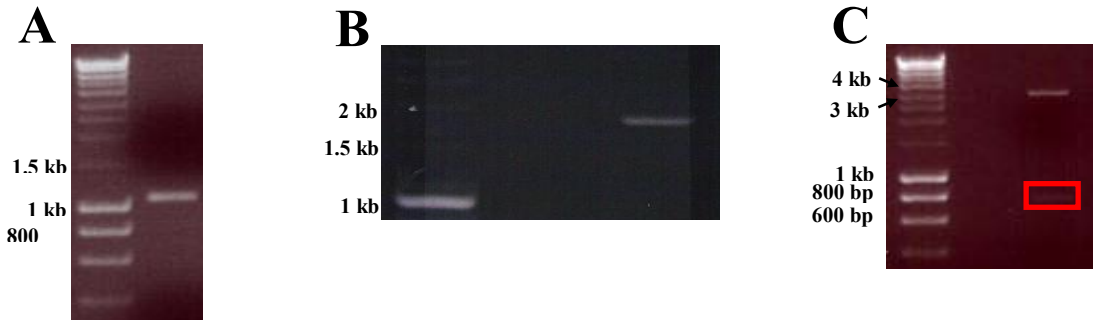


Figure 4.5 - Isolation of *TaDMC1*, *TaASY1* and *TaHORMA* fragments via enzyme restriction. Digestion with the restriction digest enzymes *Eco*RI/*Bam*HI and *Pst*I/*Nde*I resulted in the excision of *TaDMC1* (**A**), *TaASY1* (**B**) and *TaHORMA* (**C**) from their primary vectors pCR®8/GW/TOPO® and pGEM®-T Easy. A 1 kb DNA molecular weight marker is shown in all images, but not all sizes are shown.

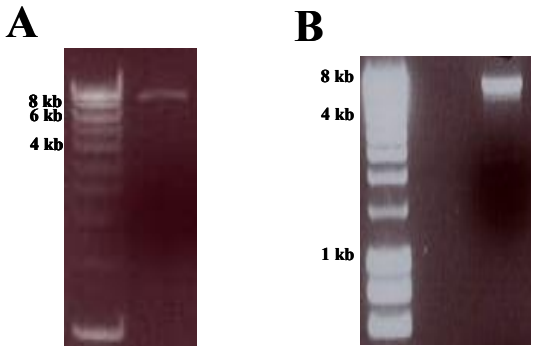


Figure 4.6 - Linerisation of pGBKT7 bait vector with *Eco*RI/*Bam*HI and *Pst*I/*Nde*I. Restriction digestion of the pGBKT7 yeast two-hybrid bait vector with either the enzymes *Eco*RI/*Bam*HI for *TaDMC1* (**A**) or with enzymes *Pst*I/*Nde*I for *TaASY1* and *TaHORMA* (**B**); enabling directional ligation of these fragments into pGBKT7. A 1 kb DNA molecular weight marker is shown in both images, but not all sizes are shown.

4.3.3 – Construction of *TaDMC1*, *TaASY1* and *TaHORMA* pGBKT7 yeast two-hybrid bait vectors and toxicity testing

Ligation and transformation of the *TaDMC1*, *TaASY1* and *TaHORMA*:pGBKT7 vectors were successful as the candidate recombinants were confirmed either through colony PCR or restriction digestion (Figure 4.7) and shown to be in the correct orientation within the pGBKT7 vectors through sequence analysis. The constructed bait vectors were all found to be non-toxic, with growth occurring on selective media plates in the absence of tryptophan (Figure 4.8 A).

However the *TaDMC1*:pGBKT7 bait vector was also found to be self activating, with growth being observed on the selective media plates in the absence of the amino acids adenine, leucine, histidine and tryptophan (Figure 4.8 B). Therefore, given the self-activation of the *TaDMC1* yeast bait vector, further analysis of protein interactors of DMC1 was discontinued.

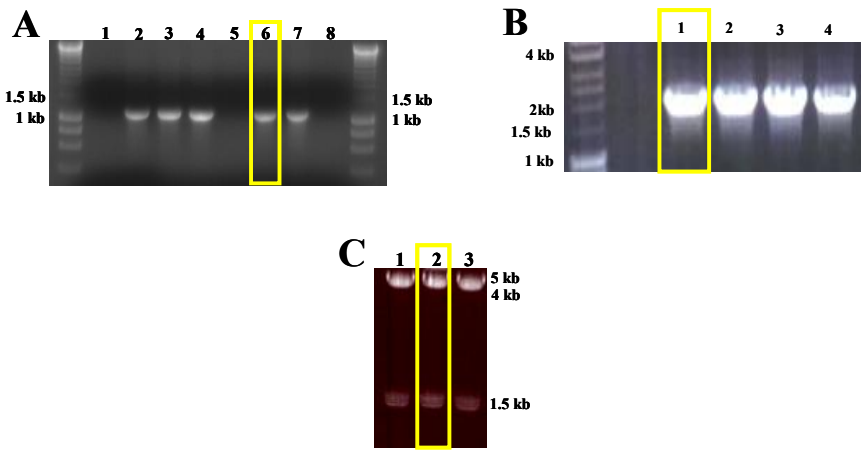


Figure 4.7 - Confirmation of *TaDMC1*, *TaASY1* and *TaHORMA* fragments within the pGBKT7 bait vectors. *TaDMC1* (**A**) and *TaASY1* (**B**) were confirmed within the pGBKT7 bait vectors (yellow boxes) via colony PCR. (**C**) Three individual colonies of *TaHORMA*:pGBKT7 were digested using the *Hind*III enzyme, with all three showing the insert and the correct orientation. Digestion resulted in the expected product at 1498 bp, 1567 bp and 4938 bp. All colonies (boxed) were further confirmed for correct orientation via sequence analysis. A 1 kb DNA molecular weight marker is shown in A and B, while comparative sizes of a 1 kb DNA molecular weight marker are labelled on the side of C, but not all sizes are shown.

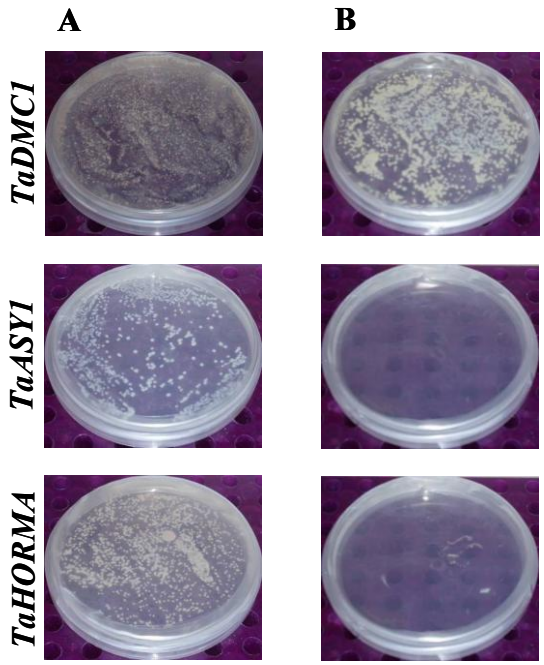


Figure 4.8 - Toxicity and self activation testing of the *TaDMC1*, *TaASY1* and *TaHORMA*:pGBKT7 yeast two-hybrid bait vectors. Yeast colony growth was observed on selective media plates lacking tryptophan (**A**) indicating that the *TaDMC1*, *TaASY1* and *TaHORMA*:pGBKT7 bait vectors are non-toxic. Yeast colony growth was not observed for *TaASY1* and *TaHORMA* on selective media lacking the amino acids adenine, leucine, histidine and tryptophan (**B**, middle and bottom) which indicates that these vectors are not self activating. However the *TaDMC1*:pGBKT7 bait vector allowed yeast colony growth (**B**, top), indicating that there is self activation of the reporter genes by the *TaDMC1* fusion bait protein within the AH109 cells.

4.3.4 – *Ta*HORMA:pGBKT7 yeast two-hybrid screen with meiotic cDNA libraries

Initial mating of the yeast strains Y187, containing the *TaHORMA*:pGBKT7 bait vector, and AH109, containing either only oligo-dT or a mix of oligo-dT and random primer early prophase I cDNA libraries, produced yeast colony growth on the selective dropout media with the absence of the amino acids adenine, leucine, histidine and tryptophan. Positive interactors were further confirmed on the selective media plates containing 5-Bromo-4-Chloro-3-Indolyl-***α***-D-galactopyranoside (X-α-gal, 2 mg mL^{-1}), the substrate for the *β*-galactosidase reporter gene**.** Four of the five colonies from the mixed library screen and three of the four colonies from the oligo-dT library screen were observed to turn blue and therefore deemed as putative positive interactors (results not shown).

4.3.5 – Isolation of *Ta*HORMA interactors

Fragments from the cDNA library vector (pGADT7) that encoded peptides that interacted with *Ta*HORMA were isolated via PCR amplification (Figure 4.8). PCR amplification of *Ta*HORMA interactor oligo-dT #9, identified from the oligo-dT library screen, resulted in two bands which were isolated separately (Figure 4.9, lane 9) and coined oligo-dT #9a (top band) and oligo-dT #9b (bottom band). Restriction analysis of the PCR products identified that two of the oligo-dT cDNA library interactors (#7 and #8) were likely to be the same sequence due to the same restriction pattern (Figure 4.10, lane 7 and 8), therefore only oligo-dT #7 was taken further as a representation of these two particular interactors.

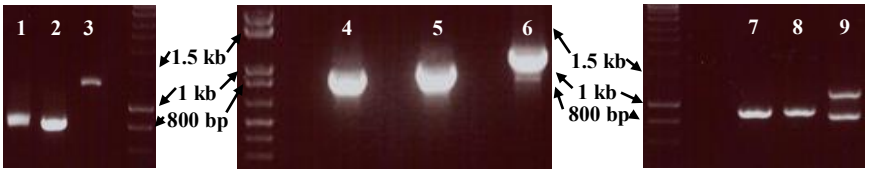


Figure 4.9 - PCR amplification of *Ta*HORMA interactors from yeast DNA extracts. Screening of the mixed oligo-dT and random primer meiotic cDNA libraries resulted in 4 colonies in which the *Ta*HORMA interacting fragments were amplified in lanes 1 through 4. Screening of the oligo-dT meiotic cDNA library resulted in 5 colonies amplified (lanes 5 through 9). Individual fragments were isolated for further confirmation; with the oligo-dT #9 multiple bands termed #9a (higher band) and #9b (lower band). A 1 kb DNA molecular weight marker is shown; but not all band sizes are listed.

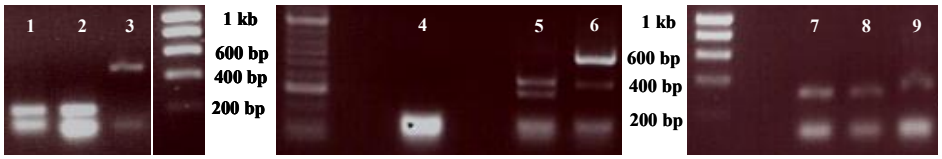


Figure 4.10 - Restriction digest analysis of *Ta*HORMA interactors. *Hae*III restriction analysis of the *Ta*HORMA interactors was conducted to identify unique interactors. The four interactors with the mixed oligo-dT and random primer meiotic cDNA libraries were unique (lanes 1-4). Two of the five interactors with the oligo-dT meiotic cDNA library (lanes 5-9) were found to have the same restriction pattern (lanes 7 and 8). A 1 kb DNA molecular weight marker is shown; but not all band sizes are listed.

Ligation of the unique *Ta*HORMA interactors into the pGEM®-T Easy vector and the subsequent transformations were successful, with multiple colonies obtained (Figure 4.11). Plasmids from two colonies for each interactor were sequenced for the assembly of the full-length fragment sequences, with all but oligo-dT #7 returning sequence information (Appendix 4.6). The sequences of the *Ta*HORMA interactors were then used to interrogate the public domain databases for the identification of putative sequence homologues (Table 4.5) using the Basic Local Alignment Search Tool (tBLASTx).

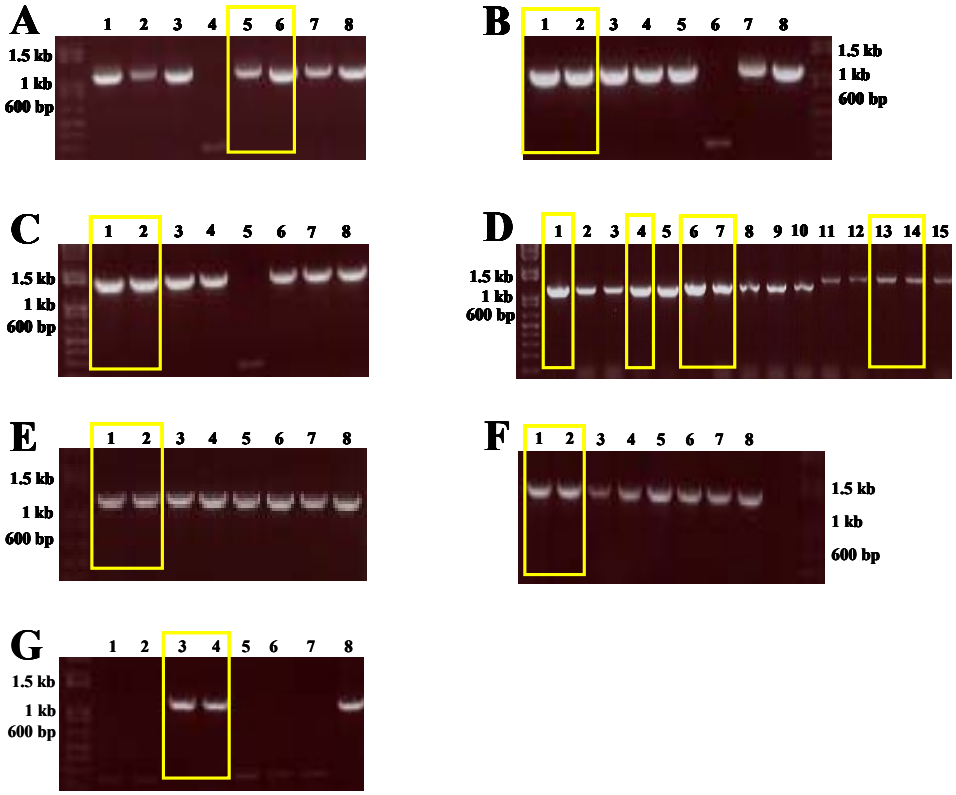


Figure 4.11 - Verification of the *Ta*HORMA interactors:pGEM®-T Easy vectors. Colony PCR of resulting transformants of the *Ta*HORMA interactors in the pGEM®-T Easy vector for the mixed #1 (**A**), mixed #2 (**B**), mixed #3 (**C**), mixed #4 (**D**, lanes 1-5), oligo-dT #5 (**D**, lanes 6-10), oligo-dT #6 (**D**, lanes 11-15), oligo-dT #7 (**E**), oligo-dT #9a (**F**), oligo-dT #9b (**G**). Highlighted yellow boxes in each of the images represent those colonies that were sequenced. A 1 kb DNA molecular weight marker is shown; but not all band sizes are listed.

Table 4.5 - *Ta*HORMA interactor sequence similarity. Nucleotide similarity of *Ta*HORMA interactors from the public database.

***Ta*HORMA interactor**	**Similar sequence homologue**	**E value**	**Identity (%)**	**Similarity (%)**
Mixed #1	*TaASY1*(EF446137)	9e-36	96	96
Mixed #2	*Hv* cDNA clone (AK250947)	2e-38	68	75
Mixed #3	*Ta*HMGB1 (AB272226)	1e-90	98	98
Mixed #4	*Hv* cDNA clone (AK250947)	8e-56	87	92
Oligo-dT #5	*Os*RPA14 (AB111915)	2e-26	63	78
Oligo-dT #6	*Hv* cDNA clone (AK250323)	8e-66	96	100
Oligo-dT #9a	*Os* cDNA clone (AK070185)	2e-105	86	92
Oligo-dT #9b	*Hv* cDNA clone (AK250947)	4e-12	84	89

*Ta*HORMA interactors mixed #2, #4 and oligo-dT #9b were found to have similarity to the same accession, and were therefore considered as one individual interactor. Mixed #2 interactor was taken forwards as the representative for these candidates. The remaining *Ta*HORMA interactors were found to be unique.

4.3.6 – Production of *Ta*HORMA interactor prey vectors for interaction confirmation

The *Ta*HORMA interactors were isolated from the pGEM®-T Easy vector (Figure 4.12 A) and ligated into the pGADT7 prey vector, and subsequently the *Ta*HORMA interactor:pGADT7 plasmids were confirmed (Figure 4.12 B). These vectors were used in co-transformation of yeast for the confirmation of the interaction with the *Ta*HORMA domain and/or *Ta*ASY1.

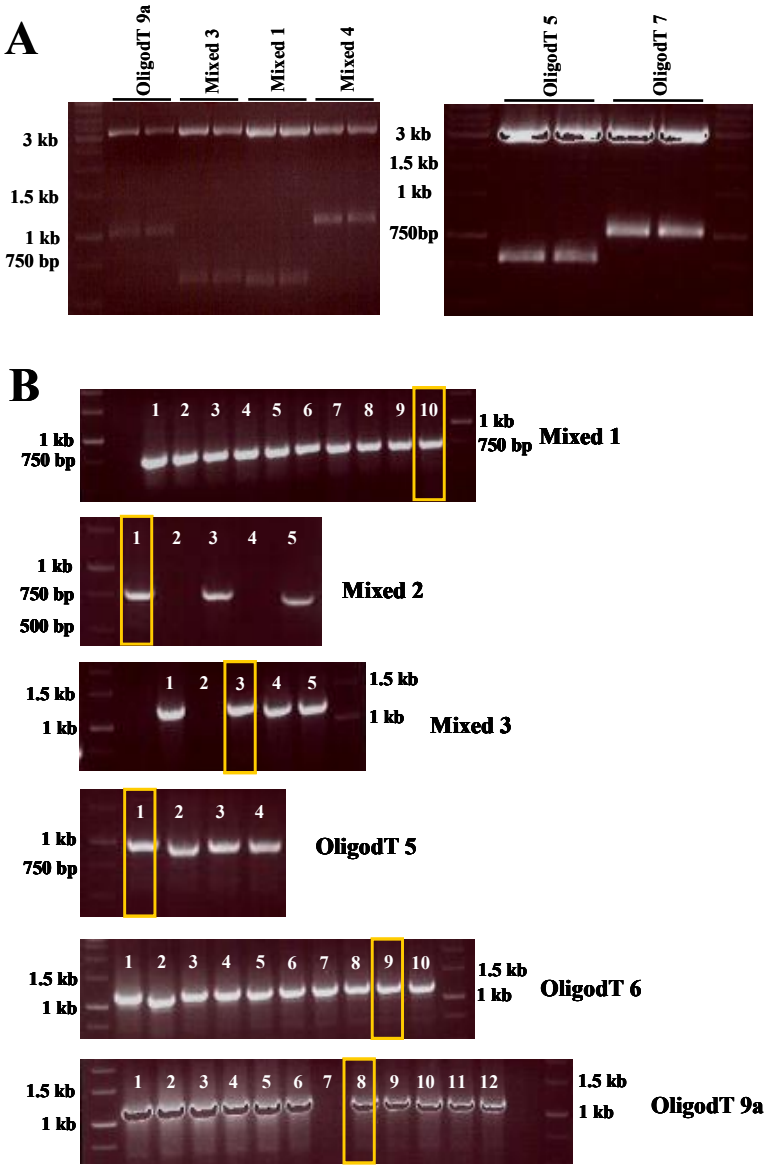


Figure 4.12 - Production of *Ta*HORMA interacting pGADT7 prey vectors. (**A**) Restriction digest analysis of the *Ta*HORMA interactors from the pGEM®-T Easy vectors. (**B**) Colony PCR verification of the *Ta*HORMA interactors in pGADT7. Boxed colonies were isolated and further confirmed through sequence analysis. A 1 kb DNA molecular weight marker is shown; but not all band sizes are listed.

4.3.7 – Co-transformation of the *Ta*HORMA interactor:pGADT7 prey with *Ta*HORMA or *Ta*ASY1:pGBKT7 bait vectors for interaction confirmation

The six proposed *Ta*HORMA interactors were successfully co-transformed with the *Ta*HORMA:pGBKT7 or *Ta*ASY1:pGBKT7 bait vectors into the AH109 yeast strain. An interaction within the yeast cell between the two translated proteins of interest, along with the presence of the vectors themselves, results in cell survival in the absence of the amino acids adenine, histidine, leucine and tryptophan; along with blue colony growth due to the production of β-galactosidase.

The proposed *Ta*HORMA interactors of mixed #1, mixed #2, oligo-dT #6 and oligo-dT # 9b were confirmed to interact with the *Ta*HORMA domain via yeast co-transformation; and with the exception of mixed #2, these interactors were also found to interact with the full length *Ta*ASY1 (Figure 4.13) – with these interactors henceforth being called HORMA Interacting Proteins (HIPs) . The remaining two proposed *Ta*HORMA interactors; mixed #3 and oligo-dT #9a, were found to be false-positives as transformations with only the *Ta*HORMA interactor:pGADT7 vectors resulted in colony growth within the AH109 cells (Figure 4.14). *TaHIPs* which were found to be interacting with both the *Ta*HORMA domain and the full length *Ta*ASY1 protein were isolated for further characterisation.

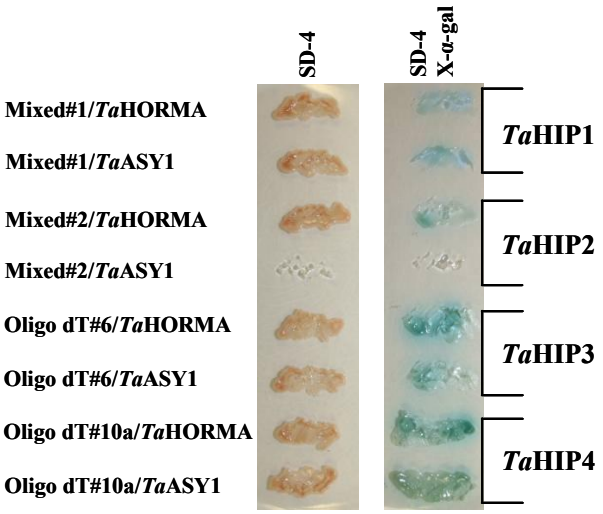


Figure 4.13 - *Ta*HORMA and *Ta*ASY1 interactions with *Ta*HIPs via co-transformation. Three of the four interactors initially found to interact with the *Ta*HORMA domain where confirmed to interact with both the *Ta*HORMA domain and the full length *Ta*ASY1 protein (*Ta*HIP1, *Ta*HIP3 and *Ta*HIP4). *Ta*HIP2 was found only to interact with the *Ta*HORMA domain and not with the full length *Ta*ASY1 protein. SD-4 = selective media without amino acids adenine, histidine, leucine and tryptophan; X-α-gal = 5-Bromo-4-Chloro-3-Indolyl-*α*-D-galactopyranoside.

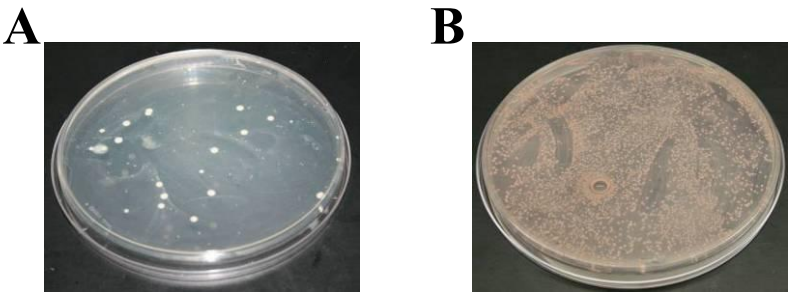


Figure 4.14 - Self activation of mixed #3 and oligo-dT #9a. The proposed *Ta*HORMA interactors of mixed #3 (**A**) and oligo-dT #9a (**B**) were found to be false positives from the initial yeast-two-hybrid screen, as colony growth on the selective media was observed in AH109 yeast, transformed exclusively with the *Ta*HORMA interacting prey vector.

4.3.8 – Isolation of full length *TaHIPs*

4.3.8.1 – *In silico* searching for *TaHIP1*, *TaHIP3* and *TaHIP4*

For the isolation of full length *Ta*HIP proteins, the public database from NCBI was interrogated using the confirmed sequences obtained in section 4.3.3 for the identification of a homologous rice ORF sequence (*E*-value ≥ -20, Table 4.6). These homologous rice ORF sequences were then used to interrogate the wheat EST database for the assembly of a sequence contig which contained the ORF of the *TaHIP* which was of comparable size to the rice homologue (Figure 4.15). *TaHIP1* resulted in no significant hits being found on the databases for a rice homologue, though a non-significant match to a rice *OsASY1* annotation was observed (Table 4.6), which is not unexpected as initial sequence analysis found *TaHIP1* to have similarity to *TaASY1*. Despite recording a significant hit of *TaHIP1* to *TaASY1,* and non-significantly to *OsASY1*, the observed similarities of the sequences were located within a small region of the *TaHIP1* sequence (212 to 307 bp); which leads to the thought that this is an unknown domain, yet no conserved domains were found in any frame translation of *Ta*HIP1.

Table 4.6 - Rice homologue annotations for *TaHIP1*, *3* and *4*. Using the confirmed *TaHIP* sequences obtained in section 4.3.3, rice homologues were identified using the tBLASTx program.

*Ta*HIP	Homologous Rice Annotation	E value
TaHIP1	*Os*ASY1 - AB109238	9e-05
TaHIP3	*Os*RPA14 - AB111915	5e-49
TaHIP4	Os07g36820	1e-36

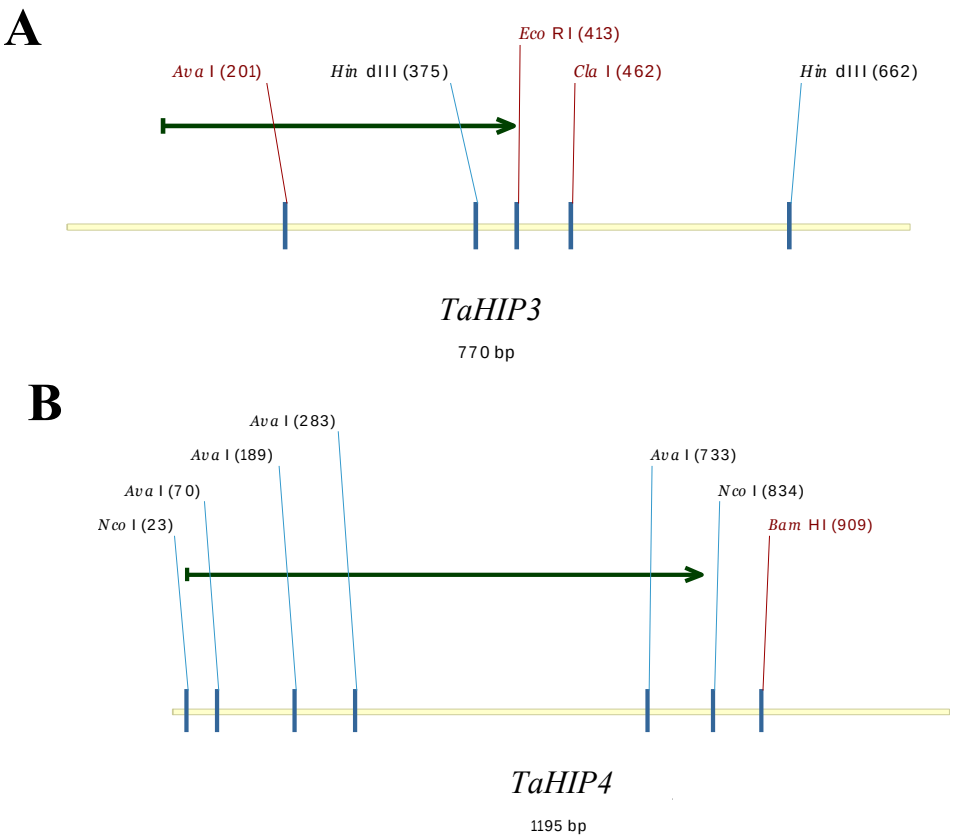


Figure 4.15 - Identification of the ORFs for *TaHIP3* and *TaHIP4*. From the *in silico* sequence searching, ORFs (*green* arrows) of 321 bp and 792 bp were identified for *TaHIP3* (**A**) and *TaHIP4* (**B**), respectively. Oligonucleotides were designed flanking the coding regions for the isolation of these candidates from wheat cDNA.

4.3.8.2 – RACE amplification for the full length *TaHIP1*

As the *in silico* search analysis was inconclusive for identifying homologues of *TaHIP1*, RACE amplification was conducted on both the 5′ and the 3′ flanking regions of sequenced *TaHIP1* in an attempt to isolate the full length gene. Oligonucleotides were designed specifically to the sequence regions flanking either side of the region of similarity to *TaASY1* to avoid amplification of *TaASY1*. For both 5′ and 3′ RACE, products were observed in the primary reactions but no successful amplification was observed in the secondary PCR (Figure 4.16). Sequence analysis of the primary products found that these fragments were

non-specific to *TaHIP1*; as alignment of these products with *TaHIP1* within Contig Express found no significant similarity between the two sequences.

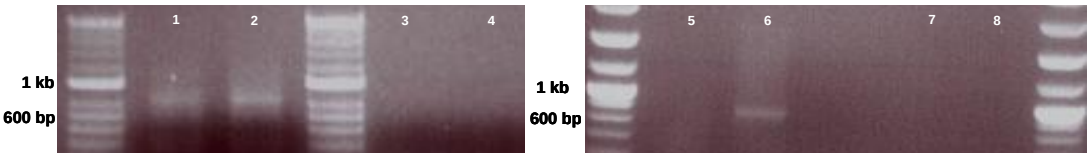


Figure 4.16 - RACE amplification of *TaHIP1*. Primary RACE amplification of the 5′ (lanes 1 and 2) and the 3′ (lanes 5 and 6) ends of *TaHIP1* with oligonucleotides *TaHIP1*_5′ (lanes 1 and 5) and *TaHIP*_3′ (lanes 2 and 6). Secondary RACE amplification of the 5′ (lanes 3 and 4) and the 3′ (lanes 7 and 8) ends of *TaHIP1* with oligonucleotides *TaHIP1*_5′ (lanes 3 and 7) and *TaHIP*_3′ (lanes 4 and 8) resulted in no amplified signal being detected.

4.3.8.3 – PCR amplification of full length *TaHIP3* and *TaHIP4*

Sequences corresponding to transcripts of *TaHIP3* and *TaHIP4* were successfully isolated from wheat meiotic cDNA (Figure 4.17 A) using oligonucleotides based on the previously obtained *in silico* sequences (section 4.3.8.1) and confirmed in the pCR8®/GW/TOPO® vector (Figure 4.17 B). Sequence analysis confirmed 98% and 97% similarity to the *in silico* identified ORFs (experimental full length sequences in Appendix 4.7). The difference observed between the sequences could be either due to the different cultivars used to amplify these sequences, with Bob White 26 used in this study, and Chinese Spring, Recital and Atlas66 representing the *in silico* obtained sequences; or due to sequencing errors from the wheat ESTs, since the differences in some amino acids in putative translations were observed.

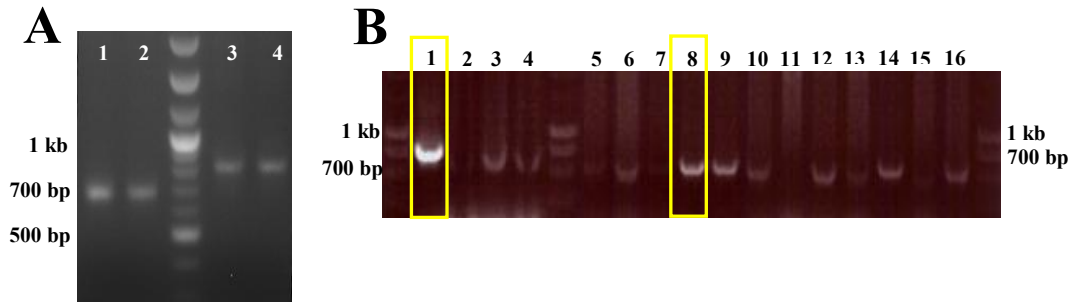


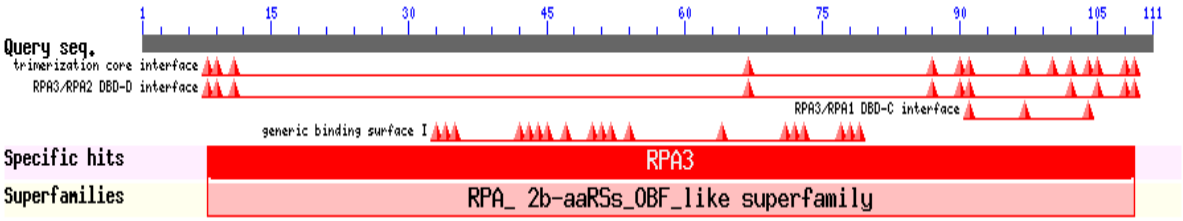
Figure 4.17 - Isolation and confirmation of full length *TaHIP3* and *TaHIP4*. (**A**) Duplicate PCR amplifications of *TaHIP3* (lanes 1 and 2) and *TaHIP4* (lanes 3 and 4). (**B**) Identification of *TaHIP4* (lanes 1-4) and *TaHIP3* (lanes 5-16) in pCR8®/GW/TOPO® vectors via colony PCR, boxed colonies indicate those that were isolated and confirmed via sequence analysis. A 2-log DNA molecular weight marker is shown; but not all band sizes are labelled.

4.3.9 – Sequence conservation of *TaHIP3* and *TaHIP4*

Homologous sequence analysis using tBLASTx determined that *TaHIP3* was similar (*E*-value ≥ 2e-40) to the 14 kDa replication protein A3 (RPA3 or RPA14) in the rice cultivar Japonica (AB111915), and the predicted *Ta*HIP3 protein was found to contain the conserved domains of the RPA3 proteins (Figure 4.18 A). *TaHIP3* from herein will be termed *TaRPA3*. Despite the conservation of the *TaRPA3* protein domains and high similarity to *Os*RPA14, there was no significant similarity to the annotated RPA3 proteins from other eukaryotes, including *Hs*RPA3 (NP_002938), *Mm*RPA3 (NP_080908) or *Sc*RPA3 (NP_010440), which can be seen in the amino acid alignments shown in Figure 4.18 B.

Phylogenetic analysis (Figure 4.18 C) shows that the amino acids encoding these proteins are conserved within the plant and animal kingdoms, with these two kingdoms forming separate clades. Since the human RPA3 protein, in conjunction with the *Hs*RPA2 protein, has been crystallised (Deng *et al.*, 2007) three dimensional modelling of *Ta*RPA3 was conducted. Figure 4.19 illustrates the conservation of tertiary structures (one β-sheet barrel and 2 α-helices) for the three RPA proteins shown; *Ta*RPA3 (green), *Os*RPA3 (yellow) and *Hs*RPA3 (red).

A



B

```
TaRPA3          ------------MDTSAPSPFVNGETLKMFPGRRVRTVVQVQ--HNEGGVLLGLSTDGHQ 46
OsRPA3          ------------MDTSGPAAFVNGEILKMFVGRRVRTVVQAQ--REEGGLLIGQSTDGHQ 46
HsRPA3          ----------MVDMMDLPRSRINAGMLAQFIDKPVCFVGRLEKIHPTGKMFILSDGEGKN 50
MmRPA3          ----------MEDIMQLPKARVNASMLPQYIDRPVCFVGKLEKIHPTGKMFILSDGEGKN 50
GgRPA3          ----------MGDVHEAPRPRIAAAQLVQHIGRPVCFVGRVEKIHPTGKLIVLSDGEGCN 50
ScRPA3          MMKGSRRTGNNTATTLNTPVVIHATQLPQHVSTDEVLQFLESFIDEKENIIDSTTMNTIS 60
                                  .   : .  *  . .         .       ::     :  .

TaRPA3          LTIRGATGAPEPP-HYIEVIGIADSSLSIRAESCTDFGEN---FDGVAFNGLCKLANDKY 102
OsRPA3          LTIKGASGAPMS--HYVEIIGIAEPNQAIRAEVCTDFGEN---FDPAPFNGLCKLANGQM 101
HsRPA3          GTIELMEPLDEEISGIVEVVGRVTAKATILCTSYVQFKEDSHPFDLGLYNEAVKIIHDFP 110
MmRPA3          GTIELMEPLDEEISGIVEVVGKVTAKATVLCASYTLFKEDTNRFDLELYNEAVKIINELP 110
GgRPA3          ATVELSEPLDEEISGILEVVGRVTNQATIMCTSYVQFREDKSPFDLELYNEALKIIHEFP 110
ScRPA3          GNAADADAAAVAN-TSLNIDTNLSSSISQLKRIQRDFKGLPPAQDFSAAPIQVSTTEKKE 119
                 .              :::      . :       *       *        .  .

TaRPA3          KYLFL------------- 107
OsRPA3          KDLFL------------- 106
HsRPA3          QFYPLGIVQHD------- 121
MmRPA3          QFFPVGLPQHE------- 121
GgRPA3          EYFPFGTGRNN------- 121
ScRPA3          TSIGVSATGGKKTTFADE 137
                    .
```

C

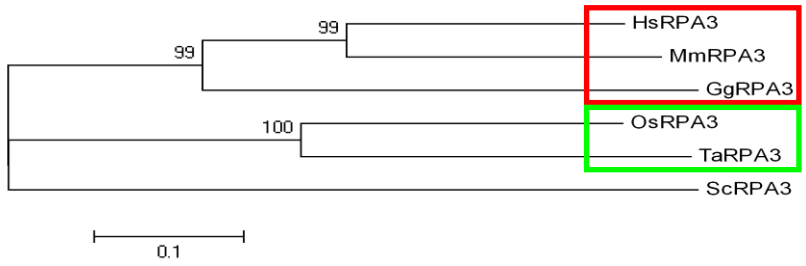


Figure 4.18 - Conserved sequence analysis of *Ta*RPA3. (**A**) *Ta*RPA3 (*Ta*HIP3) contained the conserved domains of the RPA3 proteins, including the protein binding interfaces for binding the two other sub-units of the RPA protein complex, RPA1 and RPA3. Figure generated using NCBI conserved domain predictor (**http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi**) (**B**) Amino acid alignment of the highly similar *Ta*RPA3 and *Os*RPA14 (BAD06873) proteins with the other annotated eukaryotic RPA3 proteins: *Hs*RPA3 (NP_002938), *Mm*RPA3 (NP_080908), *Gg*RPA3 (XM_418679) and *Sc*RPA3 (NP_010440). * = identical residues in all sequences in the alignment; : = conserved substitutions of amino acids with similar properties; • = semi-conserved substitutions. (**C**) Phylogenetic analysis of RPA3 shows the distinct clades between the animal (red box) and plant (green box) kingdoms. Neighbour-joining percent bootstrap values are presented for each clade while the scale bar indicates the number of amino acid substitutions per site. *Gg = Gallus gallus, Hs = Homo sapiens, Mm = Mus musculus, Os = Oryza sativa*, *Sc = Saccharomyces cerevisiae*.

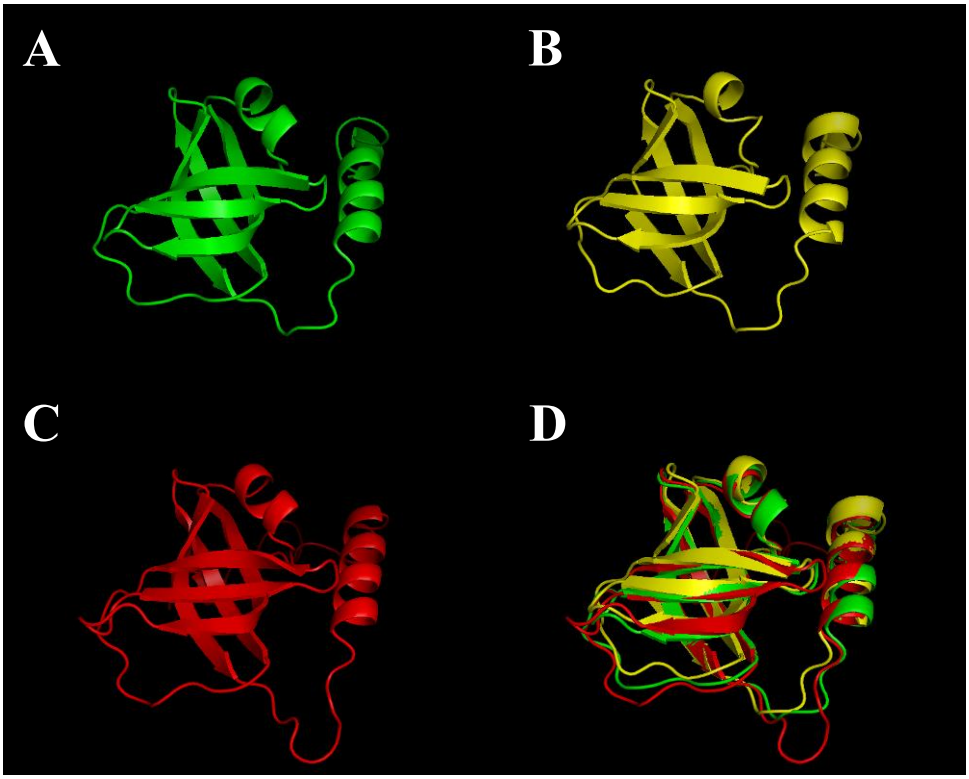


Figure 4.19 - Three dimensional modelling of RPA3 proteins. The wheat (**A**, *green*), rice (**B**, *yellow*) and human (**C**, *red*) RPA3 proteins have conserved tertiary structure of one β-sheet barrel and 2 α-helices which can be seen in the superimposition of the three models (**D**).

Sequence conservation analysis of *TaHIP4* observed significant similarity only to hypothetical plant proteins, which can be viewed in Table 4.7. The putative *Ta*HIP4 protein displays low similarity (*E*-value ≥ 2e-15) to a cysteine rich domain and a PLAC8 domain, both of which have not been characterised during meiosis (Figure 4.20 A). Phylogentic analysis of the plant orthologues of *TaHIP4* shows the predicted conservation of the proteins, with the dicot Arabidopsis being the "outlier" on the tree (Figure 4.20 B). Three dimensional modelling of *Ta*HIP4, along with the rice and Arabidopsis orthologues, was unsuccessful by the computational modelling program. Amino acids 167-246 of *Ta*HIP4 were recognised by the program, however no secondary structure was contributed to these amino acids (results not shown), which is presumably due to none of the orthologous proteins being crystallised and the structures being determined.

Table 4.7- Conserved sequence similarity of *Ta*HIP4. Homologues were obtained for *Ta*HIP4 from the public domain databases using the tBLASTx program.

Species	Accession	E value	Function
Sorghum bicolor	XM_002462961	3e-116	hypothetical protein
Oryza sativa	Os07g0553900	3e-113	hypothetical protein
Zea mays	LOC100194349	1e-112	hypothetical protein
Arabidopsis thaliana	At1g11380	9e-55	hypothetical protein

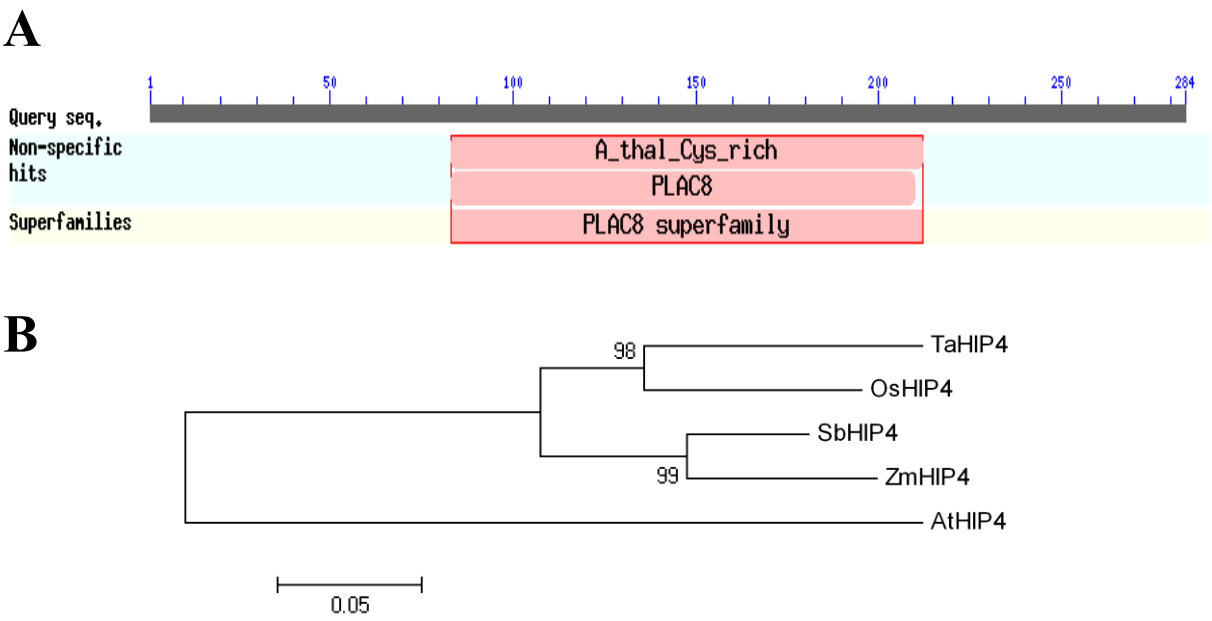


Figure 4.20 - Conservation of *Ta*HIP4. (**A**) The putative protein of *Ta*HIP4 contained a cysteine rich region and a PLAC8 domain. Figure generated using NCBI conserved domain predictor (**http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi**) (**B**) Phylogenetic analysis of HIP4 shows the predicted protein conservation within the plant species shown, with Arabidopsis (*At*HIP4) as the outlier. Neighbour-joining percent bootstrap values are presented for each clade while the scale bar indicates the number of amino acid substitutions per site. *At = Arabidopsis thaliana, Os = Oryza sativa*, *Sb = Sorghum bicolour*, *Ta = Triticum aestivum*, *Zm = Zea mays*.

4.3.10 – Expression analysis of *TaRPA3* and *TaHIP4*

Using previous microarray expression data over a reproductive time course (Crismani *et al.*, 2006), transcript expression of *TaRPA3* and *TaHIP4*, along with their interacting partner *TaASY1*, was examined. *TaASY1* and *TaRPA3* were found to be correlated (r = 0.86), with both genes having a higher level of expression during the earlier stages of meiosis when compared to the other tissue stages. This is despite there being an approximate eight-fold difference in expression during these particular stages (Figure 4.21). In contrast, *TaHIP4* was

found to have highest expression during somatic stages of the reproductive time course, with a negative correlation of r ≤ -0.70 to *TaASY1* and *TaHIP4*.

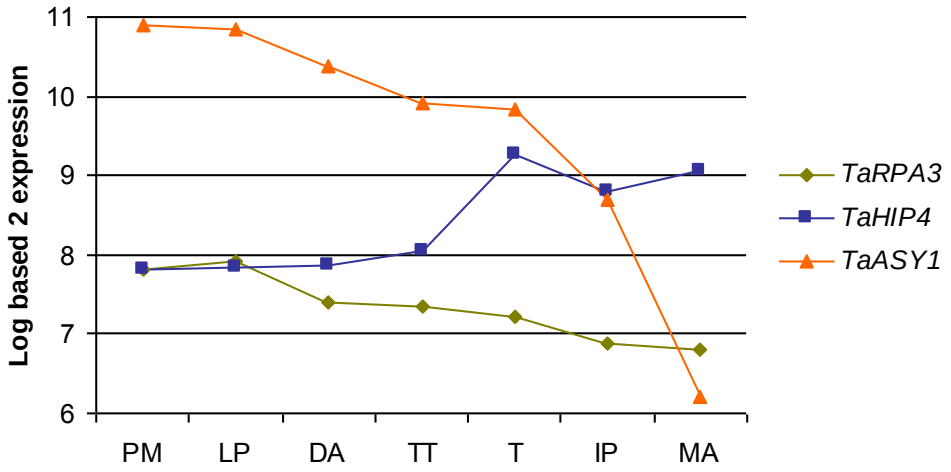


Figure 4.21 - Transcript expression of *TaASY1*, *TaRPA3* and *TaHIP4* during the seven stages of wheat reproductive development. *TaASY1* (*orange*) and *TaRPA3* (*green*) were found to have correlated expression (r = 0.86), while *TaHIP4* was negatively correlated to both *TaASY1* and *TaRPA3* (r ≥ -0.70). PM = pre-meiosis, LP = Leptotene-Pachytene, DA = Diplotene-Anaphase I, TT = Telophase I-Telophase II, T = Tetrads, IP = Immature Pollen, MA = Mature Anthers.

The expression of *TaRPA3* and *TaHIP4* was further examined in *Taasy1* knock-down mutants (as reported in Boden *et al*., 2009) and wild-type meiotic anthers, with expression levels compared against the control gene, *TaGapDH*. The RT-PCR was conducted in duplicate for each sample, and as can be seen in Figure 4.22 there was very good reproducibility for the majority of samples and observable correlation with the correspondoing *TaGapDH* control gene. Figure 4.22 illustrates that both *TaRPA3* and *TaHIP4* appear to have higher expression than *TaGapDH* in the wild-type meiotic anthers and that for both *TaRPA3* and *TaHIP4* this expression varies in a number of the *Taasy1* knock-down mutants; for example in *Taasy1.1.9.4*, *Taasy1.1.9.5*, *Taasy1.2.2.5*. In Figure 4.22 it is also observed that there appears to be an intensity gradation across the gel, which highlights the semi-quantitative nature of the RT-PCR expression analysis.

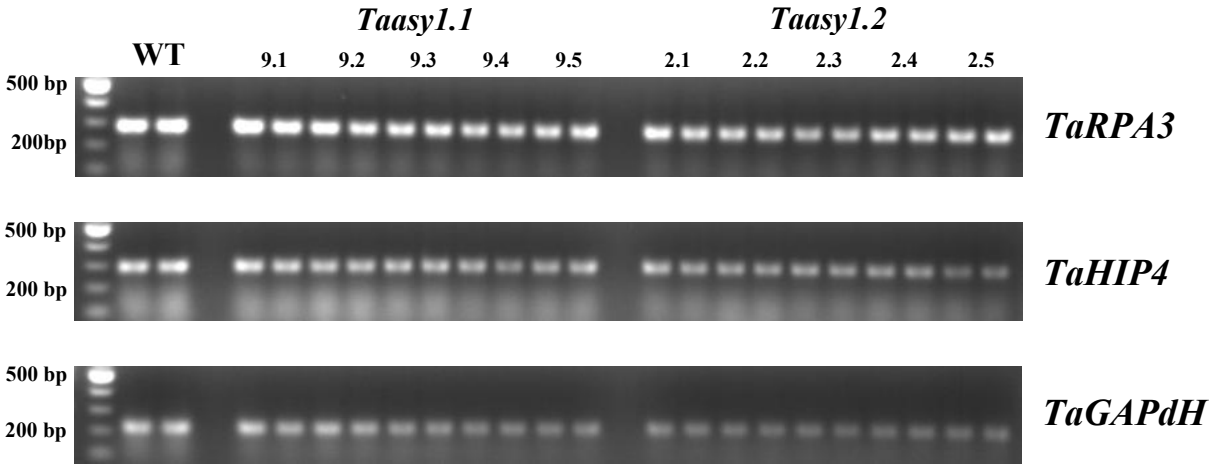


Figure 4.22 - RT-PCR expression analysis of *TaRPA3* and *TaHIP4*. Expression of both *TaRPA3* and *TaHIP4* appears to be higher than the expression of the control gene, *TaGAPdH*. This expression varies but is most pronounced in *Taasy1.1.9.4, 1.1.9.5 and 1.2.2.5*. A 1 kb DNA molecular weight marker is shown; but not all band sizes are labelled.

4.3.11 – Heterologous *Ta*RPA3 and *Ta*HIP4 protein expression

4.3.11.1 – Protein expression vector preparation

Amplification of the *TaRPA3* and *TaHIP4* ORFs (Figure 4.23 A) and the subsequent ligation and isolation of these fragments within the pCR8®/GW/TOPO® vector was successful (Figure 4.23 B and C). Sequence analysis confirmed one colony for each *TaRPA3* and *TaHIP4*, with these plasmids acting as donors for the transfer by recombination of the ORF into the protein expression vector pDEST17®. Expression vector recombinants were confirmed via colony PCR and sequence analysis (Figure 4.24).

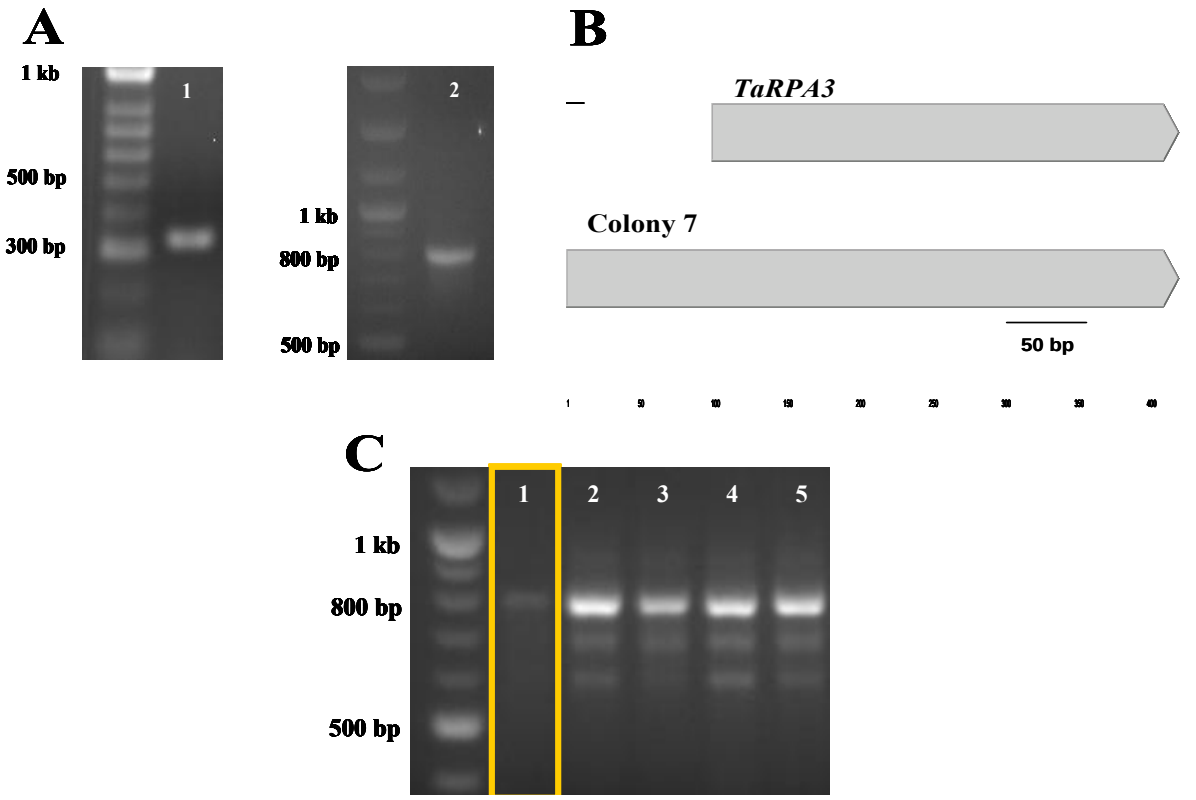


Figure 4.23 - Isolation of *TaRPA3* and *TaHIP4* and subsequent confirmation in pCR8®/GW/TOPO®. (**A**) PCR amplification of the *TaRPA3* (lane 1) and *TaHIP4* (lane 2) ORFs. (**B**) Sequence analysis confirmed the correct direction of the *TaRPA3* gene within the pCR8®/GW/TOPO® vector. (**C**) Colony PCR of *TaHIP4* in the pCR8®/GW/TOPO® vector, with sequence analysis confirming colony 1 (boxed) in the correct orientation. A 1 kb DNA molecular weight marker is shown, but not all band sizes are labelled.

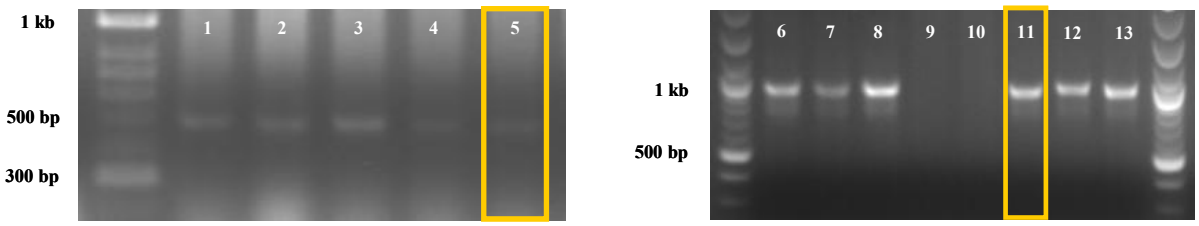


Figure 4.24 - Confirmation of *TaRPA3* and *TaHIP4* within the pDEST17® protein expression vector. Confirmation of the *TaRPA3* (lanes 1-5) and *TaHIP4* (lanes 6-13) ORFs within the protein expression vector pDEST17® via colony PCR. Plasmids from the highlighted boxes (*yellow*) were confirmed for correct orientation and translation frame. A 1 kb DNA molecular weight marker is shown, but not all band sizes are labelled.

4.3.11.2 – *Ta*RPA3 and *Ta*HIP4 protein expression and extraction

Protein expression and extraction of *Ta*RPA3 and *Ta*HIP4 was attempted at induction temperatures of 23°C and 37°C, with both induced and repressed samples. The heterologously

expressed protein of *Ta*RPA3 was expected to be approximately 11.5 kDa, while *Ta*HIP4 was expected to be approximately 28 kDa. However, the extracted elutions from the induced and repressed samples of *Ta*RPA3 and *Ta*HIP4 proteins at either 23°C or 37°C showed no discernable differences (Figure 4.25).

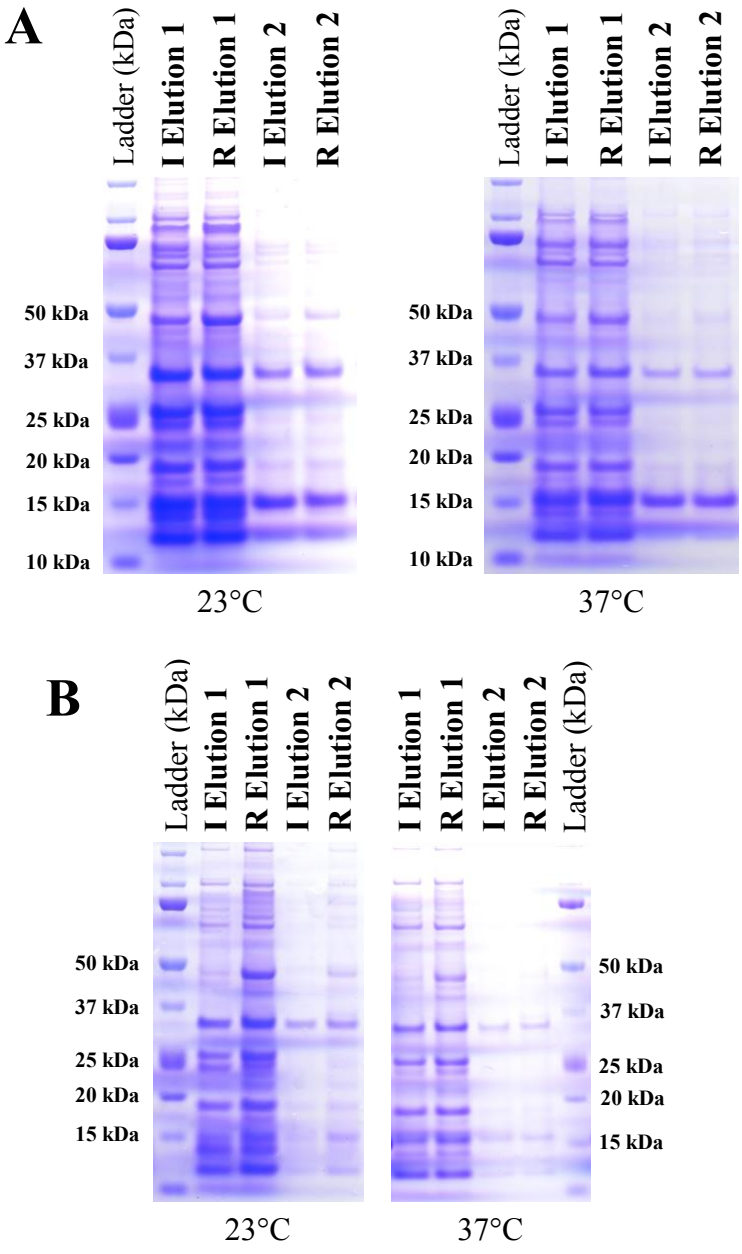


Figure 4.25 - Protein expression and extraction of *TaRPA3* and *TaHIP4*. Protein extractions from induced and repressed samples incubated at either 23°C or 37°C for *Ta*RPA3 at approximately 11.5 kDa (**A**) or *Ta*HIP4 at approximately 28 kDa (**B**) were attempted, but not successful in the isolation of the recombinant proteins as there was no discernable difference between the extracted protein bands of induced or repressed samples. I = induced; R = repressed. A BIO-RAD Precision Plus Dual Colour Protein Ladder was used, but not all sizes are shown.

4.4 – Discussion

4.4.1 – Self activation of the *Ta*DMC1 bait vector

One of the objectives of this chapter was to characterise *Ta*DMC1 protein:protein interactions via a yeast two-hybrid approach. DMC1 has previously been investigated through this approach in a number of species, where the protein has been shown to interact with Brca2 (Arabidopsis), RAD51 (mouse and Arabidopsis), RAD51C (mouse), SWI5 (transcriptional acitivator, yeast), along with a self interaction within only the C-terminus of the DMC1 protein (yeast, mouse and Arabidopsis) (Dresser *et al.*, 1997; Tarsounas *et al.*, 1999; Siaud *et al.*, 2004). These investigated interactions via the yeast two-hybrid approach, along with the other investigation showing no growth in yeast containing *Hs*DMC1, along with a non-partner protein of *Hs*FANCD, (Hussain *et al.*, 2004), indicate that the DMC1 full-length proteins from these species are not self activating with the yeast two-hybrid approach.

Of interest therefore is the result that shows the wheat DMC1 protein to be self activating within the yeast two-hybrid approach. Self activation within the yeast two-hybrid system is common, especially with transcription activating genes which are maintained in yeast, or genes which act in other processes but exhibit high transcriptional activity when tethered to a yeast promoter sequence (Du *et al.*, 1996; Hu *et al.*, 1997; Walhout and Vidal, 1999). To date the DMC1 protein has not be shown to be involved in transcriptional activity in any species, although there is the observed interaction between the yeast DMC1 protein and the yeast transcriptional activator, SWI5 (Dresser *et al.*, 1997). Therefore it is a possibility that the wheat DMC1 protein is activating the yeast system through an interaction with a native yeast transcriptional activator as there is high conservation of DMC1 sequence and structure between the wheat and yeast proteins (Chapter 3). This seems highly unlikely as the human and mouse DMC1 do not exhibit self activation and these proteins are more similar to the yeast protein when comparing wheat to yeast.

Another possible, but unlikely, explanation of the self activation exhibited by the wheat DMC1 protein is that a spontaneous mutation has arisen within the bait vector during yeast culturing. Self-activators can arise from spontaneous mutants within the interactor fusions (Bartel and Fields, 1997). As self activation for the wheat DMC1 protein was conducted in duplicate this seems improbable. Therefore it is still not fully understood why the wheat DMC1 protein exhibits self activation while the DMC1 proteins from other species are not found to be self activating.

4.4.2 – Protein interactions of *Ta*ASY1 protein and the HORMA protein domain

4.4.2.1 – Using the *Ta*HORMA domain for initial yeast two-hybrid screen

The *Ta*HORMA domain of ASY1 and other HORMA containing proteins are known to facilitate the direct interaction with chromatin during both mitosis and meiosis; often with these proteins interacting with chromatin associated DNA adducts, DNA double stranded breaks and synaptonemal complex proteins (Hollingsworth *et al.*, 1990; Aravind and Koonin, 1998; Armstrong *et al.*, 2002; Mikhailova *et al.*, 2006; Boden *et al.*, 2007). In using this domain in a yeast two-hybrid system it was therefore envisaged that a broader „pool" of proteins would be identified upon which to then confirm whether there was an interaction with the full length *Ta*ASY1 protein.

4.4.2.2 – Interactions with the *Ta*HORMA domain of *Ta*ASY1

Using the HORMA domain of the wheat *Ta*ASY1 protein, six initial interactors were identified. Upon confirmational co-transformation in yeast, four were confirmed to interact with the *Ta*HORMA domain and the remaining two interactors were found to be false positives from the initial screen. Of the four interactors identified, three were found to interact with the full length *Ta*ASY1 protein, with *Ta*HIP2 interacting only with the *Ta*HORMA domain. This interaction of *Ta*HIP2 only with the *Ta*HORMA domain and not with the full

length *Ta*ASY1 (which contains the *Ta*HORMA domain) indicates that there must be either a conformational change of the *Ta*HORMA domain or a sterical/electrostatic hinderance which leads to *Ta*HIP2‟s inability to bind to the *Ta*HORMA domain.

Protein domain conformational changes have been well documented in the literature, especially domain conformational changes resulting from ligand (eg calcium) binding and post-translational modification such as phosphorylation (Dirlam-Schatz and Attie, 1998; Jiang and McKnight, 2006). However, a conformation change in domain structure between the „isolated fragment‟ (that is, the domain only) and when residing within its full length protein has not been well discussed within the literature. Yet, one example is that the setting or protein environment of a domain can affect the conformation of a domain; where a single amino acid mutation outside and distant to the antigen binding site of antibody U10 has been shown to cause conformational change within this domain (Chien *et al.*, 1989). Therefore it is possible that the *Ta*HORMA domain does have a difference in the conformational structure between the domain in „isolation‟ and being located within the full length protein of *Ta*ASY1, resulting in the differential binding of *Ta*HIP2. It is therefore possible that *Ta*HIP2 is capable of interacting with another protein containing a HORMA domain; yet not have the ability to bind the HORMA domain within the full length *Ta*ASY1 protien.

4.4.2.3 – Interactions with *Ta*ASY1 and the *Ta*HORMA domain

The three remaining *Ta*HORMA interacting proteins (*TaHIPs*) were observed to interact with both the *Ta*HORMA domain and the full length *Ta*ASY1 protein. Sequence analysis of these three interactors through *in silico* sequence searching identified similarity to *TaASY1* for *TaHIP1*, *Os*RPA14 (or RPA3) for *TaHIP3*, and a non-annotated protein from rice for *TaHIP4*. *TaHIP3* and *TaHIP4* were successfully identified using *in silico* methods where strong similarity was found against homologues and wheat ESTs. Having designed oligonucleotides from these sequences, the products were amplified from wheat meiotic tissues and subsequently confirmed by sequence analysis.

The aforementioned similarity between *TaHIP1* and *TaASY1* suggests homodimerisation occurring with the *Ta*ASY1 protein; however despite *TaHIP1* having a significant similarity to *TaASY1*, this similarity is located only within a small region (approximately 100 bp) of the larger sequence obtained for *TaHIP1* (416 bp) with low similarity (against *TaASY1*) sequences flanking this region of similarity. Since these flanking regions have no similarity to either the coding region or genomic regions of *TaASY1*, it is concluded that *TaHIP1* is not *TaASY1* or a splice variant of *TaASY1*.

It is likely that *Ta*HIP1 shares an unknown protein domain or motif with the *Ta*ASY1 protein which has not yet been characterised. Further homologue sequence searching resulted in no significant matches for *TaHIP1* in rice, and interrogation of the wheat EST public database returned limited hits to extend the *TaHIP1* sequence. Subsequent RACE of *TaHIP1* resulted in amplification of several products using stringent cycling conditions. However through sequence analysis, these products were deemed non-specific as neither Contig Express nor BLASTn alignment found similarity between the products amplified and the previously obtained *TaHIP1* sequence that was used in first designing the oligonucleotides.

4.4.3 – Positive correlated expression of *TaRPA3* with *TaASY1* but not with *TaHIP4*

Using expression data from a wheat reproductive tissue series previously conducted by Crismani and colleagues (2006) the expression of *TaRPA3* was found to be correlated ($r = 0.86$) with the expression of *TaASY1* over the seven stages investigated. This may suggest that *TaASY1* and *TaRPA3* are co-regulated during reproduction. Having identified *TaRPA3* within the yeast two-hybrid screen adds weight to this argument. Previously characterised meiosis genes such as *HOP2* and *MND1* which are known to interact with one another in yeast, mouse and humans (Chen *et al.*, 2004; Enomoto *et al.*, 2006; Pezza *et al.*, 2006), were also identified in wheat and have correlated expression; thus suggesting co-regulation of these interacting partners (Chapter 3, section 3.3.4). Therefore it is interesting that the transcription of *TaHIP4*

was found to be negatively correlated ($r \geq -0.70$) across the reproductive tissue series with both *TaASY1* and *TaRPA3*. The difference in expression could represent a negative feedback or interaction between *Ta*ASY1, *Ta*RPA3 and *Ta*HIP4 transcription and/or protein function; with the earlier expressed *Ta*ASY1 most likely acting negatively on *Ta*HIP4.

Using reverse-transcriptase PCR (RT-PCR), expression of both *TaHIP3* and *TaHIP4* appeared to be slightly down-regulated in two independent *Taasy1* knock-down lines, in several plants. However, the control gene *TaGapDH* also appeared to have less expression in the knock-down mutants. Therefore these results may just represent a „global" down-regulation of expression in these mutants. What must be taken into close consideration here is that this type of expression analysis (RT-PCR) is only a semi-quantitative analysis, and therefore should be taken as only provisional results. In retrospect the better analysis to have conducted, and what would be highly suggested for future work, would be the more quantitative approach of Q-PCR to confirm the expression of these genes within the *Taasy1* knock-down mutants.

4.4.4 – *Ta*ASY1 interactor : *Ta*RPA3

4.4.4.1 - *Ta*RPA3 has high sequence conservation to *Os*RPA14 but not to other RPA3 proteins

TaRPA3 was successfully confirmed as a wheat homologue to the published rice replication protein A 14 kDa subunit with 87% similarity observed between *TaRPA3* and the reported *OsRPA14* (Ishibashi *et al.*, 2006). However, when comparing these two plant proteins with the previously annotated RPA proteins from other eukaryotes, there is no significant similarity (*E*-value ≥ -20) observed. This lack of similarity across the kingdoms is somewhat surprising as the replication protein A complex is found to have important roles in DNA replication, recombination and repair (Wold, 1997). Nonetheless, other rice RPA subunits have also been reported to have low similarity to the RPA subunits from other

species; for example *Os*RPA3 has 26.2%, 21.3% and 35.8% similarity to the human, yeast and *Drosophila* RPA3 proteins respectively (Sakaguchi *et al.*, 2009). Despite this lack of sequence conservation, it is evident that there is still conservation of the tertiary structure for the protein subunit based on the modelling conducted in this study (Figure 4.18). This provides evidence that the conservation of the protein structure and not the protein sequence is what is necessary to conserve the functional role of the RPA complex across organisms.

4.4.4.2 – RPA functionality and ASY1

For furthering the understanding of the role of the RPA protein complex, *Ta*RPA3 and *Ta*ASY1, heterologous protein expression of *Ta*RPA3 was attempted. Unfortunately, *Ta*RPA3 protein expression and extraction under a variety of conditions were not successful. With the heterologously expressed *Ta*RPA3 protein it was hoped to at least attempt DNA binding assays to ascertain whether the wheat RPA3 protein has the ability to bind DNA, adding more evidence for a role of RPA and an RPA:ASY1 complex within meiosis. Having a heterologously expressed RPA3 protein would also have allowed for the production of a wheat RPA3 protein antibody which could have been used for western analysis and immunolocalisation experiments for the detection of the RPA protein complex.

An alternative route to determine RPA and RPA3 protein function in meiosis would be to investigate meiosis in both *rpa3* and *asy1* knock-out Arabidopsis mutants. However this approach could very well be limited, as the *RPA3* subunit of the RPA complex has so far been the only allele of this particular subunit found; unlike subunits of *RPA1* and *RPA2* which have been found to have three alleles each. Therefore it is highly likely that any knock-out *rpa3* plant would exhibit a lethal phenotype due to the essential nature of the RPA complex in DNA replication. A more viable option would be to investigate a knock-down plant of RPA3 (if it could be obtained), and then examine the loading and location of ASY1 to determine what role the RPA complex, through RPA3, has with ASY1 in bread wheat. As mentioned above and due to the observed protein interaction with ASY1 and RPA3 (and therefore the

RPA complex), an *asy1* knock-down/knock-out Arabidopsis mutant could also be investigated. An *asy1* mutant effect may be more evident, as the RPA complex has been found to locate to the meiotic chromosomes at later stages of meiosis, for example during pachytene in mouse (Moens *et al.*, 2002), than the ASY1 protein which in wheat is observed to be located on the chromosomes during early meiosis (Boden *et al.*, 2009). However to investigate the location and loading of the RPA complex in such mutants, an antibody for any of the RPA subunits would be required; which has been unable to be produced for RPA3 here from the wheat protein due to the inability to isolate the heterologously expressed protein. Furthermore, due to the low levels of sequence similarity it would be unlikely that antibodies produced for the RPA complex in other species (for example in Moens *et al.*, 2002) would work in Arabidopsis. Therefore at the present time the function of the RPA complex, in particularly RPA3, with the ASY1 protein remains unclear.

4.4.5 – *Ta*ASY1 interactor : *Ta*HIP4

The *Ta*ASY1 interactor *Ta*HIP4 was identified as a novel plant protein, with no significant similarity to any mammalian or yeast nucleotide and protein sequences in the public databases. Given the lack of identifiable orthologues for the *TaHIP4* gene, this has limited the ability to predict the functional role it may have during meiosis in bread wheat. Three dimensional modelling of *Ta*HIP4 only returned a model for a small section of the protein (amino acids 167-246) which contained no secondary structure. Given the lack of characterised homologues, and therefore lack of crystallised proteins representing this protein, this is not surprising. Using the publically available databases it was determined that the *Ta*HIP4 protein contains an uncharacterised cysteine-rich (PLAC8) domain. While this particular cysteine-rich domain has not been characterised, cysteine rich domains in general have not been reported as yet to have a role during meiosis. However, many cysteine rich domain proteins have been found to be involved in receptors/signalling, chaperone proteins, methyltransferases and kinases (Bestor, 2000; Miernyk, 2001; Holm *et al.*, 2009; Stiegler *et*

al., 2009). Based on this knowledge it is not unreasonable to propose a role for cysteine rich domains during meiosis.

To further the understanding of *Ta*HIP4 and its role in meiosis, heterologous expression of a recombinant form of the protein was attempted. Unfortunately this was not successful; with no difference between induced and repressed samples of cell lysates being observed although multiple parameters and optimisation steps were investigated. While not within the scope of this study, another research direction that could be pursued would be to identify, obtain and functionally characterise a T-DNA Arabidopsis mutant for this particular gene. Using a knock-out or a knock-down mutant for *AtHIP4* would allow for a comprehensive set of experiments to be conducted that examine the effects of this gene on both vegetative and reproductive growth.

Chapter 5

Chapter 5 – Using Arabidopsis mutants to study novel wheat meiotic candidates

5.1 – Introduction

Within the last decade, Arabidopsis has become a valuable tool for plant scientists; with the species relatively small genome (140 Mb) being fully sequenced and annotated (Arabidopsis Genome Initiative, 2000; Haas *et al.*, 2005). This sequencing effort has allowed the production of 88,000 publically available transfer DNA (T-DNA) mutants covering roughly 73% of the expected 29,000 genes in Arabidopsis (Alonso *et al.*, 2003). These mutants allow researchers to conduct targeted reverse genetics; establishing the function of their gene of interest from the effects on an Arabidopsis plant lacking that particular gene (Krysan *et al.*, 1999; Østergaard and Yanofsky, 2004). In plants, these particular Arabidopsis mutant stocks have been readily utilised for the investigation and understanding of genes involved in meiosis (Hamant *et al.*, 2006). As such these mutants are powerful tools for the investigation of novel meiotic candidates which have been identified through other methods, such as gene expression profiling or yeast two-hybrid. This is particularly true in the case of studying more complex plant species, such as wheat, where the creation of mutants (either knock-downs/out or over-expression) is both time consuming and expensive; especially when taking into consideration that there is no guarantee that a candidate found to have meiotic expression will have a meiotic role.

From the recent study conducted by Crismani and colleagues (2006) on gene expression during wheat meiosis and microsporogenesis, 1,350 probe sets were found to have meiotic regulation. Chapter 2 of this dissertation was focused on the identification of novel transcripts from these 1,350 meiotically expressed probe sets. Using filtering steps, including expression profile clustering, sequence analysis and chromosome location; eight candidates

were identified for further analysis. These novel wheat transcripts were finally chosen from those filtered as these candidates also had homologous Arabidopsis genes. Research in this chapter is focused on the identification of T-DNA mutant lines of these Arabidopsis homologues and the functional analysis of these mutants to determine if these candidates have a role in meiosis.

5.2 – Materials and Methods

5.2.1 – Identification of T-DNA mutants for Arabidopsis homologues

T-DNA mutants of the Arabidopsis homologues (Table 5.1) were identified using T-DNA Express: Arabidopsis Gene Mapping Tool (http://signal.salk.edu/cgi-bin/tdnaexpress) from Salk Institute Genomic Analysis Laboratory (SIGnAL). Mutant lines with T-DNA insertions which reside as close to or within the first exon were preferentially chosen to increase the likelihood of obtaining a knock-out line for the gene of interest (Krysan *et al.*, 1999).

Table 5.1 - Arabidopsis homologues of novel meiotic candidates from wheat.

Novel	Transcript ID	Annotation	*At* Homologue
2	TaAffx.11970.1.S1_s_at	Putative DUF936 domain protein	At4g13370
9	Ta.10024.1.S1_s_at	Putative Homeobox-7 Transcription factor	At5g46880
11	Ta.8260.1.A1_at	Unknown protein	At4g02800
18	TaAffx.25097.1.S1_at	Unknown protein	At3g56870
24	TaAffx.84212.1.S1_at	Chromosome condensation protein-like	At5g37630
25	Ta.25933.1.S1_at	Condensin subunit (chromosome condensation)	At3g57060
26	Ta.6748.2.A1_a_at	Putative Kinesin Motor protein	At5g02370
29	TaAffx.44789.1.S1_at	Putative Histone Acetyltransferase B (HATB)	At5g56740

5.2.2 – Growth conditions

Arabidopsis plants, SALK seed stock (Alonso *et al.*, 2003) and Colombia-0 (comparative wild-type), were grown in a growth room at a constant 18°C with a 12 h day/night regime.

5.2.3 – Filtering of plants to determine a meiotic function of novel candidates

The Arabidopsis T-DNA seed stock plants were taken through a number of "filtering" steps to determine if the novels with meiotic expression have a meiotic function (Figure 5.1).

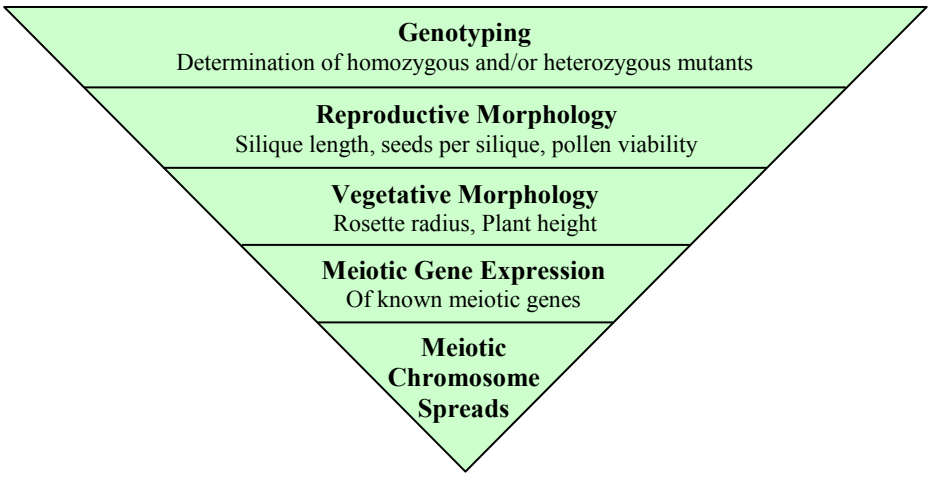


Figure 5.1 - Experimental filtering steps conducted on the Arabidopsis T-DNA seed stock plants. Using a number of "filtering" steps the Arabidopsis T-DNA stock plants, each of which represents a mutant line for a gene of interest highlighted in Table 5.1, were investigated for a meiotic role.

5.2.4 – Genotyping plants

Genomic DNA was extracted from plants, SALK seed stock and wild-type, using the REDExtract-N-AmpTM Plant PCR Kit (Sigma) as per the manufacturer"s manual except that only 40 µL of Extraction and Dilution Buffers were used instead of 100 µL. Oligonucelotides used for genotyping were those designated for each SALK line by T-DNA Express: T-DNA primer design (http://signal.salk.edu/tdnaprimers.2.html) (Appendix 5.1), for the amplification of the endogenous gene; along with the T-DNA insert primer LBb1.3 (Lb, 5′-ATTTTGCCGATTTCGGAAC-3′) from Salk Institute Genomic Analysis Laboratory (SIGnAL) for the amplification of part of the T-DNA insert. PCR amplification was conducted in 10 µL reactions containing 5 µL of Extract-N-Amp PCR ReadyMix (Sigma), 0.4 µL of either LP oligonucelotide (Appendix 5.1) or Lb oligonucloetide, 0.4 µL of RP oligonucleotide (Appendix 5.1), 3 µL of extracted DNA and 1.2 µL of sterile deionised water.

PCR conditions were initiated with denaturation for 2 minutes at 94°C; followed by 40 cycles of 96°C for 30 seconds, 55°C for 30 seconds, 68°C for 1 minute, with a final extension at 68°C for 10 minutes. Post PCR reactions were analysed by gel electrophoresis (1.5% agarose) and visualised using ethidium bromide and a UV transilluminator (FirstLight^{TM} UV illuminator, UVP, USA).

5.2.5 – Phenotype measurements

5.2.5.1 – Reproductive morphology

Reduced fertility in plants is often used as an indicator for identifying potential meiotic mutants (Ross *et al.*, 1997; Bhatt *et al.*, 1999; Sanders *et al.*, 1999; Siddiqi *et al.*, 2000; Mercier *et al.*, 2001; Caryl *et al.*, 2003). Consequently, silique length, seeds per silique and pollen viability were all recorded to determine whether the mutants being investigated within this chapter were possible meiotic mutants. Siliques were measured using a calliper digital digimax 2000, after which these siliques were broken open and seeds per silique were also recorded. Pollen viability was determined for each plant, SALK seed stock and wild-type using Alexander staining (95% ethanol (10 mL), 1% malachite green in 95% ethanol (5 mL), 1% acid fuchsine in milli-Q water (5 mL), 1% Orange G in milli-Q water (5 mL), glacial acetic acid (2 mL), glycerol (2 mL), phenol (5 g) and milli-Q water to a total volume of 100 mL).

Arabidopsis buds were removed from the plants and were placed in approximately 1 ml of 1:50 dilution of the Alexander stain on a glass microscope slide and covered with a cover slip (22 x 50 mm). Preparations were flamed briefly and incubated at room temperature for 10 minutes, and following gentle squashing to release the pollen from buds were allowed to further incubate at room temperature for 5 minutes. Following incubation, preparations were then flushed with sterile deionised water and pollen was visualised using a Leica DM LA microscope which was operated with a Leica Microscope Control Unit (Leica

Microsystems, Wetzlar, Germany). Images were obtained using a HV-C20A camera (Hitachi, Japan) and imported with Laser Microdissection System version 4.3.1.0 software (Leica Microsystems). Two tailed student‟s t-tests assuming equal variance (Microsoft® Office Excel 2003, Microsoft Corporation, USA) were conducted to compare silique length, seeds per silique and pollen viability between the Arabidopsis mutants and wild-type controls.

5.2.5.2 – Vegetative morphology

Vegetative measurements were recorded for Arabidopsis homologues where statistically significant differences were identified for the reproductive attributes measured for homozyogous/hetereozygous mutant plants and null segregants/wild-type controls. Rosette radius was measured on plants from the centre of the plant to the end of the longest leaf, using a calliper digital digimax 2000 (Figure 5.2 A); while plant height was measured from the bottom of the stem emerging from the rosette structure to the "tallest" flower (Figure 5.2 B).

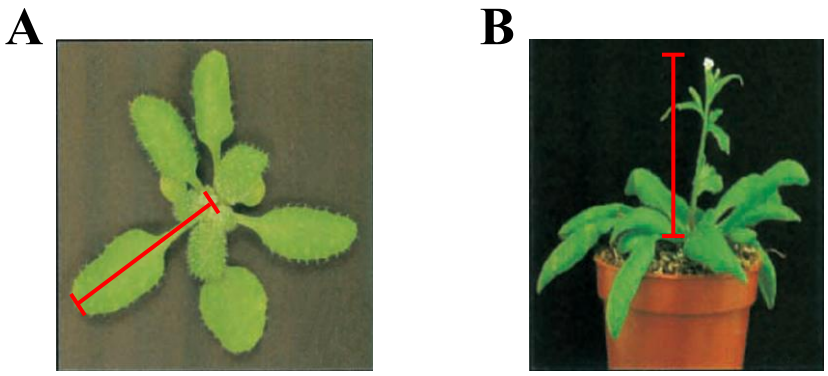


Figure 5.2 - Vegetative phenotype measurements. Rosette radius (**A**) and plant height (**B**) were measured on plants that were identified as being statistically different with respect to reproductive attributes that were measured. Pictures adapted from Boyes *et al.*, 2001.

5.2.6 – Meiotic gene expression

Since the plants which were found to have significant differences between the mutant plants and the wild-type controls in the reproductive morphologies investigated have putative gene regulatory functions; differences in gene expression of eight known meiotic candidates

from a wide-range of meiotic functions along with the control gene actin 8 (ACT8), were tested. Gene expression analysis was conducted using a semi-quantitative one-step reverse-transcriptase (RT) PCR protocol (SuperScript® One-Step RT-PCR system with Platinum® Taq DNA Polymerase, Invitrogen). Oligonucleotides used for amplification of each gene transcription product are listed in Table 5.2. RNA was collected from tissues using the Trizol reagent (Invitrogen) as per the manufacturer‟s protocol.

Table 5.2 - Oligonucleotides used to investigate meiotic gene expression. Key for gene function as follows: End Pro = DNA end processing, *Pair* = Pairing, *Recom* = Recombination, *Rep* = Repair, *Syn* = Synapsis.

Gene	Gene Function	Forward Primer 5′-3′	Reverse Primer 5′-3′
ACT8	Control	TCTGGATTGGTGGTTCTATC	TTAGTGCCTCAGGTAAGAGC
MN29		GTTATCCTGACAGGCTGCGG	GTTCAGAGGCGGGGTGAATC
ASY1	Pair/Syn	CAGAGATCACTGAGCAGGAC	GGTAATGTCAGCAGTGGAGT
DMC1	Recom	CTACAGGGTGTCAAGCTCTC	GTGAAATCCACTCGGAATAA
HOP2	Pair/Recom	AAATGCAGCTGATGCTCTAC	CACCAATGTAATCCCTTCAC
MND1	Pair/Recom	CAGGAGCAAGAGATCACG	CTTTGAGTTGCGTCAAGG
MRE11	Recom/End Pro	GTTTCAAGTCGTCAGCGACC	TGTTGAACCAGTCAGAAACA
MSH4	Recom/Rep	GGCTTCACTGCATCTATCTC	TGACTATCACACCCTTTTCC
SPO11-1	Recom	CTGGTTGTGGAAAAAGAAAC	ACGCTTCAGGTAGCATCTAA
ZYP1a	Syn	CGTACAGGACTCTCCTTTTG	CCACTTGGAATCAGTGTTTT

RT-PCR mixtures (20 µL) contained 100 ng of RNA, 12.5 µL of 2x Reaction Mix, 0.5 µL of forward oligonucleotide (10 µM), 0.5 µL of reverse oligonucloetide (10 µM), 0.5 µL of RT/ Platinum® *Taq* Mix and sterile deionised water to the final reaction volume. Reaction conditions were 50°C for 30 minutes, 94°C for 2 minutes, 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 68°C for 30 seconds, with a final extension for 10 minutes at 68°C. Post PCR reactions were analysed by gel electrophoresis (1% agarose) and visualised as per section 5.2.3.

Resulting band products were isolated from the MN29 RT-PCR analysis using the Invitrogen Gel Extraction Kit as per the manufacturer‟s instructions and sequence analysis was conducted as per Chapter 3 section 3.2.1.5, except that the aforementioned RT-PCR primers and purified PCR product as a template were used for the sequencing reaction.

Sequences obtained were analysed using the BLASTn tool on the NCBI public database (http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 174.0).

5.2.7 – Meiotic chromosome spreads

Buds from MN29 were harvested from plants and were fixed in 3:1 ethanol – glacial acetic acid for 1 hour, with refreshment of fixative three times during the hour; fixed buds were then stored at 4°C until used. Fixed flower buds were removed from fixative and placed into water and washed twice for 5 minutes; water was then replaced with 10 mM Citrate buffer and buds were washed twice for 5 minutes. Buds were then placed in a Pectolytic enzyme mixture (0.3% (w/v) cellulose RS, 0.3% (w/v) pectolyase Y23, 0.3% (w/v) cytohelicase (Sigma)) for 2.5 hours in a humi-chamber at 37°C. Tissues were then washed twice with sterile deionised water for 5 minutes, followed by a single bud (approximately 0.5 mm in size) being placed onto a microscope slide (poly L-lysine coated) in 3 µL of 10 mM citrate buffer.

The bud was then squashed with the blunt end of a dissection needle until a fine suspension had formed. To the tissue suspension, 10 µL of 60% acetic acid was added before placing the slide onto a heating block at 45°C for 2 minutes with stirring of the suspension mixture, with a further 10 µL of 60% acetic acid added after 1 minute. A "boundary" was created around the suspension mixture immediately after incubation by pipetting ice cold fresh 3:1 ethanol-acetic acid fixative around the droplet of suspension mixture, until the fixative covered the whole microscope slide. Excess solution was then removed by tilting the slide.

The slide preparation was then air-dried by being placed upright on a tissue; after which 10 µL of DAPI (4′, 6′-diamidino-2phenylindole, 2.0 µg mL^{-1} in VectaShield™, Vector Laboratories, USA) was added along with a cover slip (22 x 50 mm). Excess DAPI was removed using Whatman paper with gentle pressure before sealing the slides using clear nail

polish. Slides were stored at -20°C, light protected, and visualised as per DAPI visualisation in Boden *et al*., 2009.

5.3 – Results

5.3.1 – Identification of T-DNA Arabidopsis homologue mutants and genotyping of seed stocks

T-DNA mutant lines for the Arabidopsis homologues of the wheat novel meiotic candidates were identified (Table 5.1). Mutant lines were selected with the insertion as close as possible to the first exon, where this was feasible (Figure 5.3 and Appendix 5.2) (Krysan *et al.*, 1999). Upon plants reaching 4 weeks of age the genotype of the Arabidopsis T-DNA seed stocks was determined using PCR for the amplification of an intact endogenous gene and the insert T-DNA sequence (examples illustrated in Figure 5.4), with full genotyping results displayed in Table 5.3. Novels 2, 11, 18 were found to be all homozygous for the T-DNA insert, which is due to the seeds coming from the annotated homozygous seed stock from the SALK institute (Alonso *et al.*, 2003). Novels 24 and 25, two different types of chromosome condensation proteins, were found to be only null segregants with no T-DNA insert present; suggesting that these two particular novels are essential for development. MN09, MN26 and MN29 plants were observed to have homozygous (2 copies of the T-DNA insert), heterozygous (1 copy of the T-DNA insert and 1 copy of the intact endogenous gene) and null (2 copies of the intact endogenous gene) segregants. This variation of distribution of genotypes is most likely due to the planting of seed directly from those obtained from the SALK Insitute.

Figure 5.3 - T-DNA identification of MN18 Arabidopsis homologue At3g56870. This particular T-DNA mutant line was chosen as the insertion was located early within the first exon. Red box highlights the SALK line selected. Blue box represents the Arabidopsis chromosome, dark green arrow above the chromosome box represents the direction of the exons of the gene candidate and arrows below the chromosome box represent the orientation of the T-DNA insert within the chromosome, beginning from the left border of the insert and into the flanking plant DNA sequence. Figure derived from T-DNA express (http://signal.salk.edu/cgi-bin/tdnaexpress).

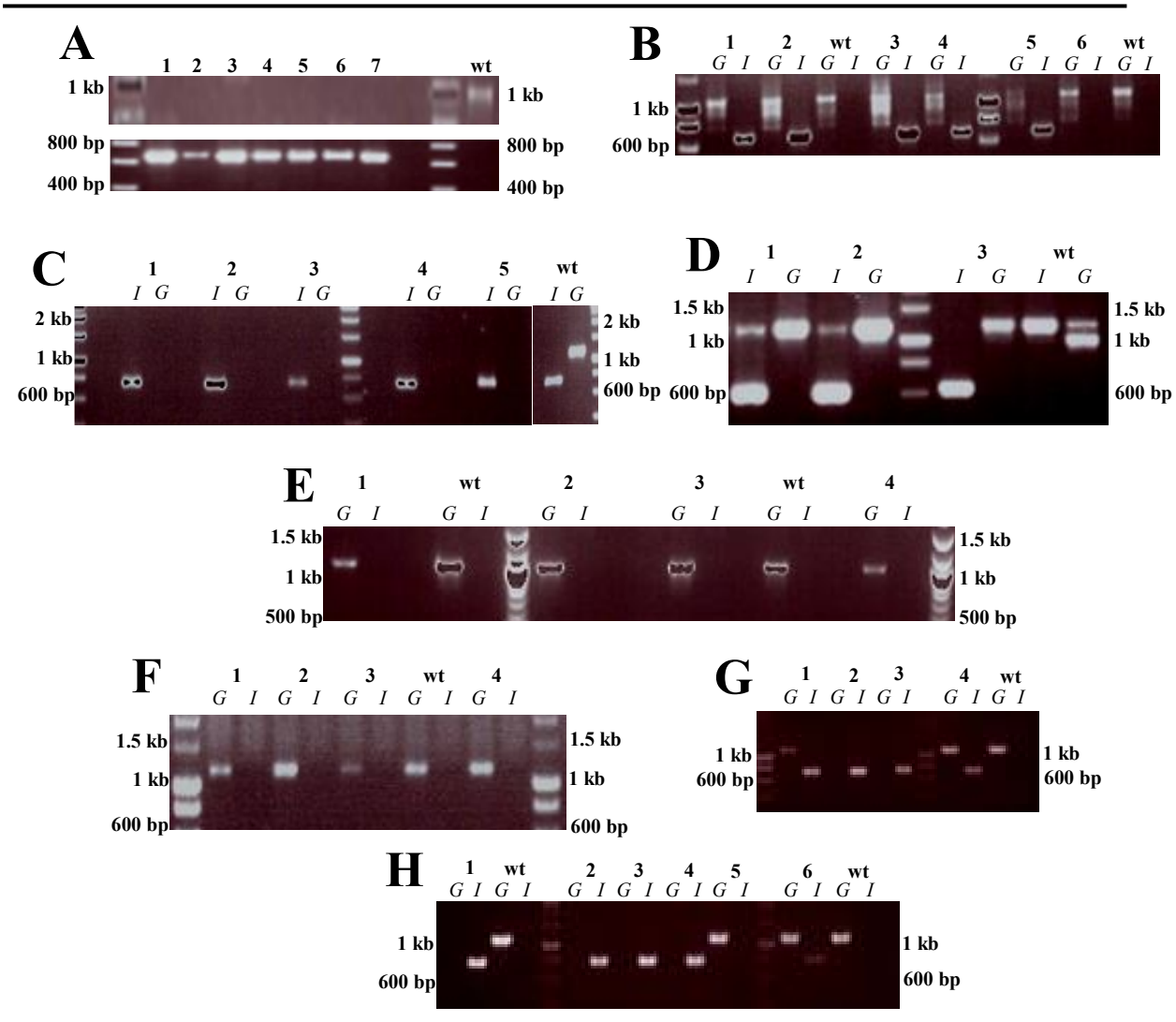


Figure 5.4 - Genotyping of Arabidopsis T-DNA mutants. PCR of the endogenous genes (*G*) and the T-DNA insert sequence (*I*) for MN02 (**A**, top – *G* and bottom – *I*), MN09 (**B**), MN11 (**C**), MN18 (**D**, with a positive control present at approximately 1.1 kb for all lanes), MN24 (**E**), MN25 (**F**), MN26 (**G**) and MN29 (**H**). Presence of the endogenous gene and an absence of insert indicate a null segregants for the T-DNA insertion (or wild-type [wt]), presence of the endogenous gene and insert fragment indicate a heterozygous segregant; and presence of the insert fragment and an absence of the endogenous gene indicate a homozygous segregant. A 1 kb DNA molecular weight marker is shown; but not all band sizes are listed.

Table 5.3 - Genotyping of Arabidopsis T-DNA seed stocks. Arabidopsis novels which were found to have homozygous and/or heterozygous plants were taken for further experimentation (*green*), while novels with only null segregants were not taken further (*red*).

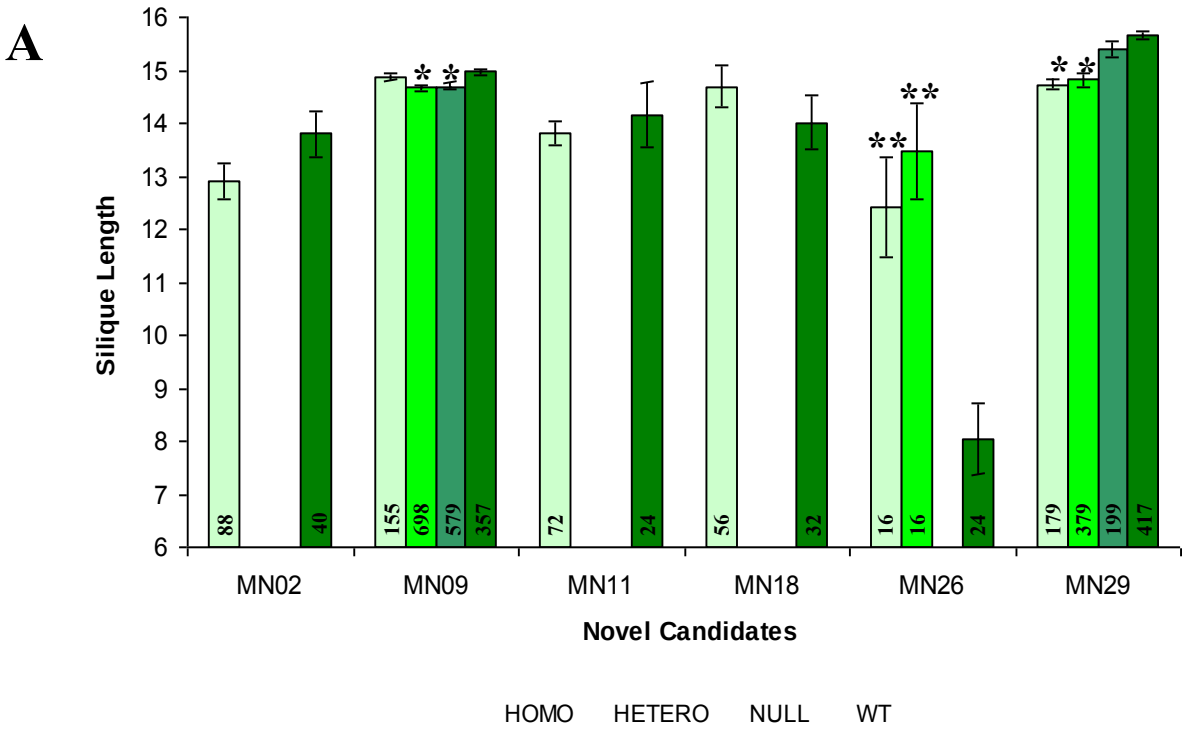
Novel	Homozygous	Heterozygous	Null	Wild-type (control)
MN02	16	-	-	5
MN09	9	36	32	19
MN11	9	-	-	3
MN18	7	-	-	4
MN24	-	-	54	24
MN25	-	-	20	13
MN26	2	2	-	4
MN29	10	22	11	25

5.3.2 – Reproductive morphology of Arabidopsis T-DNA plants

The reproductive morphology attributes of silique length, seeds per silique and pollen viability were measured for novel Arabidopsis T-DNA plants which were found to have homozygous and/or heterozygous T-DNA segregants (Figure 5.5 and 5.6). Novel MN29 had decreased silique length, seeds per silique and pollen viability in the homozygous and heterozygous compared to the wild-type controls ($p < 0.05$). Since these reproductive attributes can be used as an indicator for identifying potential meiotic genes (Ross *et al.*,

1997; Bhatt *et al.*, 1999; Sanders *et al.*, 1999; Siddiqi *et al.*, 2000; Mercier *et al.*, 2001; Caryl *et al.*, 2003); MN29 was selected for further experimentation.

Interestingly MN09 was found to have both decreased silique length and seeds per silique in the heterozygous and null segregants compared to the wild-type controls and the homozygous segregants; but pollen viability was not statistically different in any of the T-DNA stock plants when compared to the wild-type controls. MN26 T-DNA homozygous and heterozygous plants were found to have increased silique length, however the seeds per silique and pollen viability were found to be unaffected when compared to the wild-type controls. The heterozygous segregants of MN26 appear to have decreased pollen viability (Figure 5.5 C), however this was not statistically significant; most likely due to the small sample size of these plants tested for pollen viability (n=1). Novel candidates 2, 11 and 18 did not display any significant differences in the reproductive attributes that were measured.



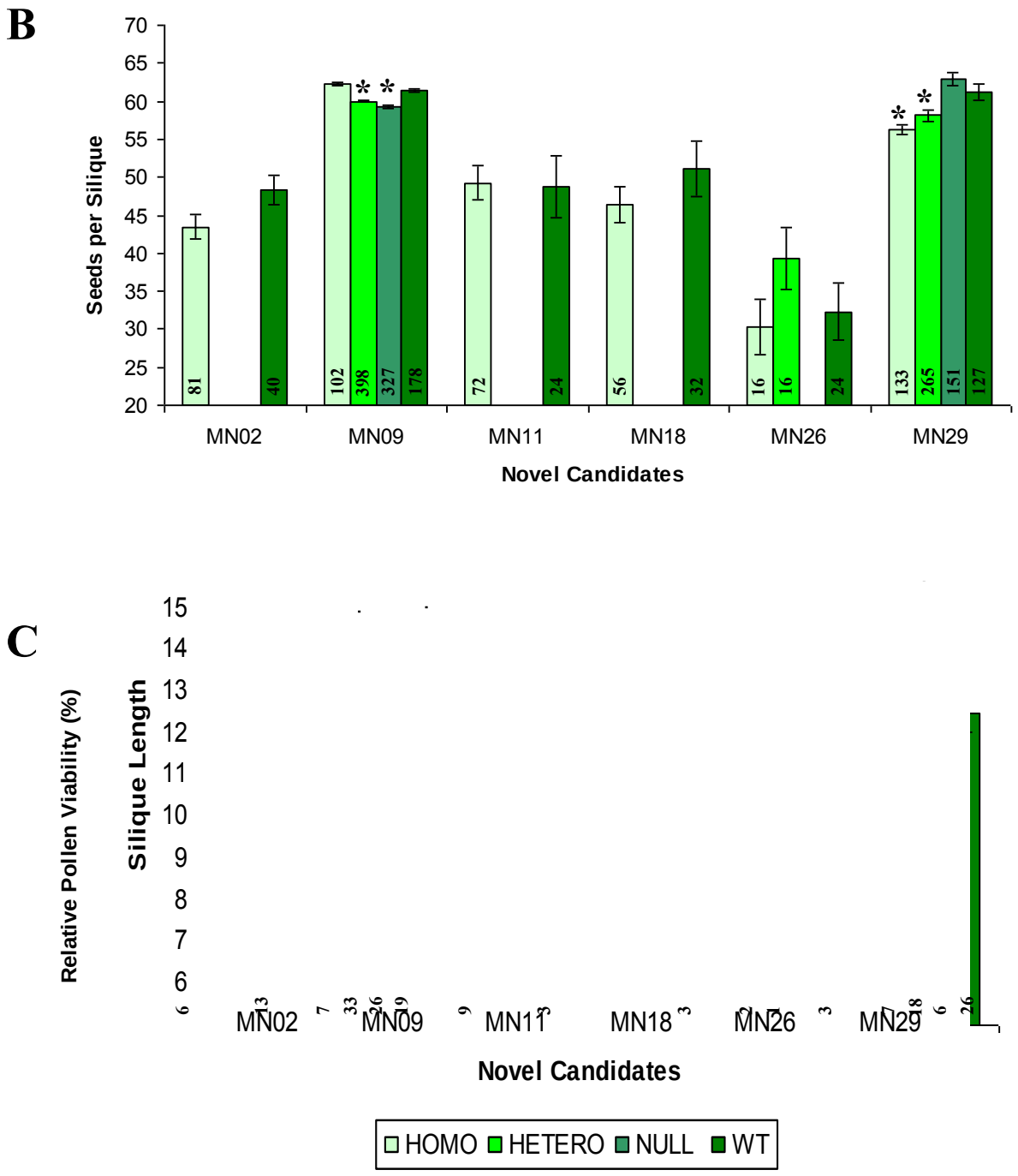


Figure 5.5 - Reproductive morphology of the T-DNA mutant Arabidopsis plants. (**A**) Silique length was found to be significantly decreased (*, $p < 0.05$) in MN09 (hetero and nulls) and MN29 (homo and hetero) and significantly increased (**, $p < 0.05$) in MN26 (homo and hetero). (**B**) Seeds per silique were found to be significantly decreased (*, $p < 0.05$) in MN09 (hetero and nulls) and MN29 (homo and hetero). (**C**) Pollen viability was found to be significantly decreased (*, $p < 0.05$) in MN29 (homo and hetero). Values within columns represent sample size (n) for silique length (**A,**

n=silique number), seeds per silique (**B,** n=silique number) and pollen viability (**C,** n=plants, 500 pollen grains per plant). Homo = homozygous, hetero = heterozygous and wt = wild-type.

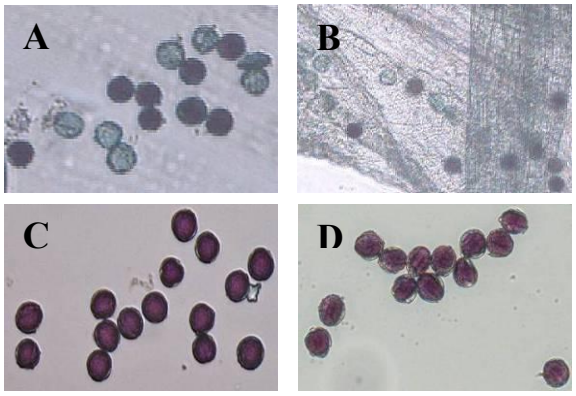


Figure 5.6 - Pollen Viability of MN29 Arabidopsis T-DNA plants. Pollen viability of the Arabidopsis T-DNA plants was determined using Alexander staining (Alexander, 1969). Red = viable pollen, while green = non-viable pollen. (**A**) homozygous, (**B**) heterozygous (with plant tissue), (**C**) null and (**D**) wild-type. Magnification of A, C and D is 400x and B is 200x.

5.3.3 – Vegetative morphology of MN29 Arabidopsis T-DNA plants

Vegetative measurements for MN29 were conducted (Figure 5.7). Both rosette radius and plant height were significantly decreased, $p<0.05$ and $p<0.1$ respectively, in the MN29 homozygous and heterozygous T-DNA plants when compared to the wild-type and null segregants. These results indicate that the MN29 HAT-B gene may also have a possible role in vegetative development and growth.

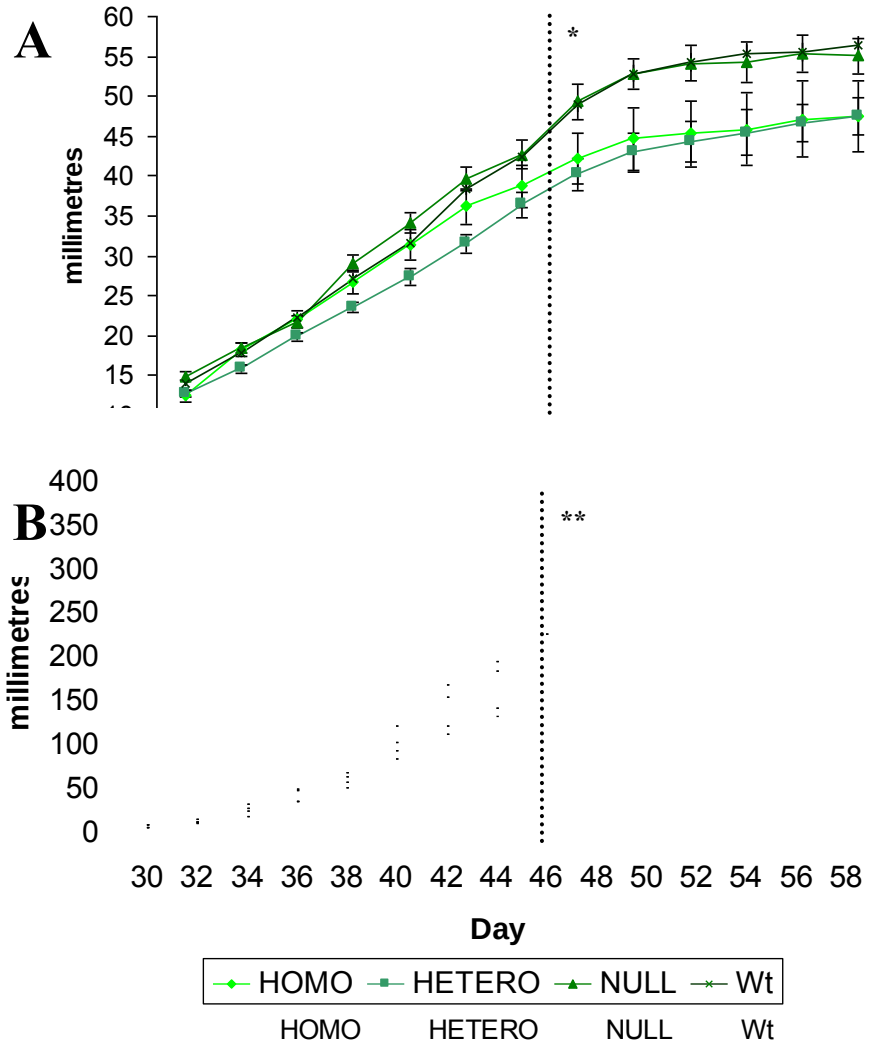


Figure 5.7 - Vegetative morphology of MN29 T-DNA Arabidopsis plants. MN29 homozygous and heterozygous T-DNA plants were found to have a significantly smaller (* $p<0.05$ and ** $p<0.1$) rosette radius (**A**) and plant height (**B**) when compared to null and wild-type plants. Homo = homozygous, hetero = heterozygous, null = null segregants and wt = wild-type.

5.3.4 – Expression of MN29 in Arabidopsis T-DNA plants

The expression of the *MN29 HAT-B* gene was investigated in the Arabidopsis T-DNA mutant plants and wild-type controls using semi-quantitative reverse transcriptase (RT) PCR expression analysis (Figure 5.8). *MN29* expression was decreased in homozygous tissues but not completely „knocked-out" (Figure 5.8). Confirmation of what was being amplified in these reactions compared to what was expected (e.g. knocked out expression of MN29), was confirmed through sequence analysis of the RT-PCR products. Heterozygous T-DNA segregants were also observed to have decreased leaf *MN29* expression from null segregants

and wild-type plants, but not decreased to the same extent as the homozygous T-DNA segregant plants (Figure 5.8). Interestingly *MN29* expression in bud tissues from heterozygous T-DNA segregants appears to be at the same expression level as is shown in the null segregants and wild-type plants (Figure 5.8). As expected, the null segregants showed a similar level of *MN29* expression in both leaf and bud tissues when compared to wild-type plants. Despite the addition of the same amount of RNA within the RT-PCR expression analysis there appears to be some differential expression occurring between the replicates of the genotyes, which is especially evident in the homozougous lines.

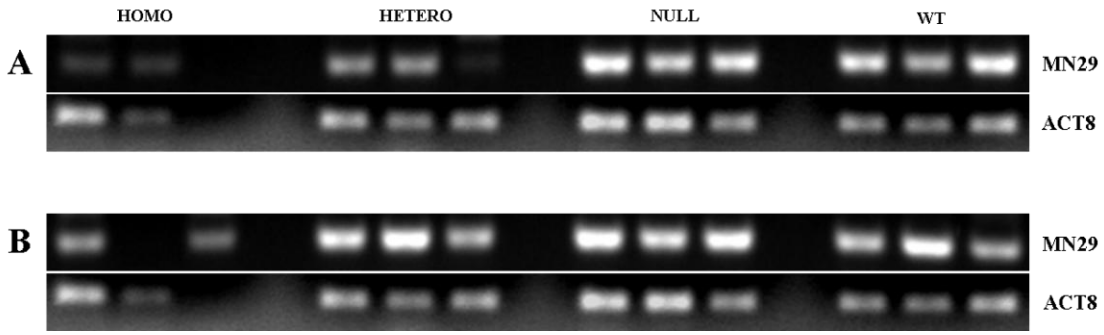


Figure 5.8 - MN29 gene expression analysis in MN29 Arabidopsis T-DNA plants. Expression analysis of MN29 transcripts was conducted in leaf (**A**) and bud (**B**) tissue from homozygous (HOMO), heterozygous (HETERO), null (NULL) segregants and wild-type (WT) plants. MN29 expression is present in both homozygous and heterozygous plants in leaf and bud tissues. Homozygous MN29 expression is decreased when compared to heterozygous, null and wild-type MN29 expression in both leaf and bud tissues; while heterozygous MN29 expression appears to be when compared to null and wild-type expression in leaf tissue only.

5.3.5 – Expression of meiotic genes in *MN29* Arabidopsis T-DNA plants

The expression of eight known meiotic genes was investigated in both leaf and bud tissues of the *MN29* T-DNA plants, using the *actin 8* (*ACT8*) gene as a control (Figure 5.9). RT-PCR analysis has shown an apparent decrease in expression of *HOP2* and *MND1* in leaf;

while in bud tissue, *SPO11*, *ASY1* and *DMC1* have an apparent increase in expression (Figure 5.9).

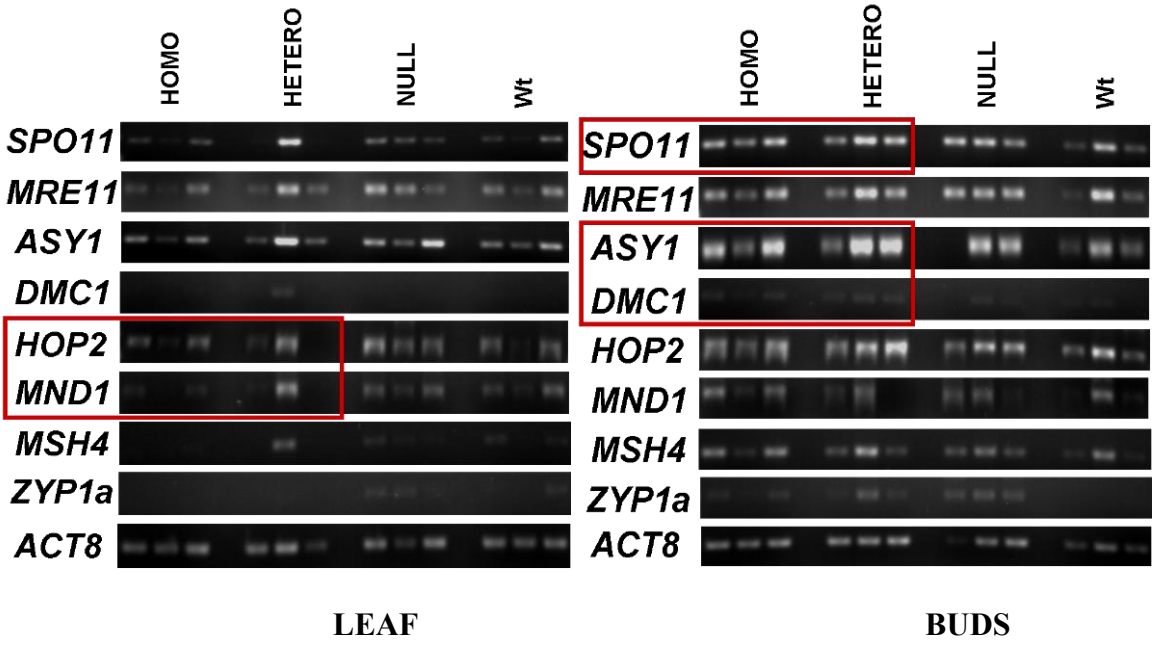
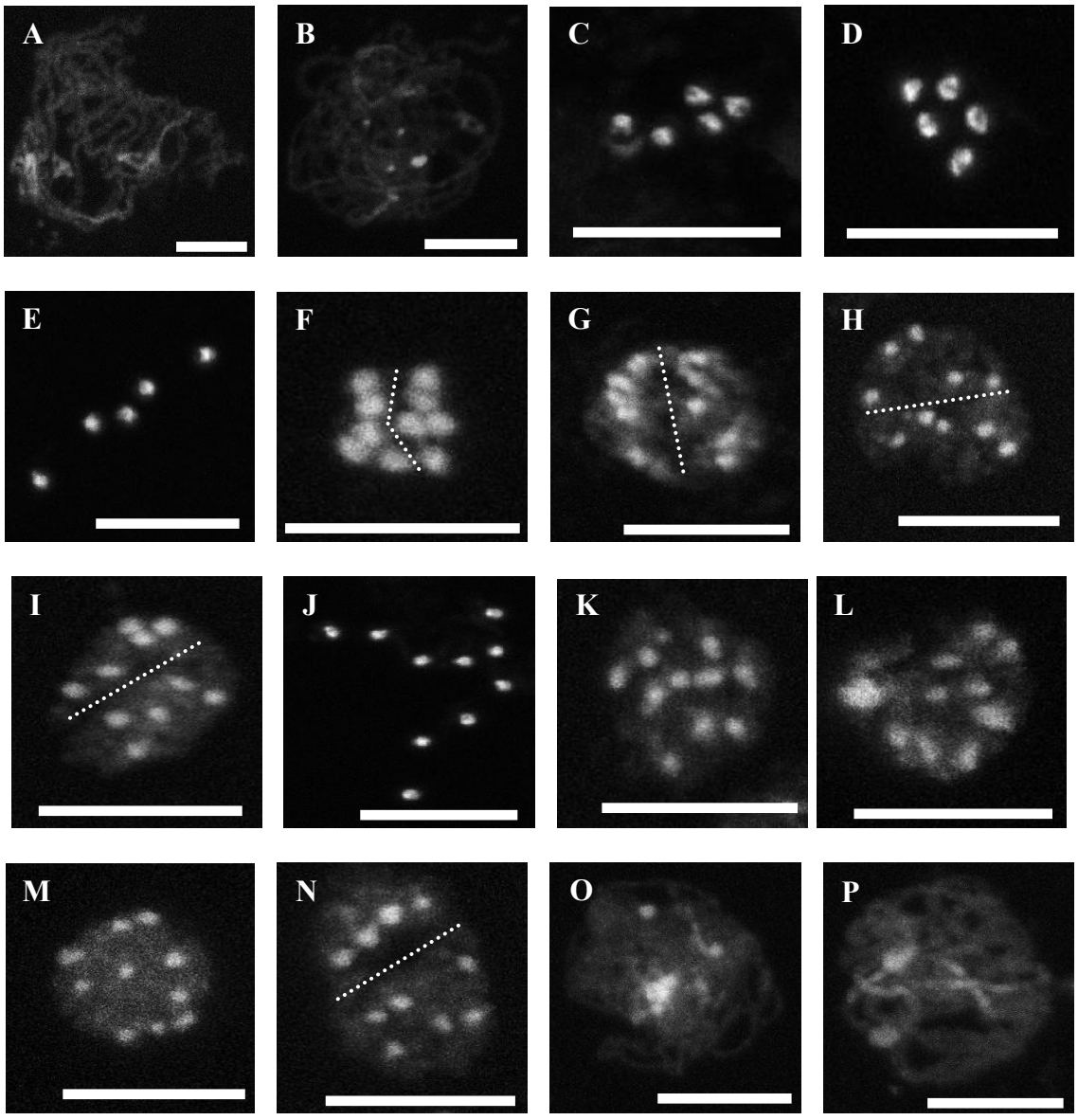


Figure 5.9 - Meiotic gene expression in MN29 T-DNA plants. RT-PCR expression analysis was examined using leaf and bud tissue, across eight known meiotic genes. An apparent decrease in *HOP2* and *MND1* expression in leaf tissue (*left-side, red box*) is observed, while an apparent increase in expression in *SPO11*, *ASY1* and *DMC1* expression has occurred in bud tissue (*right-side, red boxes*).

5.3.6 – Meiotic chromosome spreads of MN29 Arabidopsis T-DNA plants

Given that the semi-quantitative RT-PCR showed apparent differences in the meiotic gene expression of *SPO11*, *ASY1* and *DMC1* within the MN29 T-DNA plants, meiotic chromosome spreads using the Arabidopsis MN29 T-DNA plants were conducted. Figure 5.9 displays images of the visualised meiotic chromosome spreads for each of the Arabidopsis T-DNA segregants and the comparative wild-type plants. In homozygous and heterozygous T-DNA Arabidopsis, early meiosis appeared to be unaffected (Figure 5.10 A and B) when compared to wild-type plants (Figure 5.10 O and P). Correct formation of bivalent structures during late diakinesis (Figure 5.10 C and D), and meiotic stages of metaphase I (Figure 5.10 E) and anaphase I (Figure 5.10 F-I) were observed within the homozygous and heterozygous T-DNA segregants; as well as in the wild-type and null segregants plants (Figure 5.10 Q-V).

However, and interestingly, nuclei during late diakinesis with univalents (Figure 5.10 J-M) and uneven segregation during anaphase I (Figure 5.10 N) were also observed in homozygous and heterozygous T-DNA plants. This occurred in a ratio of approximately 1:3 for incorrect segregation (univalents and uneven anaphase I) to correct segregation (bivalents/metaphase I/anaphase I) (Figure 5.11).



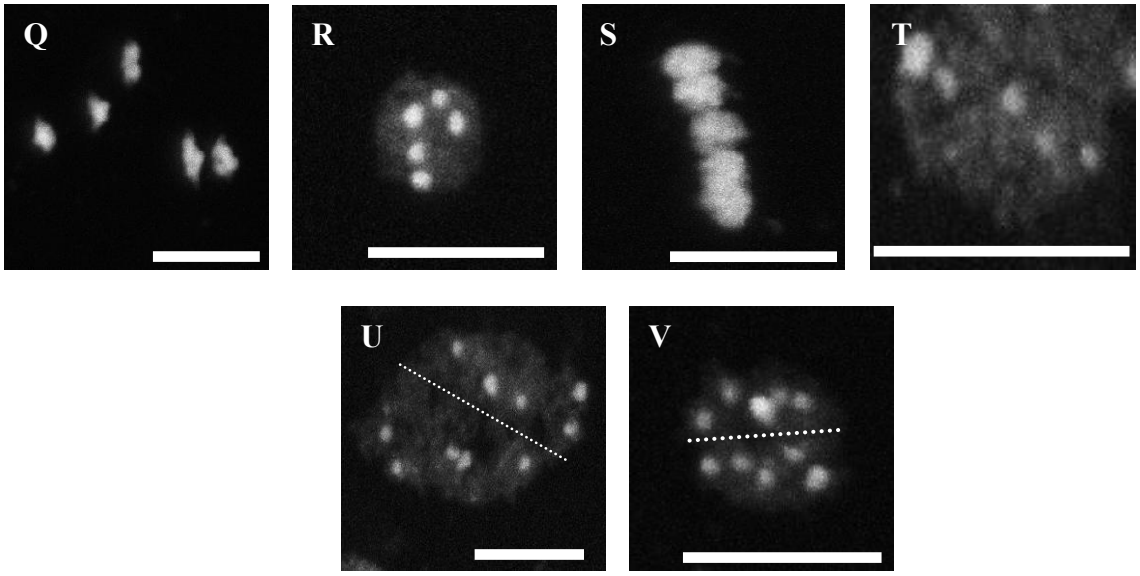


Figure 5.10 - Meiotic chromosome spreads of Arabidopsis MN29 T-DNA plants. Meiotic chromosome spreads of (**A**) Homo pachytene, (**B**) Hetero pachytene, (**C**) Homo late diakinesis, (**D**) Hetero late diakinesis, (**E**) Homo metaphase I, (**F**) and (**G**) Homo correct anaphase I, (**H**) and (**I**) Hetero anaphase I, (**J**) and (**K**) Homo incorrect late diakinesis with univalents, (**L**) and (**M**) Hetero incorrect late diakinesis with univalents, (**N**) Homo uneven anaphase I, (**O**) Wt zygotene, (**P**) Wt pachytene, (**Q**) Wt late diakinesis, (**R**) Null late diakinesis, (**S**) Wt metaphase I, (**T**) Null metaphase I, (**U**) Wt anaphase I and (**V**) Null anaphase I. Homo = homozygous segregant, hetero = heterozygous segregant, null = null segregant, wt = wild-type. Scale bar, 10µm.

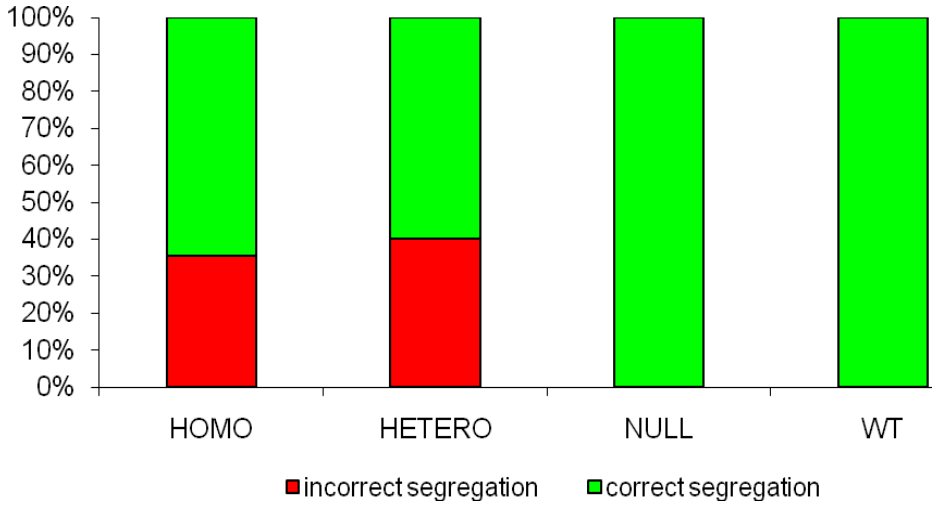


Figure 5.11 - Incorrect segregation within Arabidopsis MN29 T-DNA plants. Relative presence of nuclei undergoing incorrect segregation (univalents and uneven anaphase I) or nuclei with correct segregation (bivalents, metaphase I and anaphase I), was observed in a ratio of approximately 1:3 in Homozygous (HOMO) and Heterozygous (HETERO) MN29 T-DNA segregates when compared to null segregants and wild-type plants (WT). n = 50 nuclei per genotype.

5.4 – Discussion

5.4.1 – Filtering of Arabidopsis T-DNA candidates

In Chapter 2, novel candidates with wheat meiotic expression were filtered down to eight potential targets for further preliminary investigation using the appropriate T-DNA Arabidopsis mutant homologues. To determine if these candidate genes have an important role in meiosis (at least in Arabidopsis) the T-DNA mutant plants were further filtered using the experimental design that has been reported in the results. This has been summarised below in Table 5.4 for ease of reference to the following discussion.

Table 5.4 - Results of filtering the Arabidopsis T-DNA mutants. Candidates which were homozygous (H or Homo) or heterozygous (h or Hetero) for the T-DNA insert were taken further, along with their null segregants (N or Null). Only one mutant line that showed significantly decreased reproductive morphology attributes measured in the homo or hetero segregants (compared to the nulls and wild-type) was identified. Consequently, MN29 was investigated in full. X = no significant difference or unaffected, ↑ = significantly increased, ↓ = significantly decreased, A = affected.

Novel	Genotype	Reproductive Morphology									Vegetative Morphology						Meiotic Gene Expression			Meiotic Chromosome Spreads		
		Silique Length			Seeds per Silique			Pollen Viability			Rosette Radius			Plant Height								
		H	h	N	H	h	N	H	h	N	H	h	N	H	H	N	H	h	N	H	H	N
24	Null																					
25	Null																					
2	Homo	X		X	X		X	X		X												
11	Homo	X		X	X		X	X		X												
18	Homo	X		X	X		X	X		X												
26	Homo & Hetero	↑	↑	X	X	X	X	X	X	X												
9	Homo & Hetero	X	↓	↓	X	↓	↓	X	X	X												
29	Homo & Hetero	↓	↓	X	↓	↓	X	↓	↓	X	↓	↓	X	↓	↓	X	A	A	X	A	A	X

The first filtering step found that the seed stocks for Novels 24 and 25, which were proposed chromosome condensation proteins, only gave rise to null segregants (that is, no T-DNA insert); suggesting a vital role for these candidate genes within vegetative growth, most likely in chromosome condensation during mitosis. This is highly likely as Novels 24 and 25 have putative homology to the previously characterised chromosome proteins CAP-G and

CAP-D2, respectively, which have been shown to be involved in chromosome condensation during mitosis in other species and/or systems (Schmiesing *et al.*, 2000; Lavoie *et al.*, 2002 ; Ono *et al.*, 2003). However, other independent (different placement of T-DNA) lines of these particular candidates would need to be investigated to confirm these finding of lethality. For the remaining novels (2, 9, 11, 18, 26 and 29) homozygous and/or heterozygous segregants of the T-DNA insert were identified (Table 5.4).

Reproductive attributes including silique length, seeds per silique and pollen viability were used as the next filtering step. This is due to reduced fertility, indicated in these phenotypes, being regularly used as an indicator for identifying meiotic mutants (Ross *et al.*, 1997; Bhatt *et al.*, 1999; Sanders *et al.*, 1999; Siddiqi *et al.*, 2000; Mercier *et al.*, 2001; Caryl *et al.*, 2003) as these are the „end products" of meiosis. Homozygous plants from Novels 2, 11 and 18 were not found to have a significant difference in any of the three attributes measured and were therefore discontinued from further preliminary analysis at this time. The homozygous and heterozygous plants of the MN26 candidate were actually found to have increased silique length, yet were not found to have a corresponding increase in seeds per silique, while pollen viability was unaffected. This candidate was also discontinued from further preliminary analysis at this time.

Interestingly MN09 heterozygous and null segregants were found to have significantly lower silique lengths and seeds per silique compared to the homozygous segregants and wild-type plants (which were not significantly different). Perhaps a most likely explanation of a significant difference between hetero/null and homo/wt segregants is the small sample size of the homozygous plants (n=9), thereby not providing enough statistical power to determine a difference to the wild-type plants. Despite this observed difference in the silique morphology, the pollen viability in all three segregants was not found to be significantly affected; which either indicated that there is an effect of MN09 in female meiosis and not male meiosis, or most likely that this is an observable „artefact" from the MN09 T-DNA seed stock which leads to the differences observed in silique length and seeds per silique.

Only the MN29 T-DNA plants were found to have a significant decrease in all three reproductive attributes investigated when compared to both the null segregants and wild-type plants. Of particular note is that the pollen viability levels were reduced to approximately 75% when compared to the wild-type and null levels. Pollen viability has been shown to be varied in Arabidopsis mutants, ranging from total sterility to only a reduction of approximately 10% (Bleuyard *et al.*, 2004; Bleuyard and White, 2004; Li *et al.*, 2004; Kaur *et al.*, 2006; Berchowitz *et al.*, 2007). The three mutants of recombination repair proteins Atrad51, Atrad51 and Atxrcc3 were sterile in male meiosis (Bleuyard *et al.*, 2004; Bleuyard and White, 2004; Li *et al.*, 2004); while mutants of other recombination proteins *Atmsh4* and *Ataml5* (Kaur *et al.*, 2006; Berchowitz *et al.*, 2007) along with a meiotic master regulator *Atmus81* (Berchowitz *et al.*, 2007) show reduced male fertility (50%, 87% and 60-70% respectively).

Therefore it is possible for disruption of an important meiotic gene to only reduce pollen viability, and not by much, and not result in sterility. Another possibility of only observing reduced pollen viability instead of sterility, is that the expression of MN29 is only knocked-down and not knocked-out (discussed in more detail following). Correlation of expression of a given gene and pollen viability has been observed previously, where the reduction in mRNA of *AtPTEN1* has been closely related to the level of observed pollen viability (Gupta *et al.*, 2002); therefore it could be possible for the T-DNA mutants of MN29 to only be causing a reduction in pollen viability.

From this preliminary investigation of the eight candidates only one, MN29, showed the desired reproductive phenotypic difference to wild-type. While it could be that MN29 is the *only* candidate which has a meiotic role, this seems unlikely given that all eight candidates were chosen due to their meiotic expression/regulation and putative functions. It is likely that the other candidates may play a role during meiosis, yet due to the preliminary nature of this invesitigation their role has not been „noticed" due to a number of different reasons. One import reason may be the low number of plants which were used for investigation in some of the candidate lines, with the small number of plants not allowing for a significant difference

being able to be detected between the mutants and wild-type plants. Another possible cause for a number of lines not showing mutant phenotypes could be the inherent unreliability of the T-DNA knockouts in Arabidopsis, such as T-DNA inserts in promoter regions, introns and ends of the genes not having as much of an effect as thoses in the start of genes (discussed in more detail later). Therefore it is possible that some of the candidates which despite containing the T-DNA insert may not be affected by the insertions and therefore will not display a phenotype. Consequently multiple independent lines for a candidate are often looked at at one time, to minimise the chance of having a line which is not affected by the T-DNA insert and also as verification of any phenotypical differences observed. Therefore going forward for further investigation of these eight candidates, including MN29, this should be taken into consideration, with further analysis starting with a larger number of plants from multiple independent lines (with a number lines being available for each candidate) which are known to have altered gene expression or protein production.

5.4.2 – *MN29* Expression in Arabidopsis T-DNA segregants

Interestingly the expression of *MN29* in the Arabidopsis T-DNA segregants, while affected, was not „knocked-out"; with amplified PCR products being confirmed as *MN29* transcript through sequence analysis. The homozygous segregants were found to have decreased expression of *MN29* from the null segregants and wild-type plants in both the leaf and bud tissues, with expression in buds appearing to be stronger, indicating that there is differential expression of *MN29* between the two tissues. Heterozygous segregant *MN29* expression in leaf was decreased from null segregants and wild-type plants, but not to the same extent as the homozygous T-DNA segregants. This „middle ground" expression in the heterozygous T-DNA mutant is most likely due to the one normal allele of the MN29 gene present, where normal expression of MN29 could still be occurring; much like the T-DNA disrupted rice gene of *OsCHLH* (Jung *et al.*, 2003). The MN29 expression levels for homozygous T-DNA < heterozygous T-DNA < null/wild-type, provides strong evidence that

the MN29 gene is being expressed from both alleles in leaf. Interestingly, MN29 expression in bud tissue was almost the same level as the null segregants and wild-type plants, indicating that there is differential expression regulation occurring between tissues. Also interestingly is that there appeared to be variability between the replicates within the genotypes, despite the addition of the same amount of RNA within the RT-PCR analysis. This varied expression, especially in the homozygous genotype of MN29, could be evidence of differential expression of the MN29 gene due to the T-DNA insert, though this seems highly unlikely since the T-DNA will reside within the exact same spot within the gene since these plants were from the same T-DNA line. What seems the most likely explanation of this varied expression between replicates of genotypes is that it is the semi-quantitative nature of the RT-PCR analysis; therefore the quantitative method of Q-PCR should be used for further analysis.

Only achieving a knock-down of expression of *MN29* was unexpected as the majority of T-DNA mutagenesis in coding regions/exons results in a knock-out effect of expression (Krysan *et al.*, 1999; Wang, 2008). While it is unexpected the T-DNA insert in *MN29* has resulted in a knock-down effect, it is not uncommon to still observe a „mutant" phenotype (Wang, 2008). A number of genes containing T-DNA inserts in their coding region (including one within the first exon) have previously been observed to still have expression of the gene of interest, sometimes even increased expression or differential transcript size, in the homozygous T-DNA mutants; however no protein is produced within these homozygous mutants (Monte *et al.*, 2003; Delatte *et al.*, 2005; Pastuglia *et al.*, 2006).

Therefore it is highly possible that a T-DNA insertion in the second exon of *MN29*, the most 5′ T-DNA line available for *MN29*, could result in a knocked-down level of transcript expression in the homozygous and heterozygous T-DNA segregants. This reduced MN29 expression could be contributing to a reduction in protein production and/or the T-DNA insertion could give rise to no protein production or an aberrant or truncated protein as a T-DNA insert will by chance almost certainly contain stop codons which result in early translation termination (Krysan *et al.*, 1999; Wang, 2008). Unfortunately since there is no

published antibody for this particular protein, protein measurement in these mutants was not conducted.

Whether it is the observed decreased transcript expression or perhaps no protein production, the reduction in *MN29* expression has led to an observable difference/phenotype for the homozygous/heterozygous T-DNA segregants. Interestingly there was no significant difference in the measured phenotypes, both vegetative and reproductive, between the homozygous and heterozygous T-DNA segregants; presenting an apparent paradox since the *MN29* expression in buds of hetero plants is almost indistinguishable from the control plants. This perhaps provided some eveidence that it is the presence of the T-DNA insert within a transcript, which is observed to be produced since MN29 expression is present in homozygous T-DNA plants, which is the cause of the mutant phenotypes, most likely through a method of effecting translation of said transcript. It is unlikely to be through a method of RNAi-like feedback since transcripts were found to be present with RNA extractions. Further analysis to confirm the expression of the MN29 transcripts in these plants, or in other independent lines of MN29 T-DNA mutants, the quantitative method of Q-PCR should and would be used.

5.4.3 – Meiotic gene expression in MN29 Arabidopsis T-DNA segregants

Since the loci of MN29 in Arabidopsis corresponds to a Histone Acetyltransferase B protein, which have been reported to work in gene regulation (Eberharter and Becker, 2002), the expression of eight known meiotic genes was investigated in the MN29 Arabidopsis T-DNA segregants and wild-type plants. The eight meiotic genes examined covered the main processes of meiosis including pairing/recombination (*DMC1*, *HOP2* and *MND1*) recombination (*SPO11*, *MRE11* and *MSH4*) and synapsis (*ASY1*, *ZYP1*). These particular meiotic genes were chosen as they are „major" players within their given process or they are of particular interest to our research group, and therefore would like to know if they are affected by the absence/reduction of *MN29* expression. RT-PCR analysis of leaf and bud

tissue showed that the meiotic expression of the eight known meiotic genes was not dramatically affected from null segregants and wild-type plants. Observed differences in gene expression between the homozygous/heterozygous segregants when compared to the null segregants and wild-type plants were seen for *HOP2* and *MND1*. Expression of these two candidates appeared to be down-regulated in leaf tissues, while expression of *SPO11*, *ASY1* and *DMC1* appeared to be up-regulated in bud tissue. A possible reason for gene expression not being dramatically affected could be due to MN29 expression itself not being knocked-out, therefore minimising the effect on gene expression as some MN29 protein may still be produced and be functional.

It is more likely that direct interaction/regulation by MN29 with a given meiotic gene will result in an increase in expression, as histone acetyltransferance and the acetylation of histones play a positive role in gene expression (Eberharter and Becker, 2002). This suggests that a decrease in *MN29* would result in decreased acetylation of histones and thus a decrease in expression, which appears to be happening in *HOP2* and *MND1* in leaf tissues. The HOP2 and MND1 proteins are known to interact in a protein complex (Enomoto *et al.*, 2006; Pezza *et al.*, 2006) and are co-expressed (Chapter 3). This may imply that they are co-regulated and therefore it is not surprising that both genes together have been affected. This apparent down regulation of *HOP2* and *MND1* was not observed in bud tissue in the homozygous/heterozygous T-DNA plants, suggesting that *HOP2* and *MND1* may have differential regulation of expression in these tissues by MN29. This differential expression of genes by MN29 between tissues is most likely to be through the „complexing" of the MN29 protein with other proteins to assist in gene regulation; for example if MN29 forms a regulatory complex in leaf with protein X to regulate *HOP2* and *MND1*, however protein X is not expressed and produced in bud tissues and therefore MN29 is unable to exert regulatory control of *HOP2* and *MND1* in this tissue.

The up-regulation of *SPO11*, *ASY1* and *DMC1* is not likely due to direct regulation of MN29, as previously mentioned that the absence of MN29 would lead to decreased

acetylation of histones and therefore a decrease in gene expression. More likely, MN29 regulation of these genes is conducted indirectly; with MN29 regulating the regulator of the meiotic genes. It is interesting that there was increased expression of the *SPO11* gene, which is involved in double strand break (DSB) formation, and the *DMC1* gene, which is a protein involved in the repair of these DSBs. The increase in both *SPO11* and *DMC1* leads to the thought that MN29 may only be involved in the regulation of *SPO11* and due to the increase in the resulting DSBs, there is a subsequent increase (independent of MN29 regulation) of *DMC1* to repair the more than normal *SPO11* induced DSBs.

Again, as with the expression analysis results from Chapter 4, these results highlight the semi-quantitative nature of RT-PCR experiental analysis. These preliminary gene expression results would need to be further confirmed through the more quantitative expression analysis method of Q-PCR. More quantitative measurement of gene expression, especially within a complete gene knock-out of MN29, will allow for a definitive detemination of the effect of MN29 acetylation on the expression of meiotic genes in reproductive and vegetative tissues.

5.4.4 – Understanding the role of MN29 during meiosis

Meiotic chromosome spreads of the Arabidopsis MN29 T-DNA segregants and wild-type plants were conducted to determine the particular stages or process in which MN29 regulation is targeted. Early stages during meiosis I, zygotene and pachytene, appeared to be unaffected between the homozygous and heterozygous T-DNA segregates compared to the null segregants and wild-type plants (Figure 5.10 A, B, O, P). However later meiosis I was found to be different between the homozygous/heterozygous and null segregants and wild-type plants. Bivalent formation and correct later stages of metaphase I and anaphase I was observed in all plants: yet MN29 homozygous/heterozygous T-DNA segregants were observed to have a higher percentage of univalents and/or uneven segregation of chromosomes at anaphase I than null segregants and wild-type. Interestingly, the percentage

(approximately 75%) of correct chromosome segregation in homozygous/heterozygous T-DNA segregants is the same level observed for pollen viability.

Presence of univalents and uneven segregation of chromosomes is consistent with the notion that there is a problem within late recombination and synapsis since the chromosomes have not been able to „stick" together. Since the early meiotic stages of zygotene and pachytene appear to be unaffected in the homozygous and heterozygous segregants; providing evidence that later recombination or repair of recombination might be affected by the absence of MN29. An example of a gene with relatively normal early meiosis I but with observable univalents is the post-synaptic Arabidopsis mutant of *Atptd*, where reduced fertility and the presence of univalents were observed in mutants but they also had relatively normal synapsis and recombination nodules occurring (Wijeratne *et al.*, 2006). The *Atptd* mutant was proposed to have a role in crossover formation (Wijerante *et al*., 2006); therefore it is possible that MN29 also plays such a role.

Recently the histone acetyltransferase *Meiotic Control of Crossovers1* (*MCC1*), a member of the histone acetyltransferase GNAT/HAT1 superfamily to which the Arabiodpsis orthologue of MN29 (or *At*HAG2) belongs to (Pandey *et al*., 2002), was found to be involved in meiotic recombination and chromosome segregation in Arabidopsis (Parrella *et al.*, 2010). Overexpression in Arabidopsis results in the hyperacetylation of lysine residues 9 and 16 on histone 3, leading to an observable shift in chiasma distribution and the reduction of crossovers for chromosome 1 and 2 (Parrella *et al*., 2010). Mutations of *AtGCN5/HAG1*, another member of the GNAT/HAT1 superfamily, has also been found to affect plant development including sterility or partial fertility; however the mechanism underlying this remains unclear and, in particular, meiosis was not directly invesitigated (Bertrand *et al*., 2003; Vlachonasios *et al*., 2003; Parella *et al*., 2010). Differential acetylation of histones in certain regions of the yeast genome during meiosis has also been found to affect the recombination rates in these areas (Yamada *et al*., 2004; Merker *et al*., 2008). These results,

along with others in yeast, demonstrate compelling evidence of a link between histone de/acetylation and normal meiotic progression (Parrella *et al*., 2010).

As previously mentioned the MN29 T-DNA Arabidopsis homozygous and heterozygous segregants were shown to have significant differences in their reproductive architecture, aberrant meiotic chromosomes spreads and observable differences in expression of known meiotic genes. Since histone acetyltransferse proteins have also been found to be involved in gene expression and regulation (Eberharter and Becker, 2002), the role of MN29 in meiosis may also be acted through the gene regulation of other meiotic genes. Target meiotic genes of MN29 regulation would likely involved in late recombination and crossover formation and resolution; since during early meiosis no differences between mutants and nulls or wild-type were seen in the meiotic chromosome spreads. The aberrant reproductive morphology attributes and the defect in meiosis just mentioned provide strong evidence that the *MN29* gene is involved in Arabidopsis meiosis; however, analysis so far has been unable to pin-point an exact role for MN29.

Chapter 6

Chapter 6 – General Discussion and Future Directions

A common theme throughout this dissertation has been that the process of meiosis is highly complex and integrated. The main objective of this study has been to characterise candidates which are known to be involved in meiosis in other species or that are novel, yet meiotically expressed, in the polyploid wheat species *Triticum aestivum* or the model diploid species *Arabidopsis thaliana*.

6.1 – Using transcriptomics to investigate meiosis in wheat

6.1.1 – Identification of novel candidates with meiotic expression

Of the previously known 1,350 transcripts which have been observed to have meiotic regulation during wheat reproduction (Crismani *et al.*, 2006), sequence analysis at the time found that 81% (1,094) of these transcripts had either no annotation (826) or predicted annotations (268). These candidates therefore represent a cache of potentially important novel candidates involved in meiosis in bread wheat, and in other higher eukaryotes. Consequently further investigation of these meiotically expressed candidates has been focused on identifying candidates which have a high probability of being involved in meiosis. The novel candidates were taken through a number of filtering steps designed to maximise the possibility of obtaining a candidate which is involved in meiosis. Filtering steps included examining the expression profiles and how clusters formed with the known meiotic genes of *TaASY1* and *TaDMC1* (which were also present on the microarray). Expression profile clustering of the novel candidates with the known meiotic genes was conducted, as many genes with similar functions display similar expression patterns (Kaminski and Friedman, 2002).

Ultimately, to determine whether the refined list of novel candidates were involved in meiosis, homologues for these candidates in Arabidopsis needed to be identified and mutants

for those genes obtained from the SALK Institute‟s insertion mutant collection. In identifying eight novel candidates and the respective Arabidopsis homologues, this formed the basis for future experimentation reported later in the dissertation.

6.1.2 – Future meiotic candidates

While the current investigation has focused on novel candidates with similar expression profiles to the known meiotic genes of *TaASY1* and *TaDMC1*, there may be many genes which have roles in meiosis which do not have expression profiles similar to these known genes. Both *ASY1* and *DMC1* have been investigated through microarray analysis in yeast, with both the genes being classed as early expressed meiotic genes (Chu *et al.*, 1998; Primig *et al.*, 2000). The early expressed meiotic gene class is but one class of expression for meiotic and sporulation genes (Chu *et al.*, 1998; Primig *et al.*, 2000). Therefore it could be possible that the filtering steps within Chapter 2 have created a bias for early meiotic genes. While early meiotic genes are interesting for the processes of chromosome pairing, recombination and synapsis, other novel candidates which have later meiotic expression would also provide a valuable resource for further research in other meiotic processes.

6.2 – Genetic and functional analysis within polyploid wheat

In the following discussion of the analysis of known meiotic genes it must also be remembered that this preliminary analysis is not the „full story‟ of these genes, meaning that it is not the three homoeologues that were isolated and characterised but the expressed gene; since the genes were initially isolated from cDNA, the representation of expressed genes. However this approach in a preliminary investigation is actually a very good starting point since the functionality of the gene, especially those which function at the protein level, starts with the expression/regulation of the said gene. Therefore the genes isolated and investigated within this project represent either the particularly expressed but undefined homoeologue of

the individual genes *or* one of the two or three expressed homoeologous of the individual genes.

While the homoeologues of these genes will be slightly different at the sequence level, since they belong to the different progenitor genomes of wheat, it would not be expected that they would be vastly different enough to cause structural and functional differences especially since conservation between different species and kingdoms is also at a high level. However there might be a small amount of difference which contributes to slightly different expression, regulation or functionality and therefore should not be discounted. In further analysis of genes within the polyploidy organism of wheat it will be important to isolate and characterise *all three* homoeologues of each particular gene of interest to get the full understanding of the roles and functions of these genes.

6.3 – Investigation of known meiotic genes in bread wheat

In investigatng novel genes with meiotic expression and regulation, the known meiotic genes of *TaASY1* and *TaDMC1* were used as guides for the identification of possible novel early meiotic genes as they were also on the wheat microarray chip. While homologues of these genes have been extensively studied in diploid species for many years (Hollingsworth and Byers, 1989; Hollingsworth *et al.*, 1990; Bishop *et al.*, 1992; Dresser *et al.*, 1997; Li *et al.*, 1997; Doutriaux *et al.*, 1998; Nara *et al.*, 1999; Tarsounas *et al.*, 1999; Caryl *et al.*, 2000; Ding *et al.*, 2001; Hong *et al.*, 2001; Armstrong *et al.*, 2002; Nonomura *et al.*, 2004; Kant *et al.*, 2005), it has only been recently that the wheat ASY1 homologue was identified and characterised (Boden *et al.*, 2007; Boden, 2008; Boden *et al.*, 2009). In contrast, the DMC1 homologue had only been identified in wheat and was yet to be characterised in-depth (Petersen and Seberg, 2002; Petersen *et al.*, 2006; Khoo *et al.*, 2008).

6.3.1 – DMC1 and the HOP2:MND1 protein complex

DMC1 has been shown to be important for meiosis in the model diploid species of yeast, Arabidopsis and rice (Bishop *et al.*, 1992; Doutriaux *et al.*, 1998; Ding *et al.*, 2001). The homologous strand invasion capacity of the DMC1 protein is further enhanced by the protein complex of HOP2 and MND1 (Chen *et al.*, 2004; Petukhova *et al.*, 2005; Enomoto *et al.*, 2006; Pezza *et al.*, 2006), with these two proteins also being reported to be essential for meiosis in the model species (Leu *et al.*, 1998; Gerton and DeRisi, 2002; Schommer *et al.*, 2003; Kerzendorfer *et al.*, 2006; Panoli *et al.*, 2006). The homologous recognition by DMC1, enhanced by HOP2:MND1, made these meiotic genes interesting candidates to investigate given the strict homologous chromosome pairing process which occurs in wheat.

6.3.1.1 – *DMC1*, *HOP2* and *MND1* are conserved in wheat

The three meiotic genes of *DMC1*, *HOP2* and *MND1* were identified and isolated from the bread wheat cultivar Bob White 26 and these homologues were found to have high similarity to those identified in the diploid species genes and their corresponding proteins; especially DMC1, with the wheat protein displaying 53% identity to the yeast DMC1 protein. Further sequence analysis of the *Ta*DMC1 protein through computational three-dimensional modelling illustrated that this conservation of sequence, both at the DNA and protein level, corresponds to a high conservation of the tertiary structure of the DMC1 protein. Modelling of the HOP2 and MND1 proteins was not conducted; as unlike DMC1 (Kinebuchi *et al.*, 2004; Hikiba *et al.*, 2008) no homologue of either HOP2 or MND1 has been crystallised. HOP2 and MND1 protein crystallisation for the isolated wheat homologues was not a viable option during this project, as the HOP2 protein (which was heterologously expressed and extracted) was not at high enough purity or concentration for crystallisation work to commence.

Heterologously expressed DMC1 and HOP2 also possessed the ability to bind DNA, with both showing a preference for single stranded DNA which is in accordance with the

orthologous mouse proteins. However, the DNA binding ability of the MND1 protein, and therefore the HOP2:MND1 protein complex, was unable to be confirmed, given that MND1 heterologous expression and extraction was not successful. Therefore it remains unknown if the HOP2:MND1 protein complex displays the same preference for double stranded DNA as the homologue protein complex in yeast and humans, or for single stranded DNA as in mice. Also unable to be tested, since the *Ta*DMC1 protein extractions were not pure and MND1 was unable to be confirmed to be isolated, was the strand exchange and homology recognition ability of the three proteins. This would have been of interest as DMC1 and the HOP2:MND1 protein complex might be expected, in bread wheat, to play roles in the homology recognition of chromosomes in the homologous versus homoeologous chromosome pairing processes.

6.3.1.2 – DMC1 and the HOP2:MND1 protein complex - future directions for the hexaploid wheat proteins

Future investigation of the three meiotic proteins would firstly centre upon designing strategies for optimising heterologous protein expression. Expression and native extractions of these three proteins have been conducted using bacterial cells which are lysed to harvest the proteins, however since native state recombinant proteins are required, extraction methods are „gentle" to minimise protein disruption. This „gentle" extraction method could be limiting the amount of protein harvested, since heterologously expressed proteins can often be sequestered into inclusion bodies which are not accessed through this „gentle" extraction method. A possible way to improve this problem, along with the previously mentioned purity issues, could be to use a yeast protein expression and secretion system, such as the Invitrogen PichiaPink™ Yeast Expression System. Such a system allows for high protein expression as the protein expression vector, and thus the recombinant protein, contains secretion signals which exports the protein from the cell into the extracellular media. Along with having secretion signals, if the vector also contains a His-Tag this would allow for the "double

isolation" of the recombinant proteins of interest; with the protein being secreted into the extra-cellular media and then purified through a nickel column.

Once a high concentration of pure protein sample is achieved a number of investigations could be conducted. Firstly, with a high concentration and pure samples of HOP2 and MND1 proteins, crystallisation could be conducted to determine the protein structure of these two proteins; either individually/together/together binding DNA/in complex with DMC1; or all of the above. Having the structure of the protein complex resolved would provide further understanding of how these two proteins work, especially with respect to their role in binding DNA and DMC1. Secondly, having functional protein samples for all three meiotic proteins would enable the DNA binding and strand invasion ability for the proteins and complexes to be investigated.

Confirming strand invasion ability would then lead to investigating homology recognition of the three proteins. Within the *in vitro* strand invasion assay, the level of mis-match at the DNA level which is recognised by DMC1, and the HOP2:MND1 protein complex, could provide strong evidence that these proteins are involved in discriminating between homologous and homoeologous chromosome pairs. This could be conducted through substituting homologous (100% similarity) or non-homologous (25% similarity) single-stranded, labelled, invasion oligonucleotides with a homeologous (80%-95% similarity) oligonucleotide and observing if DMC1 is able to recognise the DNA mis-matches. If DMC1 was able to recognise the difference between the homoeologous strands then the assimilation of the labelled oligonucleotide would not occur or be visualised, thereby providing strong evidence the DMC1 plays a role in homologous chromosome recognition. In contrast, if DMC1 is not able to recognise the differences, the homoeologous oligonucleotide would be assimilated and therefore observed. Such a novel result would then enable the addition of the HOP2:MND1 protein complex during strand invasion, to determine if the protein complex, along with enhancing DMC1, is able to effect the homology recognition ability of the DMC1 protein. The HOP2 protein has been shown to have an effect on homology recognition, with

Leu and colleagues (1998) observing multiple chromosome associations in yeast meiosis in *hop2* mutants; therefore providing strong evidence that the HOP2:MND1 protein complex is involved in the homology recognition pathway.

If the meiotic proteins of DMC1 and HOP2:MND1 were found to be involved in homology recognition these proteins could then be manipulated to affect recognition. Manipulation could be achieved by producing knock-down mutants of the three genes in preference of producing knock-outs, which would more than likely be lethal. Wheat knock-down mutants for these genes, done individually or together, may reduce the strict chromosome pairing control. This would subsequently lead to a degree of homoeologous chromosome associations taking place, therefore providing a means of increasing genetic diversity in plant breeding programs.

6.3.2 – *Asynapsis 1* (*ASY1*)

6.3.2.1 – Interacting partners of *Ta*ASY1

The interacting protein partners of *Ta*ASY1 were investigated to further the understanding of the role of ASY1 during meiosis in wheat. The HORMA domain of the *Ta*ASY1 protein was used in an initial screen of an early meiotic cDNA library via a yeast two-hybrid approach. The yeast two-hybrid analysis identified four proteins which interacted with the isolated HORMA domain, and three of these were then further confirmed to interact with the full-length *Ta*ASY1 protein. Two of these interacting proteins were of particular interest; with *Ta*ASY1 physically interacting with the RPA 3 (*R*eplication *P*rotein *A*) protein of the RPA protein complex and also with a previously uncharacterised plant protein, termed HIP4 (*H*ORMA *I*nteracting *P*rotein).

Expression data of *TaRPA3* and *TaHIP4* were available through the previously conducted microarray analysis of wheat reproduction (Crismani *et al.*, 2006); with the *TaRPA3* early meiotic expression profile correlating with its interacting partner of *TaASY1*; in

contrast the late meiotic expression profile of *TaHIP4* was negatively correlated to *TaASY1* and *TaRPA3*. Expression of the *Ta*ASY1 interacting proteins were also affected in *Taasy1* RNAi mutants, indicating a feedback loop between the interacting proteins. Functional characterisation of these novel interacting partners of *Ta*ASY1 showed that the *Ta*RPA3 protein was highly conserved both in sequence and structure to other RPA3 homologues; whereas characterisation of *TaHIP4* found that this candidate has putative homology only to other plant proteins, all of which are also uncharacterised.

6.3.2.2 – Further characterisation of novel *Ta*ASY1 protein interactors and advancing the understanding of *Ta*ASY1 during meiosis

Since the *Ta*ASY1 interactor of *Ta*HIP4 is novel, further characterisation of this gene is vital. Perhaps one of the best ways forward in characterising this potential meiotic candidate would be through the analysis of the homologous Arabidopsis T-DNA mutant, such as the *Athip4* mutant line 13-005-1 from the Riken Bioresource center (since there is no SALK line currently available), which would provide insight into the role of this gene. Further research could focus on the production of an antibody for this protein. A *Ta*HIP4 antibody could be used for protein localisation of *Ta*HIP4 in wild-type tissues, where co-localisation with *Ta*ASY1 would be expected based on their confirmed interaction. The *Ta*HIP4 antibody could also be used to investigate the protein in the *Taasy1* RNAi mutants to observe what effect reduced *ASY1* expression has on *Ta*HIP4.

In contrast to *Ta*HIP4, homologues of *Ta*RPA3 have been previously described within the scientific literature. RPA3 is the smallest sub-unit, 14 kDa, of the RPA protein complex which also consists of RPA1, a 70 kDa sub-unit, and RPA2, a 32 kDa sub-unit (Wold, 1997). Mutants of two Arabidopsis RPA homologues, RPA1B (70b) and RPA2A (32a), are sensitive to methyl methanesulfonate (MMS) and UV-B, indicating a possible role for these particular sub-units in DNA repair/recombination repair (Ishibashi *et al.*, 2005; Xia *et al.*, 2006).

The RPA protein complex has been reported to be co-located with the recombinase proteins RAD51/DMC1; with the RPA complex appearing to gradually replace RAD51/DMC1 foci (Moens *et al.*, 2002). This replacement of RAD51/DMC1 with the RPA3 complex obviously relies on the recruitment of the complex to the chromosomes, which could occur through either a known interaction with the recombination proteins (Sakaguchi *et al.*, 2009) and/or the confirmed interaction with synaptonemal complex protein ASY1 (through an interaction with RPA3, as shown in Chapter 4). Synapsis of chromosomes and the synaptonemal complex (SC) has been proposed to, along with forming a physical attachment between a chromosome pair, help stabilise chromosomes during the process of recombination where there is a large amount of DNA rearrangement (Zickler and Kleckner, 1998). Therefore an interaction between the SC protein ASY1 and the recombination protein RPA, and its interaction with the recombinase proteins RAD51/DMC1, could provide further evidence of the SC helping to stabilise a chromosome pair during the recombination and repair process (Figure 6.1).

Despite the knowledge of RAD51 and DMC1, the SC element protein ASY1 and the RPA protein complex, the particular function of the smallest sub-unit (RPA3) is still largely unknown. Further understanding the role of RPA3, the RPA complex and its activity with ASY1 could readily be achieved through mutant analysis in a T-DNA Arabidopsis homologue. Using the publically available T-DNA mutant stocks, the protein location of the RPA protein complex and ASY1 within the Arabidopsis mutants of ASY1 and individual RPA sub-units, respectively, could be investigated. However, it is likely that any RPA3 mutant will be lethal, since RPA3 is important for the formation of the RPA protein complex (Henricksen *et al.*, 1994); and since to date the RPA3 sub-unit has found to be a single-gene locus, there is no chance of redundancy compensating for the absence of RPA3 (Sakaguchi *et al.*, 2009). However a problem which might then be faced is that the absence of the single gene copy of RPA3 could result in lethality, therefore it may be useful to use an RNAi

approach with the RPA3 gene to knock-down expression and to observe the subsequent effect either in Arabidopsis or wheat.

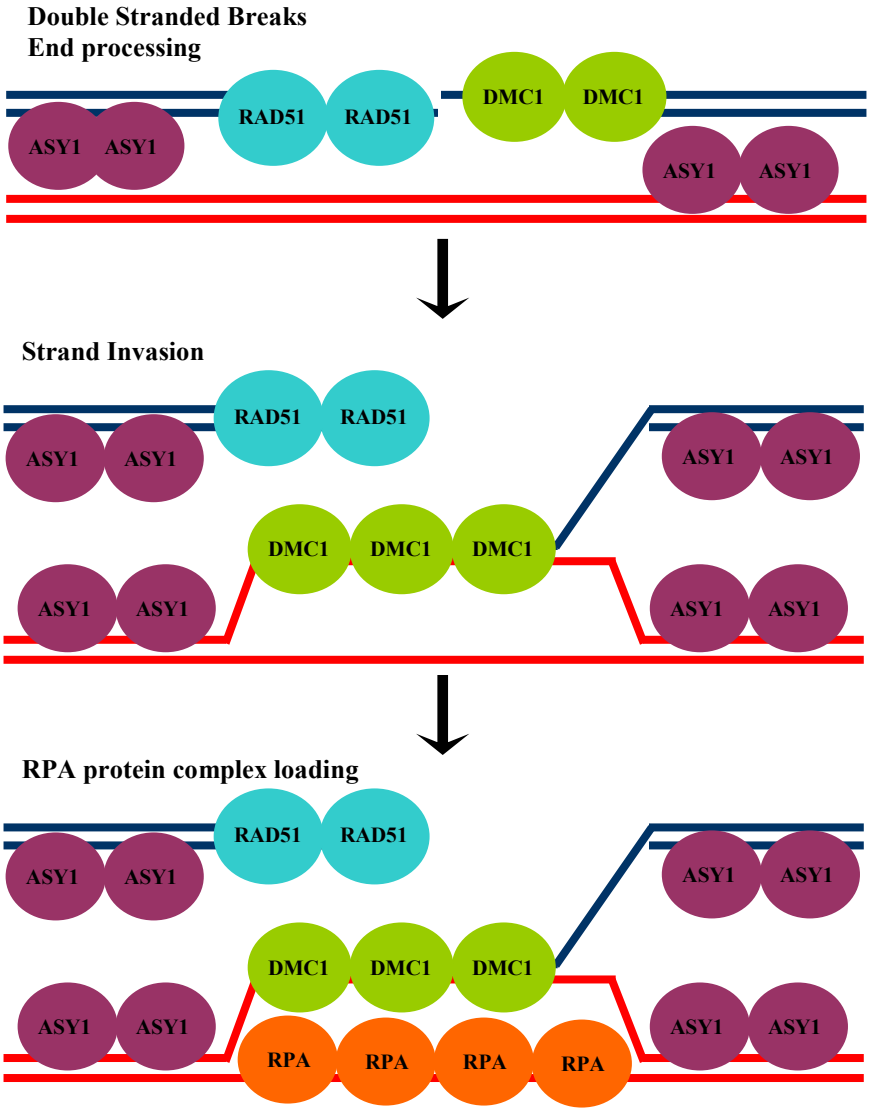


Figure 6.1 - Association of RPA3, ASY1 and recombination proteins during early meiosis. RAD51 and DMC1 are recruited to the processed single stranded overhangs, while concurrently ASY1 is loaded along the chromosomes as axial/lateral elements of the synaptonemal complex (SC). Following homologous strand invasion, conducted by DMC1, the recombination proteins and/or the SC element protein ASY1 possibly recruit the RPA protein complex, which is known to play a role in recombination repair.

6.4 – Using Arabidopsis homologues to investigate novel candidates with meiotic expression

As discussed previously, the microarray analysis of wheat reproduction has provided a powerful platform for the investigation of both known and novel candidates with meiotic expression and regulation. 1,094 candidates which were placed through a number of filtering steps resulted in eight Arabidopsis homologue candidates, also uncharacterised during meiosis, being identified for further study. Using individual insertional T-DNA mutants for each candidate, analysis was focused on the identification of a candidate which had a role during meiosis.

6.4.1 – MN29 plays a role in meiosis?

Of the eight Arabidopsis candidates which were taken through for further investigation, the MN29 (a histone acetyletransferase B) protein was identified as having a putative, yet un-determined, role in meiosis. Interestingly MN29 transcript expression was observed not to be knocked-out in homozygous and heterozygous T-DNA mutants but had knocked-down expression. This reduced expression (and not knocked-out expression) may be contributing to why the following analyses were not observed to have dramatic differences, which are often seen in meiotic mutants (Ross *et al.*, 1997; Caryl *et al.*, 2000; Grelon *et al.*, 2001).

MN29 Arabidopsis T-DNA homozygous and heterozygous mutants had reproductive attributes that were significantly different to wild-type and nulls. This included decreased silique length, seeds per silique and pollen viability. T-DNA mutants were also observed to have decreased vegetative growth, measured through rosette radius and plant height, indicating that the role of MN29 is not exclusive to meiosis. Further expression analysis of eight meiotic genes in these mutants found that there were small variable differences in leaf and bud expression of *AtHOP2* and *AtMND1* (reduced in leaf) and *AtSPO11*, *AtASY1* and

AtDMC1 (increased in bud). An increase in *SPO11*, which is involved in catalysing double stranded breaks, would result in increased formation of double stranded breaks during meiosis, therefore it is interesting that the proteins involved in the repair of these breaks, *DMC1* or *RAD51*, are also up-regulated; which could be due to feedback between the double stranded breaks and the mechanisms which repair them. Also interesting is the observed reduction in both *AtHOP2* and *AtMND1* expression in leaf. The HOP2 and MND1 proteins have previously been confirmed to form a protein complex (Petukhova *et al.*, 2005; Enomoto *et al.*, 2006; Pezza *et al.*, 2006) and are co-expressed in wheat (Chapter 3). This co-expression of HOP2 and MND1 indicates that they may also be co-regulated. Could it be that MN29 has a role in this regulation?

Since the reproductive attributes and meiotic genes were affected in the MN29 Arabidopsis T-DNA mutants, this provides preliminary evidence that MN29 does have a role during meiosis. To further investigate whether this is true and if so, when MN29 action is important during meiosis, meiotic chromosome spreads were conducted. Early stages of prophase I, zygotene and pachytene, appeared to be unaffected during meiosis in the homozygous and heterozygous T-DNA mutants compared to null segregants and wild-type plants. However, some spread preparations for the homozygous and heterozygous T-DNA mutants clearly identified multiple nuclei with univalents and uneven segregation of chromosomes during anaphase I, when compared to null segregants and wild-type plants.

6.4.2 – Further investigating MN29 to understand its role during meiosis

Future analysis of the MN29 candidates should include the identification of other Arabidopsis T-DNA MN29 mutants which have knocked-out MN29 gene expression. A complete knock-out of MN29 would most likely provide a stronger phenotype; which would allow for a more definitive characterisation of the phenotype observed. In testing gene expression, a wider range of meiotic genes – with focus on the later recombination proteins – would be needed to identify which meiotic genes are targeted by MN29. Also in conducting

gene expression analysis, a more quantitative methodology such as Q-PCR could be undertaken for a definitive understanding of which genes are regulated by MN29 and by how much. In-depth recombination and crossover activity assays could also be conducted in a MN29 mutant to determine how much recombination has been altered when compared to wild-type. Upon confirmation of MN29 activity during meiosis in Arabidopsis, MN29 could then be investigated in wheat; looking at mutations of MN29 (either via EMS and TILLING or RNAi) and conducting such previously mentioned techniques to confirm a similar role in wheat meiosis.

6.5 – The bigger picture

The fundamental objectives of the investigation of meiotic genes in wheat, especially those involved in homologous chromosome pairing and recombination, are to 1) increase the understanding of meiotic processes in higher eukaryotes and 2) provide the ability to access genetic diversity from closely related but distinct species via meiosis and recombination in a controlled and viable manner. Achieving these objectives will result in increasing the genetic diversity available for plant breeding programs. Using the wheat species of *Triticum aestivum*, an allohexaploid with multiple genomes which behaves like a diploid at meiosis, enables these objectives to be investigated. Furthermore, by identifying and characterising novel genes involved in homologous chromosome pairing and recombination within the hexaploid genetic environment, insight into how these genes and proteins function in diploid species will be generated.

References

References

Able, J.A., Langridge, P., Milligan, A.S (2007) Capturing diversity in the cereals: many options but little promiscuity. *Trends in Plant Science* **12**, 71-79.

Agarwal, S., Roeder, G. (2000) *Zip3* provides a link between recombination and synaptonemal complex proteins. *Cell* **102**, 245-255.

Al-Kaff, N., Knight, E., Bertin, I., Foote, T., Hart, N., Griffiths, S., Moore, G. (2008) Detailed dissection of the chromosomal region containing the *Ph1* locus in wheat *Triticum aestivum*: with deletion mutants and expression profilling. *Annuals of Botany* **101**, 863-872.

Alexander, M.P. (1969) Differential staining of aborted and nonaborted pollen. *Biotechnic & Histochemistry* **44**, 117-122.

Alonso, J.M., Stepanova, A.N.*, et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657.

Andrews, J., Bouffard, G.G., Cheadle, C., Lu, J., Becker, K.G., Oliver, B. (2000) Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Research* **10**, 2030-2043.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.

Aragón-Alcaide, L., Reader, S., Miller, T., Moore, G. (1997) Centromeric behaviour in wheat with high and low homoeologous chromosomal pairing. *Chromosoma* **106**, 327-33.

Aravind, L., Koonin, E. (1998) The HORMA domain: a common structural denominator in mitotic checkpoints, chromosome synapsis and DNA repair. *Trends in Biochemical Science* **23**, 284-286.

Armstrong, S.J., Caryl, A.P., Jones, G.H., Franklin, F.C.H. (2002) Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in *Arabidopsis* and *Brassica*. *Journal of Cell Science* **115**, 3645-3655.

Armstrong, S.J., Franklin, F.C.H., Jones, G.H. (2003) A meiotic time-course for *Arabidopsis thaliana*. *Sex Plant Reproduction* **16**, 141-149.

Armstrong, S.J., Jones, G.H. (2003) Meiotic cytology and chromosome behaviour in wild-type *Arabidopsis thaliana*. *Journal of Experimental Botany* **54**, 1-10.

Averbeck, N., Sunder, S., Sample, N., Wise, J.A., Leatherwood, J. (2005) Negative control contributes to an extensive program of meiotic splicing in fission yeast. *Molecular Cell* **18**, 491-498.

Azumi, Y., Liu, D., Zhao, D., Li, W., Wang, G., Hu, Y., Ma, H. (2002) Homolog interaction during meiotic prophase I in *Arabidopsis* requires the *SOLO DANCERS* gene encoding a novel cyclin-like protein. *EMBO Journal* **21**, 3081-3095.

Bartel, P.L., Fields, S. (Eds) (1997) 'The yeast two-hybrid system.' Oxford University Press, New York

Berchowitz, L.E., Francis, K.E., Bey, A.L., Copenhaver, G.P. (2007) The role of *AtMUS81* in interference-insensitive crossovers in *A. thaliana*. *PLoS Genetics* **3**, e132.

Bertrand, C., Bergounioux, C., Domenichini, S., Delarue, M., Zhou, D.X. (2003) *Arabidopsis* histone acetyltransferase *At*GCN5 regulates the floral meristem activity through the WUSCHEL/AGAMOUS pathway. *Journal of Biological Chemistry* **278**, 28246-28251.

Bestor, T.H. (2000) The DNA methyltransferases of mammals. *Human Molecular Genetics* **9**, 2395-2402.

Bhatt, A.M., Lister, C., Page, T., Fransz, P., Findlay, K., Jones, G.H., Dickinson, H.G., Dean, C. (1999) The *DIF1* gene of Arabidopsis is required for meiotic chromosome

segregation and belongs to the REC8/RAD21 cohesin gene family. *Plant Journal* **19**, 463-472.

Bishop, D.K., Park, D., Xu, L., Kleckner, N. (1992) DMC1: a meiosis-specific yeast homolog of *E. coli recA* required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69**, 439-456.

Bleuyard, J.-Y., Gallego, M.E., White, C.I. (2004) Meiotic defects in the *Arabidopsis rad50* mutant point to conservation of the MRX complex function in early stages of meiotic recombination. *Chromosoma* **113**, 197-203.

Bleuyard, J.-Y., White, C.I. (2004) The Arabidopsis homologue of Xrcc3 plays an essential role in meiosis. *EMBO Journal* **23**, 439-449.

Blumenthal, T., Spieth, J. (1996) Gene structure and organization in *Caenorhabditis elegans*. *Current opinions in Genetics and Development* **6**, 692-698.

Boden, S.A. (2008) Investigating chromosome pairing in bread wheat using *ASYNAPSIS I*. PhD thesis, The University of Adelaide.

Boden, S.A., Langridge, P., Spangenberg, G., Able, J.A. (2009) *Ta*ASY1 promotes homologous chromosome interactions and is affected by deletion of *Ph1*. *Plant Journal* **57**, 487-497.

Boden, S.A., Shadiac, N., Tucker, E.J., Langridge, P., Able, J.A. (2007) Expression and functional analysis of *TaASY1* during meiosis of bread wheat (*Triticum aestivum*). *BMC Molecular Biology* **8**, 65.

Bowers, J.E., Chapman, B.A., Rong, J., Paterson, A.H. (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature Structural and Molecular Biology* **422**, 433-438.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., Görlach, J. (2001) Growth stage–based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell* **13**, 1499-1510.

Burns, N., Grimwade, B., Ross-Macdonald, P.B., Choi, E.-Y., Finberg, K., Roeder, G.S., Snyder, M. (1994) Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes and Development* **8**, 1087-1105.

Carballo, J.A., Johnson, A.L., Sedgwick, S.G., Cha, R.S. (2008) Phosphorylation of the axial element protein Hop1 by Mec1/Tel1 ensures meiotic interhomolog recombination. *Cell* **132**, 758-770.

Caryl, A.P., Armstrong, S.J., Jones, G.H., Franklin, F.C.H. (2000) A homologue of the yeast *HOP1* gene is inactivated in the *Arabidopsis* meiotic mutant *asy1*. *Chromosoma* **109**, 62-71.

Caryl, A.P., Jones, G.H., Franklin, F.C.H. (2003) Dissecting plant meiosis using *Arabidopsis thaliana* mutants. *Journal of Experimental Botany* **54**, 25-38.

Chelysheva, L., Gendrot, G., Vezon, D., Doutriaux, M.-P., Mercier, R., Grelon1, M. (2007) Zip4/Spo22 Is required for class I CO formation but not for synapsis completion in *Arabidopsis thaliana*. *PLoS Genetics* **3**, e83.

Chelysheva, L., Vezon, D., Belcram, K., Gendrot, G., Grelon, M. (2008) The Arabidopsis BLAP75/Rmi1 homologue plays crucial roles in meiotic double-strand break repair. *PLoS Genetics* **4**, e1000309.

Chen, C.-B., Xu, Y.-Y., Ma, H., Chong, K. (2005) Cell biological characterization of male meiosis and pollen development in rice. *Journal of Integrative Plant Biology* **47**, 734−744.

Chen, Y.-K., Leng, C.-H.*, et al.* (2004) Heterodimeric complexes of Hop2 and Mnd1 function with Dmc1 to promote meiotic homolog juxtaposition and strand assimilation. *Proceeding of the National Academy of Sciences of the United States of America* **101**, 10572-10577. Copyright (2004) National Academy of Sciences, U.S.A.

Chien, N.C., Robertso, V.A., Giusti, A.M., Scharff, M.D., Getzoff, E.D. (1989) Significant structural and functional change of an antigen-binding site by a distant amino acid

substitution: proposal of a structural mechanism. *Proceeding of the National Academy of Sciences of the United States of America* **86**, 5532-5536.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., Herskowitz, I. (1998) The transcriptional program of sporulation in budding yeast *Science* **282**, 699-705.

Cohen, P.E., Pollack, S.E., Pollard, J.W. (2004) Genetic analysis of chromosome pairing, recombination, and cell cycle control during first meiotic prophase in mammals. *Endocrine Reviews* **27**, 398–426.

Cohen, P.E., Pollard, J.W. (2001) Regulation of meiotic recombination and prophase I progression in mammals. *BioEssays* **23**, 996-1009.

Crismani, W., Baumann, U., Sutton, T., Shirley, N., Webster, T., Spangenberg, G., Langridge, P., Able, J.A. (2006) Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics* **7**, 267.

Deng, X., Habel, J.E., Kabaleeswaran, V., Snell, E.H., Wold, M.S., Borgstahl, G.E. (2007) Structure of the full-length human RPA14/32 complex gives insights into the mechanism of DNA binding and complex formation. *Journal of Molecular Biology* **374**, 865-876.

Dernburg, A.F., McDonald, K., Moulder, G.B., R., Dresser, M., Villeneuve, A.M. (1998) Meiotic recombination in *C. elegans* initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. *Cell* **94**, 387-398.

Ding, Z.-j., Wang, T., Chong, K., Bai, S. (2001) Isolation and characterization of *OsDMC1*, the rice homologue of the yeast *DMC1* gene essential for meiosis. *Sex Plant Reproduction* **13**, 285-288.

Dirlam-Schatz, K.A., Attie, A.D. (1998) Calcium induces a conformational change in the ligand binding domain of the low density lipoprotein receptor. *Journal of Lipid Research* **39**, 402-411.

Doutriaux, M.-P., Couteau, F., Bergounioux, C., White, C. (1998) Isolation and characterisation of the *RAD51* and *DMC1* homologs from *Arabidopsis thaliana*. *Molecular and General Genetics* **257**, 283-291.

Dresser, M.E., Ewing, D.J., Conrad, M.N., Dominguez, A.M., Barstead, R., Jiaugt, H., Kodadekt, T. (1997) DMCl functions in a *Saccharomyces cervsiae* meiotic pathway that is largely independent of the RAD51 pathway. *Genetics* **147**, 533-544.

Driscoll, C.J. (1972) Genetic suppression of homoeologous chromosome pairing in hexaploid wheat. *Candian Journal of Genetic Cytology* **14**, 39-42.

Driscoll, C.J. (1973) Minor genes affecting homoeologous pairing in hybrids between wheat and related genera. *Genetics* **74**, s566.

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.

Durrant, W.E., Wang, S., Dong, X.N. (2007) Arabidopsis SNI1 and RAD51D regulate both gene transcription and DNA recombination during the defense response. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4223-4227.

Eberharter, A., Becker, P.B. (2002) Histone acetylation: a switch between repressive and permissive chromatin. *EMBO Journal* **3**, 224-229.

Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14863-14868.

Enomoto, R., Kinebuchi, T., Sato, M., Yagi, H., Kurumizaka, H., Yokoyama, S. (2006) Stimulation of DNA strand exchange by the human TBPIP/Hop2-Mnd1 complex. *Journal of Biological Chemistry* **281**, 5575-5581.

Esposito, R.E., Frink, N., Bernstein, P., Esposito, M.S. (1972) The genetic control of sporulation in Saccharomyces. *Molecular and General Genetics* **114**, 241-248.

Gerton, J.L., DeRisi, J., Shroff, R., Lichten, M., Brown, P.O., Petes, T.D. (2000) Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proceeding of the National Academy of Sciences of the United States of America* **97**, 11383-11390.

Grelon, M., Vezon, D., Gendrot, G., Pelletier, G. (2001) *AtSPO11-1* is necessary for efficient meiotic recombination in plants. *EMBO Journal* **20**, 589-600.

Griffiths, S., Sharp, R., Foote, T., bertin, I., Wanous, M., Reader, S., Colas, I., Moore, G. (2007) Molecular charaterization of Ph1 as a major chromosomes pairing locus in ploypoid wheat. *Nature* **439**, 749-752.

Guan, H., Kiss-Toth, E. (2008) Advanced technologies for studies on protein interactomes. *Advances in Biochemical Engineering / Biotechnology* **110**, 1-24.

Gupta, R., Ting, J.T.L., Sokolov, L.N., Johnson, S.A., Luan, S. (2002) A tumor suppressor homolog, AtPTEN1, is essential for pollen development in Arabidopsis. *Plant Cell* **14**, 2495-2507.

Haas, B.J., Wortman, J.R., Ronning, C.M., Hannick, L.I., Smith Jr., R.K., Maiti, R., Chan, A.P., Yu, C., Farzad, M., Wu, D., White, O., Town, C.D. (2005) Complete reannotation of the Arabidopsis genome: methods, tools, protocols and the final release. *BMC Biology* **3**, 7.

Habu, T., Wakabayashi, N., Yoshida, K., Yomogida, K., Nishimune, Y., Morita, T. (2004) p53 protein interacts specifically with the meiosis-specific mammalian RecA-like protein DMC1 in meiosis. *Carcinogenesis* **25**, 889-893.

Hamant, O., Ma, H., Cande, W.Z. (2006) Genetics of meiotic prophase I in plants. *Annual Review of Plant Biology* **57**, 267-302.

Hanai, T., Hamada, H., Okamoto, M. (2006) Application of bioinformatics for DNA microarray data to bioscience, bioengineering and medical fields. *Journal of Bioscience and Bioengineering* **101**, 377-384.

Harlow, E., Lane, D. (Eds) (1988) 'Antibodies: a laboratory manual'. Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.

Haruta, N., Kurokawa, Y., Murayama, Y., Akamatsu, Y., Unzai, S., Tsutsui, Y., Iwasaki, H. (2006) The Swi5–Sfr1 complex stimulates Rhp51/Rad51- and Dmc1-mediated DNA strand exchange *in vitro*. *Nature Structural and Molecular Biology* **13**, 823-830.

Henricksen, L.A., Umbricht, C.B., Wold, M.S. (1994) Recombinant replication protein A expression: complex formation, and functional characterization. *Journal of Biological Chemistry* **269**, 11121-11132.

Henry, J.M., Camahort, R., Rice, D.A., Florens, L., Swanson, S.K., Washburn, M.P., Gerton, J.L. (2006) Mnd1/Hop2 facilitates Dmc1-dependent interhomolog crossover formation in meiosis of budding yeast. *Molecular and Cellular Biology* **26**, 2913-2923.

Higgins, J.D., Armstrong S.J., Franklin, F.C.H, Jones, G.H. (2004) The *Arabidopsis MutS* homolog *AtMSH4* functions at an early step in recombination: evidence for two classes of recombination in *Arabidopsis*. *Genes and Development* **18**, 2557-2570.

Higgins, J.D., Sanchez-Moran, E., Armstrong, S.J., Jones, G.H., Franklin, F.C.H. (2005) The Arabidopsis synaptonemal complex protein is required for chromosome synapsis and normal fidelity of crossing-over. *Genes and Development* **19**, 2488-25000.

Hikiba, J., Hirota, K., Kagawa, W., Ikawa, S., Kinebuchi., T., Sakane, I., Takizawa, Y., Yokoyama, S., Mandon-Pepin, B., Nicolase, A., Shibata, T., Ohta, K., Kurumizaka, H. (2008) Structural and functional analyses of the DMC1-M200V polymorphism found in the human population. *Nucleic Acids Research*, 1-10.

Hollingsworth, N., Goetsch, L., Byers, B. (1990) The HOP1 gene encodes a meiosis-specific component of yeast chromosomes. *Cell* **61**, 73-84.

Hollingsworth, N.M., Byers, B. (1989) HOP1: a yeast meiotic pairing gene. *Genetics* **121**, 445-462.

Holm, D., Fink, D.R., Grønlund, J., Hansen, S., Holmskov, U. (2009) Cloning and characterization of SCART1, a novel scavenger receptor cysteine-rich type I transmembrane molecule. *Molecular Immunology* **46** 1663-1672.

Hong, E.L., Shinohara, A., Bishop, D.K. (2001) Saccharomyces cerevisiae Dmc1 protein promotes renaturation of single-strand DNA (ssDNA) and assimilation of ssDNA into homologous super-coiled duplex DNA. *Journal of Biological Chemistry* **276**, 41906-41912.

Hu, G., Zhang, S., Vidal, M., Baer, J.L., Xu, T., Fearon, E.R. (1997) Mammalian homologs of seven in absentia regulate DCC via the ubiquitin–proteasome pathway. *Genes and Development* **11**, 2701-2714.

Hurst, L.D., Pál, C., Lercher, M.J. (2004) The evolutionary dynamics of eukaryotic gene order. *Nature Reviews* **5**, 299-310.

Hussain, S., Wilson, J.B., Medhurst, A.L., Hejna, J., Witt, E., Ananth, S., Moses, R., West, S.C., de Winter, J.P., Ashworth, A., Jones, N.J., Mathew, C.G. (2004) Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Human Molecular Genetics* **13**, 1241–1248.

Ishibashi, T., Kimura, S., Sakaguchi, K. (2006) A higher plant has three different types of RPA heterotrimeric complex. *Journal of Biochemistry* **139**, 99-104.

Iwabata, K., Koshiyama, A.*, et al.* (2005) DNA topoisomerase II interacts with Lim15/Dmc1 in meiosis. *Nucleic Acids Research* **33**, 5809-5818.

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V., Kim, S.K. (2001) Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhbditis elegans*. *Proceeding of the National Academy of Sciences of the United States of America* **98**, 218-223.

Jiang, Z.G., McKnight, C.J. (2006) A phosphorylation-induced conformation change in dematin headpiece. *Structure* **14**, 379-387.

Jung, K.-H., Hur, J., Ryu, C.-H., Choi, Y., Chung, Y.-Y., Miyao, A., Hirochika, H., An, G. (2003) Characterization of a rice chlorophyll-deficient mutant using the T-DNA gene-trap system. *Plant Cell Physiology* **44**, 463-472.

Kaback, D.B., Feldberg, L.R. (1985) *Saccharomyces cerevisiae* exhibits a sporulation-specific temporal pattern of transcript accumulation. *Molecular and Cellular Biology* **5**, 751-761.

Kaminski, N., Friedman, N. (2002) Practical approaches to analyzing results of microarray experiments. *American Journal of Respiratory Cell and Molecular Biology* **27**, 125-132.

Kant, C.R., B.J.Rao, J.K.Sainis (2005) DNA binding and pairing activity of *Os*Dmc1, a recombinase from rice. *Plant Molecular Biology* **57**, 1-11.

Kaur, J., Sebastian, J., Siddiqi, I. (2006) The *Arabidopsis-mei2-Like* genes play a role in meiosis and vegetative growth in *Arabidopsis*. *Plant Cell* **18**, 545-559.

Keeny, S. (2001) Mechanism and control of meiotic recombination initiation. *Current Topics in Developmental Biology* **52**, 1-53.

Kerzendorfer, C., Vignard, J., Pedrosa-Harand, A., Siwiec, T., Akimcheva, S., Jolivet, S., Sablowski, R., Armstong, S., Schweizer, D., Mercier, R., Schlögelhofer, P. (2006) The *Arabidopsis thaliana MND1* homologue plays a key role in meiotic homologous pairing, synapsis and recombination. *Journal of Cell Science* **119**, 2486-2496.

Khoo, K.H.P., Jolly, H.R., Able, J.A. (2008) The RAD51 gene family in bread wheat is highly conserved across eukaryotes, with RAD51A upregulated during early meiosis. *Functional Plant Biology* **35**, 1267-1277.

Kinebuchi, T., Kagawa, W., Enomoto, R., Tanaka, K., Miyagawa, K., Shibata, T., Kurumizaka, H., Yokoyama, S. (2004) Structural basis for octameric ring formation and DNA interaction of the human homologous-pairing protein Dmc1. *Molecular Cell* **14**, 363-374.

Krysan, P.J., Young, J.C., Sussman, M.R. (1999) T-DNA as an insertional mutagen in Arabidopsis. *Plant Cell* **11**, 2283–2290.

Kupiec, M., Byers, B., Esposito, R.E., Mitchell, A.P. (1997). In 'The Molecular and cellular biology of the yeast Saccharomyces: Cell cycle and cell biology'. (Eds JR Pringle, JR Broach and EW Jones) pp. 889-1036. Cold Springs Habor Laboratory Press: Cold Spring Habor, NY.

Lalonde, S., Ehrhardt, D.W., Loque, D., Chen, J., Rhee, S.Y., Frommer, W.B. (2008) Molecular and cellular approaches for the detection of protein–protein interactions: latest techniques and current limitations. *Plant Journal* **53**, 610-635.

Lavoie, B.D., Hogan, E., Koshland, D. (2002) In vivo dissection of the chromosome condensation machinery: reversibility of condensation distinguishes contributions of condensin and cohesin. *Journal of Cell Biology* **156**, 805-815.

Lee, J.M., Sonnhammer, E.L.L. (1993) Genomic gene clustering analysis of pathways in eukaryotes. *Genome Research* **13**, 875-882.

Leu, J.-Y., Chua, P.R., Roeder, G.S. (1998) The meiosis-specific Hop2 protein of *S. cerevisiae* ensures synapsis between homologous chromosomes. *Cell* **94**, 375-386.

Li, W., Chen, C., Markmann-Mulisch, U., Timofejeva, L., Schmelzer, E., Ma, H., Reiss, B. (2004) The *Arabidopsis AtRAD51* gene is dispensable for vegetative development but required for meiosis. *Proceeding of the National Academy of Sciences of the United States of America* **101**, 10596-10601.

Li, Z., Golub, E.I., Gupta, R., Radding, C.M. (1997) Recombination activities of HsDmc1 protein, the meiotic human homolog of RecA protein. *Proceeding of the National Academy of Sciences of the United States of America* **94**, 11221-11226.

Libby, B.J., De La Fuente, R., O'Brien, M.J., Wigglesworth, K., Cobb, J., Inselman, A., Eaker, S., Handel, M.A., Eppig, J.J., Schimenti, J.C. (2002) The mouse meiotic mutation *mei1* disrupts chromosome synapsis with sexually dimorphic consequences for meiotic progression. *Developmental Biology* **242**, 174-187.

Lin, Z., Kong, H., Nei, M., Ma, H. (2006) Origins and evolution of the recA/RAD51 gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. *Proceeding of the National Academy of Sciences of the United States of America* **103**, 10328-10333.

Lloyd, A.H., Milligan, A.S., Langridge, P., Able, J.A. (2007) *TaMSH7*: A cereal mismatch repair gene that affects fertility in transgenic barley (*Hordeum vulgare* L.). *BMC Plant Biology* **7**, 67.

Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Eun, A., He, Y., Ma, W., Wang, W., Wen, Z., Peng, J. (2003) 15,000 unique zebrafish EST clusters and their future use in microarray for profiling gene expression patterns during embryogensis. *Genome Research* **13**, 455-466.

Loidl, J., Klein, F., Scherthan, H. (1994) Homologous pairing is reduced but not abolished in asynaptic mutants of yeast. *Journal of Cell Biology* **125**, 1191-1200.

Ma, J., Skibbe, D.S., Fernandes, J.F., Walbot, V. (2008) Male reproductive development: gene expression profiling of mazie anther and pollen ontogeny. *Genome Biology* **9**, R181.

MacQueen, A.J., Colaiacovo, M.P., McDonald, K., Villeneuve, A.M. (2002) Snapsis-dependent and -independent mechanisms stabilizehomolog pairing during prophase in C. elegans. *Genes and Development* **16**, 2428-2442.

Martínez-Pérez, E., Shaw, P., Reader, S., Aragón-Alcaide, L., Miller, T., Moore, G. (1999) Homologous chromosome pairing in wheat. *Journal of Cell Science* **112**, 1761-1769.

Masson, J.-Y., Davies, A.A., Hajibagheri, N., Dyck, E.V., Benson, F.E., Stasiak, A.Z., Stasiak, A., West, S.C. (1999) The meiosis-specific recombinase hDmc1 forms ring structures and interacts with hRad51. *EMBO Journal* **18**, 6552 - 6560.

Masterson, J. (1994) Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* **264**, 421-424.

Mata, J., Bahler, J. (2003) Correlations between gene expression and gene conservation in fission yeast. *Genome Research* **13**, 2686-2890.

Mata, J., Lyne, R., Burns, G., Bähler, J. (2002) The transcriptional program of meiosis and sporulation in fission yeast. *Nature Genetics* **32**, 143-147.

McKim, K.S., Green-Marroquin, B.L., Sekelsky, J.J., Chin, G., Steinberg, C., Khodosh, R., Hawley, R.S. (1998) Meiotic synapsis in the absence of recombination. *Science* **279**, 876-878.

Mercier, R., Vezon, D., Bullier, E., Motamayor, J.C., Sellier, A., Lefevre, F., Pelletier, G., Horlow, C. (2001) Switch1 (Swi1): a novel protein required for the establishment of sister chromatid cohesion and for bivalent formation at meiosis. *Genes and Development* **15**, 1859-1871.

Merker, J.D., Dominska, M., Greewell, P.W., Rinella, E., Bouck, D.C., Shibata, Y., Strahl, B.D., Mieczkowski, P., Petes, T.D. (2008) The histone methylase Set2p and the histone deacetylase Rpd3p repress meiotic recombination at the HIS4 meiotic recombination hotspot in *Saccharomyces cerevisiae*. *DNA Repair* **7**, 1298-1308.

Miernyk, J.A. (2001) The J-domain proteins of Arabidopsis thaliana: An unexpectedly large and diverse family of chaperones. *Cell Stress Chaperones* **6**, 209-218.

Mikhailova, E., Phillips, D., Sosnikhina, S., Lovtsyus, A., Jones, R., Jenkins, G. (2006) Molecular assembly of meiotic proteins Asy1 and Zyp1 and pairing promiscuity in rye (*Secale cereale* L.) and its synaptic mutant *sy10*. *Genetics* **10**, 1247-1258.

Moens, P.B., Kolas, N.K., Tarsounas, M., Marcon, E., Cohen, P.E., Spyropoulos, B. (2002) The time course and chromosomal localization of recombination-related proteins at meiosis in the mouse are compatible with models that can resolve the early DNA-DNA interactions without reciprocal recombination. *Journal of Cell Science* **115**, 1611-1622.

Moore, D.P., Orr-Weaver, T.L. (1998) Chromosome segregation during meiosis: building an unambivalent bivalent. *Current Topics in Developmental Biology* **37**, 263-299.

Moore, G. (2000) Cereal chromosome structure, evolution and pairing. *Annual Review of Plant Physiology and Plant Molecular Biology* **51**, 195-222.

Moore, G. (2002) Meiosis in allopoylploids - the importance of „Teflon" chromosomes. *Trends in Genetics* **18**, 456-463.

Motamayor, J.C., Vezon, D., Bajon, C., Sauvanet, A., Grandjean, O., Marchand, M., Bechtold, N., Pelletier, G., Horlow, C. (2000) Switch (swi1), an *Arabidopsis thaliana* mutant affected in the female meiotic switch. *Sex Plant Reproduction* **12**, 209-218.

Mukherjee, S., Bal, S., Saha, P. (2001) Protein-interaction maps using yeast two-hybrid assay. *Current Science* **81**, 458-464.

Muyt, A.D., Mercier, R., Mézard, C., Grelon, M. (2009) Meiotic recombination and crossovers in plants. *Genome Dynamics* **5**, 14-25.

Muyt, A.D., Vezon, D., Gendrot, G., Gallois, J.-L., Stevens, R., Grelon, M. (2007) *At*PRD1 is required for meiotic double strand break formation in *Arabidopsis thaliana*. *EMBO Journal* **26**, 4126-4137.

Nabeshima, K., Kakihara, Y., Hiraoka, Y., Nojima, H. (2001) A novel meiosis-specific protein of fission yeast, Meu13p, promotes homologous pairing independently of homologous recombination. *EMBO Journal* **20**, 3871-3881.

Nag, D.K., Pata, J.D., Sironi, M., Flood, D.R., Hart, A.M. (2006) Both conserved and non-conserved regions of Spo11 are essential for meiotic recombination initiation in yeast. *Molecular Genetics and Genomics* **276**, 313-321.

Nara, T., Saka, T., Sawado, T., Hotta, H.T.Y.I.Y., Sakaguchi, K. (1999) Isolation of a *LIM15/DMC1* homolog from the basidiomycete *Coprinus cinereus* and its expression in relation to meiotic chromosome pairing. *Molecular and General Genetics* **262**, 781-789.

Neale, M.J., Keeney, S. (2006) Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature* **442**, 153-158.

Neyton, S., Lespinasse, F., Moens, P.B., Paul, R., Gaudray, P., Paquis-Flucklinger, V., Santucci-Darmanin, S. (2004) Association between MSH4 (MutS homologue 4) and the DNA strand-exchange RAD51 and DMC1 proteins during mammalian meiosis. *Molecular Human Reproduction* **10**, 917-924.

Niu, H., Wan, L., Baumgartner, B., Schaefer, D., Loidl, J., Hollingsworth, N.M. (2005) Partner choice during meiosis is regulated by Hop1-promoted dimerization of Mek1. *Molecular Biology of the Cell* **16**, 5804-5818.

Nonomura, K.I., Nakano, M., Murata, K., Miyoshi, K., Eiguchi, M., Miyao, A., Hirochika, H., Kurata, N. (2004) An insertational mutation of rice PAIR2, the ortholog of Arabidopsis ASY1, results in a defect in homologous chromosome pairing during meiosis. *Molecular Genetics and Genomics* **271**, 121-129.

Ono, T., Losada, A., Hirano, M., Myers, M.P., Neuwald, A.F., Hirano, T. (2003) Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell* **115**, 109-121.

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

Pandey, R., Müller, A., Napoli, C.A., Selinger, D.A., Pikaard, C.S., Richards, E.J., Bender, J., Mount, D.W., Jorgensen, R.A. (2002) Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Research* **30**, 5036-5055.

Pang, A.L.Y., Johnson, W., Ravindranath, N., Dym, M., Rennert, O.M., Chan, W.-Y. (2006) Expression profiling of purified male germ cells: stage-specific expression patterns related to meiosis and postmeiotic development *Physiological Genomics* **24**, 75-85.

Panoli, A.P., Ravi, M., Sebastian, J., Nishal, B., Reddy, T.V., Marimuthu, M.P., Subbiah, V., Vijaybhaskar, V., Siddiqi, I. (2006) *AtMND1* is required for homologous pairing during meiosis in *Arabidopsis*. *BMC Molecular Biology* **7**, 24.

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., Lagudah, E., Somers, D., Kilian, A., Alaux, M., Vautrin, S., Bergès, H., Eversole, K., Appels, R., Safar, J., Simkova, H., Dolezel, J., Bernard, M., Feuillet, C. (2008) A physical map of the 1-Gigabase bread wheat chromosome 3B. *Science* **322**, 101-104.

Pawlowski, W.P., Cande, W.Z. (2005) Coordinating the events of the meiotic prophase. *Trends in Cell Biology* **15**, 674-681.

Pawlowski, W.P., Inna N. Golubovskaya, Timofejeva, L., Meeley, R.B., Sheridan, W.F., Cande, W.Z. (2004) Coordination of meiotic recombination, pairing, and synapsis by PHS1. *Science* **303**, 89-92.

Perrella, G., Consiglio, M.F., Aises-Cigliano, R., Cremona, G., Sanchez-Moran, E., Barra, L., Errico, A., Bressan, R.A., Franklin, C.H., Conicella, C. (2010) Histone hyperacetylation affects meiotic recombination and chromosome segregation in Arabidopsis. *The Plant Journal* **62**, 796-806.

Petersen, G., Seberg, O. (2002) Molecular Evolution and Phylogenetic Application of *DMC1*. *Molecular Phylogenetics and Evolution* **22**, 43-50.

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

Petukhova, G.V., Pezza, R.J., Vanevski, F., Ploquin, M., Masson, J.-Y., Camerini-Otero, R.D. (2005) The Hop2 and Mnd1 proteins act in concert with Rad51 and Dmc1 in meiotic recombination. *Nature Structural and Molecular Biology* **12**, 449-453.

Pezza, R.J., Petukhova, G.V., Ghirlando, R., Camerini-Otero, R.D. (2006) Molecular Activities of Meiosis-specific Proteins Hop2, Mnd1, and the Hop2-Mnd1 Complex. *Journal of Biological Chemistry* **281**, 18426-18434.

Primig, M., Williams, R.M., Winzeler, E.A., Tevzadze, G.G., Conway, A.R., Hwang, S.Y., Davis, R.W., Esposito, R.E. (2000) The core meiotic transcriptome in budding yeast. *Nature Genetics* **26**, 415-423.

Rabitsch, K.P., Toth, A., Galova, M., Schleiffer, A., Schaffner, G., Aigner, E., Rupp, C., Penkner, A.M., Moreno-Borchart, A.C., Primig, M., Esposito, R.E., Klein, F., Knop , M., Nasmyth, K (2001) A screen for genes required for meiosis and spore formation based on whole-genome expression. *Current Biology* **11**, 1001–1009.

Reinke, V., Smith, H.E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S.J.M., Davis, E.B., Scherer, S., Ward, S., Kim, S.K. (2000) A global profile of germline gene expression in *C. elegans*. *Molecular Cell* **6**, 605-616.

Riley, R., Chapman, V. (1958) Genetic control of the cytologically diploid behaviour of hexaploid wheat. *Nature* **182**, 713-715.

Rockmill, B., Roeder, G.S. (1990) Meiosis in asynaptic yeast. *Genetics* **126**, 563-574.

Ronceret, A., Doutriaux, M.-P., Golubovskaya, I.N., Pawlowski, W.P. (2009) PHS1 regulates meiotic recombination and homologous chromosome pairing by controlling the transport of RAD50 to the nucleus. *Proceeding of the National Academy of Sciences of the United States of America* **106**, 20121-20126.

Ross, K.J., Fransz, P., Armstrong, S.J., Vizir, I., Mulligan, B., Franklin, F.C.H., Jones, G.H. (1997) Cytological characterization of four meiotic mutants of *Arabidopsis* isolated from T-DNA-transformed lines. *Chromosome Research* **5**, 551-559.

Sakaguchi, K., Ishibashi, T., Uchiyama, Y., Iwabata, K. (2009) The multi-replication protein A (RPA) system - a new perspective. *FEBS Journal* **276**, 943-963.

Sanchez-Moran, E., Santos, J.-L., Jones, G.H., Franklin, F.C.H. (2007) ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in *Arabidopsis*. *Genes and Development* **21**, 2220-2233.

Sanders, P.M., Bui, A.Q., Weterings, K., McIntire, K.N., Hsu, Y.-C., Lee, P.Y., Truong, M.T., Beals, T.P., Goldberg, R.B. (1999) Anther developmental defects in Arabidopsis thaliana male-sterile mutants. *Sexual Plant Reproduction* **11**, 297-322.

Sarai, N., Kagawa, W., Kinebuchi, T., Kagawa, A., Tanaka, K., Miyagawa, K., Ikawa, S., Shibata, T., Kurumizaka, H., Yokoyama, S. (2006) Stimulation of Dmc1-mediated DNA strand exchange by the human Rad54B protein. *Nucleic Acids Research* **34**, 4429-4437.

Sarkar, P.K., Florczyk, M.A., McDonough, K.A., Nag, D.K. (2002) *SSP2*, a sporulation-specific gene necessary for outer spore wall assembly in the yeast *Saccharomyces cerevisiae*. *Molecular Genetics and Genomics* **267**, 348-358.

Schena, M. (2003) Microarray analysis. pp. 13-23, Wiley-Liss.

Schlecht, U., Demougin, P., Koch, R., Hermida, L., Wiederkehr, C., Descombes, P., Pineau, C., Jegou, B., Primig, M. (2004) Expression profiling of mammalian male meiosis and gametogenesis identifies novel candidate genes for roles in the regulation of fertility. *Molecular Biology of the Cell* **15**, 1031-1043.

Schmiesing, J.A., Gregson, H.C., Zhou, S., Yokomori, K. (2000) A human condensin complex containing hCAP-C-hCAP-E and CNAP1, a homolog of *Xenopus* XCAP-D2, colocalizes with phosphorylated Histone H3 during the early stage of mitotic chromosome condensation *Molecular and Cellular Biology* **20**, 6996-7006.

Schommer, C., Beven, A., Lawrenson, T., Shaw, P., Sablowski, R. (2003) *AHP2* is required for bivalent formation and for segregation of homologous chromosomes in *Arabidopsis* meiosis. *Plant Journal* **36**, 1-11.

Schwacha, A., Kleckner, N. (1997) Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway. *Cell* **90**, 1123-1135.

Schwarzacher, T. (1997) Three stages of meiotic homologous chromosome pairing in wheat: cognition, alignment and synapsis. *Sex Plant Reproduction* **10**, 324-331.

Sears, E.R. (1952) Homoeologous chromosomes in *Triticum aestivium*. *Genetics* **37**, 624.

Sears, E.R. (1976) Genetic control of chromosome pairing in wheat. *Annual Review of Genetics* **10**, 31-51.

Sears, E.R. (1977) An induced mutant with homoeologous pairing in common wheat. *Canadian Journal of Genetics and Cytology* **19**, 585-593.

Sears, E.R. (1982) A wheat mutation conditioning an intermediate level of homoeologous chromosome pairing. *Canadian Journal of Genetics and Cytology* **24**, 715-719.

Sehorn, M.G., Sigurdsson, S., Bussen, W., Unger, V.M., Sung, P. (2004) Human meiotic recombinase Dmc1 promotes ATP-dependent homologous DNA strand exchange. *Nature* **429**, 433-437.

Shinohara, A., Shinohara, M. (2004) Roles of *RecA* homologues *Rad51* and *Dmc1* during meiotic recombination. *Cytogentics and Genomes Research* **107**, 201-207.

Siaud, N., Dray, E., Gy, I., Gerard, E., Takvorian, N., Doutriaux, M.-P. (2004) Brca2 is involved in meiosis in *Arabidopsis thaliana* as suggested by its interaction with Dmc1. *EMBO Journal* **23**, 1392-1401.

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

Sorrells, M.E., Rota, M.L., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Miftahudin, Mahmoud, A., Ma, X., Gustafson, P.J., Qi, L.L., Echalier, B., Gill, B.S., Matthews, D.E., Lazo, G.R., Chao, S., Anderson, O.D., Edwards, H.,

Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Zhang. D., Nguyen, H.T., Peng, J., Lapitan, N.L.V., Gonzalez- Hernandez, J.L., Anderson, J.A., Hossain, K., Kalavacharla, V., Kianian, S.F., Choi, D., Close, T.J., Dilbirligi, M., Gill, K.S., Steber, C., Walker-Simmons, M.K., McGuire, P.E., Qualset, C.O. (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Research* **13**, 1818-1827

Stiegler, A.L., Burden, S.J., Hubbard, S.R. (2009) Crystal structure of the frizzled-like cysteine-rich domain of the receptor tyrosine kinase MuSK. *Journal of Molecular Biology* **393**, 1-9.

Sutton, T., Whitford, R., Baumann, U., Dong, C.M., Able, J.A., Langridge, P. (2003) The *Ph2* pairing homoeologous locus of wheat (*Triticum aestivum*): identification of candidate meiotic genes using a comparative genetics approach. *Plant Journal* **36**, 443-456.

Tamura, K., Dudley, J., Nei, M., Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology of Evolution* **24**, 1596-1599.

Tanaka, T., Nakamura, T., Takagi, H., Sato, M. (1997) Molecular cloning and characterisation of a novel TBP-1 interacting protein (TBPIP): enhancement of TBP-1 action on Tat by TBPIP. *Biochemical and Biophysical Research Communications* **239**, 176-181.

Tanaka, T.S., Kunath, T., Kimber, W.L., Jaradat, S.A., Stagg, C.A., Usuda, M., Yokota, T., Niwa, H., Rossant, J., Ko, M.S.H. (2002) Gene expression profiling of embryo-derived stem cells reveal candidate genes associated with pluripotency and lineage specificity. *Genome Research* **12**, 1921-1928.

Tarsounas, M., Morita, T., Pearlman, R.E., Moens, P.B. (1999) RAD51 and DMC1 form mixed complexes associated with mouse meiotic chromosome cores and synaptonemal complexes. *Journal of Cell Biology* **147**, 207-219.

Teichmann, S.A., Veitia, R.A. (2004) Genes encoding subunits of stable complexes are clustered on the yeast chromosomes: an interpretation from a dosage balance perspective. *Genetics* **167**, 2121-2125.

Terasawa, M., Shinohara, A., Hotta, Y., Ogawa, H., Ogawa, T. (1995) Localization of RecA-like recombination proteins on chromosomes of the lily at various meiotic stages. *Genes and Development* **9**, 925-934.

Thompson, D.A., Stahl, F.W. (1999) Genetic control of recombination partner preference in yeast meiosis: Isolation and characterization of mutants elevated for meiotic unequal sister-chromatid recombination. *Genetics* **153**, 621-641.

Tsubouchi, H., Roeder, G.S. (2002) The Mnd1 protein forms a complex with Hop2 to promote homologous chromosome pairing and meiotic double-strand break repair. *Molecular and Cellular Biology* **22**, 3078–3088.

Udall, J.A., Wendel, J.F. (2006) Ploypoidy and crop improvement. *Crop Science* **46**, S3-14.

Ushizawa, K., Herath, C.B., Kaneyama, K., Shiojima, S., Hirasawa, A., Takahashi, T., Imai, K., Ochiai, K., Tokunaga, T., Tsunoda, Y., Tsujimoto, G., Hashizume, K. (2004) cDNA microarray analysis of bovine embryo gene expression profiles during the pre-implantation period. *Reproductive Biology and Endocrinology* **2**, 77.

Vignard, J., Siwiec, T., Chelysheva, L., Vrielynck, N., Gonord, F., Armstrong, S.J., Schlogelhofer, P., Mercier, R. (2007) The interplay of RecA-related proteins and the MND1-HOP2 complex during meiosis in *Arabidopsis thaliana*. *PLoS Genetics* **3**, e176.

Vlachonasios, K.E., Thomashow, M.F., Triezenberg, S.J. (2003) Distruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect *Arabidopsis* growth, development, and gene expression. *Plant Cell* **15**, 626-638.

Walhout, A.J.M., Vidal, M. (1999) A genetic strategy to eliminate self-activator baits prior to high-throughput yeast two-hybrid screens. *Genome Research* **9**, 1128-1134.

Wang, Y.H. (2008) How effective is T-DNA insertional mutagenesis in Arabidopsis? *Journal of Biochemistry Techniques* **1**, 11-20.

Wang, Z., Liang, Y., Li, C., Xu, Y., Lan, L., Zhao, D., Chen, C., Xu, Z., Xue, Y., Chong, K. (2005) Microarray analysis of gene expression involved in anther development in rice (*Oryza sativa* L.). *Plant Molecular Biology* **58**, 721-737.

Wijeratne, A.J., Chen, C., Zhang, W., Timofejeva, L., Ma, H. (2006) The *Arabidopsis thaliana* PARTING DANCERS gene encoding a novel protein is required for normal meiotic homologous recombination. *Molecular Biology of the Cell* **17**, 1331-1343.

Wilson, P.J., Riggs, C.D., Hasenkampf, C.A. (2005) Plant chromosome homology: hypotheses relating rendezvous, recognition and reciprocal exchange. *Cytogentics and Genomes Research* **109**, 190-197. S. Karger AG, Basel

Wold, M.S. (1997) REPLICATION PROTEIN A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annual Review of Biochemistry* **66**, 61-92.

Wu, S.-Y., Culligan, K., Lamers, M., Hays, J. (2003) Dissimilar mispair-recognition spectra of Arabidopsis DNA-mismatch-repair proteins MSH2·MSH6 (MutSα) and MSH2·MSH7 (MutSγ). *Nucleic Acids Research* **31**, 6027-6034.

Xu, L., Ajimura, M., Padmore, R., Klein, C., Kleckner, N. (1995) NDT80, a meiosis-specific gene required for exit from pachytene in *Saccharomyces cerevisiae*. *Molecular and Cellular biology* **15**, 6572-6581.

Yamada, T., Mizuno, K.I., Hirota, K., Kon, N., Wahls, W.P., Hartsuiker, E., Murofushi, H., Shibata, T., Ohta, K. (2004) Roles of histone acetylation and chromatin remodelling factor in a meiotic recombination hotspot. *EMBO Journal* **23**, 1792-1803.

Yin, Y., Cheong, H., Friedrichsen, D., Zhao, Y., Hu, J., Mora-Garcia, S., Chory, J. (2002) A crucial role for the putative Arabidopsis topoisomerase VI in plant growth and development. *Proceeding of the National Academy of Sciences of the United States of America* **99**, 10191-10196.

Zarembinski, T.I., Hung, L.-W., Mueller-Dieckmann, H.-J., Kim, K.-K., Yokota, H., Kim, R., Kim, S.-H. (1998) Structure-based assignment of the biochemical function of a

hypothetical protein: A test case of structural genomics. *Proceeding of the National Academy of Sciences of the United States of America* **95**, 15189-15193.

Zickler, D., Kleckner, N. (1998) The leptotene-zygotene transition of meiosis. *Annual Review of Genetics* **32**, 619-97.

Zickler, D., Kleckner, N. (1999) Meiotic chromosomes: integrating structure and function. *Annual Review of Genetics* **33**, 603-754.

Appendix

Appendix 2.1 – Expressed sequence tags (ESTs) corresponding to candidates

EST accession numbers corresponding to Affymetrix Transcript IDs. EST sequences were used to search the public database NCBI (BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI-GenBank Flat File Release 160.0) for annotating candidates.

Novel	Transcript ID	Corresponding EST Accession
1	Ta.1226.1.S1_at	BE420387
2	TaAffx.11970.1.S1_s_at	BG606837
3	Ta.2952.1.S1_at	BJ208107
4	Ta.9089.1.A1_at	BJ304059
5	Ta.24603.1.A1_at	BQ170856
6	Ta.25984.1.S1_at	CD454262
7	Ta.3397.1.S1_at	BJ211174
8	Ta.10020.1.S1_at	BJ236889
9	Ta.10024.1.S1_s_at	BJ236932
10	TaAffx.11970.2.S1_at	BJ309260
11	Ta.8260.1.A1_at	BQ161665
12	Ta.13989.1.S1_at	BQ788879
13	TaAffx.16307.1.S1_at	BQ800749
14	Ta.14043.1.S1_at	BQ805612
15	TaAffx.87249.1.S1_at	CA593367
16	TaAffx.58843.1.S1_at	CA614416
17	Ta.18725.1.S1_at	CA641260
18	TaAffx.25097.1.S1_at	CA706541
19	TaAffx.24409.1.S1_x_at	CA719845
20	TaAffx.20193.1.A1_at	CD905557
21	TaAffx.60258.2.S1_s_at	CA729180
22	TaAffx.38287.1.S1_at	CD907277
23	Ta.12761.1.A1_at	CK205455
24	TaAffx.84212.1.S1_at	CA642301
25	Ta.25933.1.S1_at	CD454563
26	Ta.6748.2.A1_a_at	BJ276809
27	Ta.16123.1.A1_at	CK169168
28	Ta.6748.1.A1_at	BJ306952
29	TaAffx.44789.1.S1_at	CA652796
30	Ta.16224.1.S1_at	CA602888
31	Ta.18287.1.S1_at	CA633298
32	Ta.9286.1.S1_at	CD454567
33	TaAffx.65731.1.A1_at	BJ230292
34	Ta.1673.1.S1_at	BE446763
35	Ta.11783.1.A1_at	BQ170792
36	Ta.3093.2.A1_at	CA499821
37	Ta.25584.1.S1_at	CD454169
38	TaAffx.124106.1.S1_at	CA600406
39	Ta.9807.1.A1_at	BJ237913

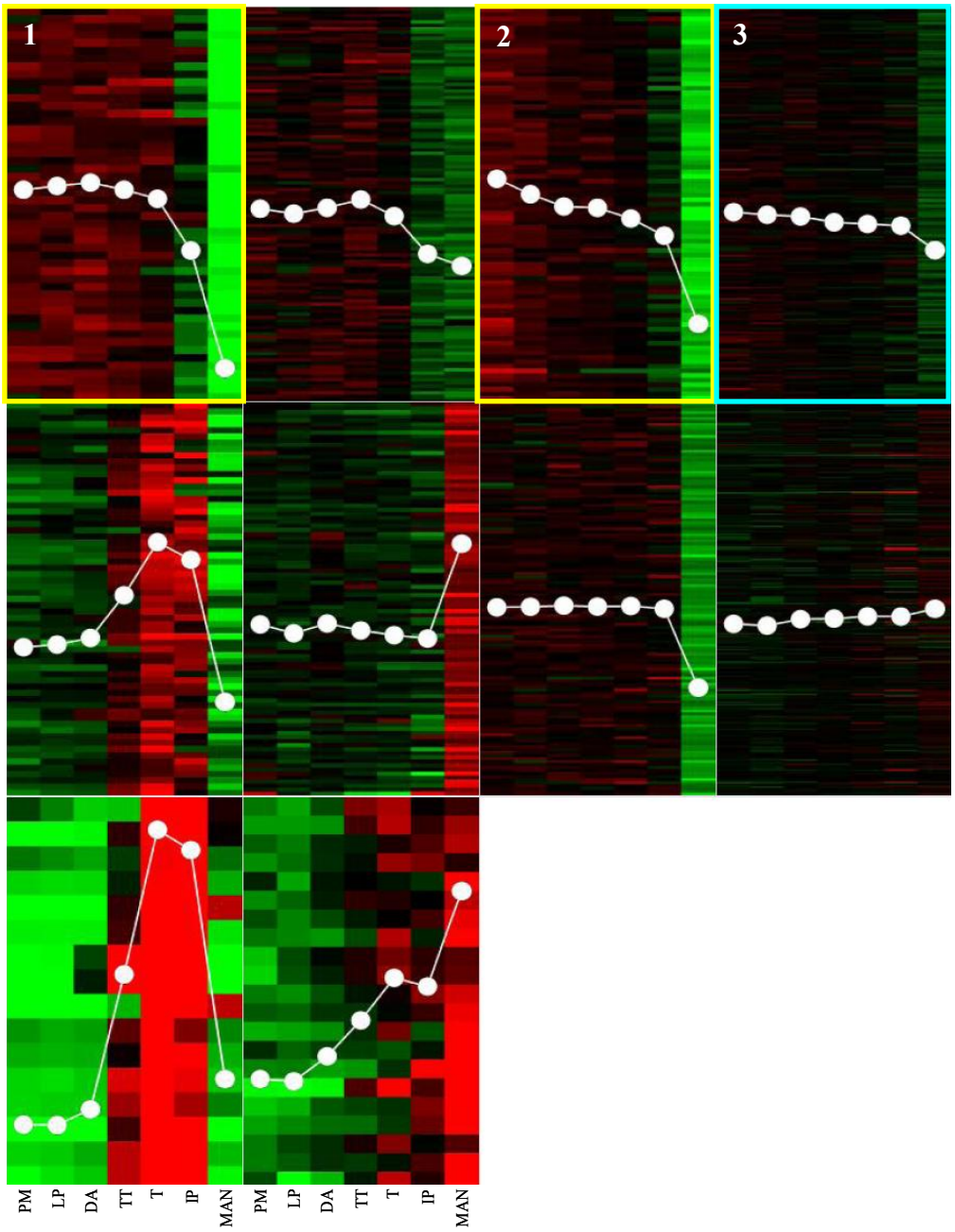
Appendix 2.2 – Oligonucleotides for amplification of candidates

Novel	Transcript ID	Sense	Antisense	Fragment Size (bp)
1	Ta.1226.1.S1_at	AGCCAGGCAGCTGTTGCTCG	TTGTTGTCTTGCACGGCCAAAGTA	334
2	TaAffx.11970.1.S1_s_at	GGCCATGGATGTGGGCTTCC	AGCTTGAAGCATCCGGTCAGCTC	271
3	Ta.2952.1.S1_at	GCTCAGGAAGAAGGCTGCTGCTT	CACACCTCAATCGCTAAACCCACC	262
4	Ta.9089.1.A1_at	GCCGCTGAGGAAGAAGGTTTCTG	ACACAGCAGACTGCATCTCCGATATC	262
5	Ta.24603.1.A1_at	CGGCGAAGACATCACCGAAGAG	GCAGGTTCGGATCTCAGTAATCGC	355
6	Ta.25984.1.S1_at	AAGGCTTCTGGTGCTCCGGC	AACACACAGCATACTGCATCTCCGAG	246
7	Ta.3397.1.S1_at	CTTCTCCTGAGAAGAAGGTGCGGA	GCACCAACACATCAGATCAGCTAACC	379
8	Ta.10020.1.S1_at	GACGAGGCAGGTGCTCCCGT	GCCATGCCCCGCAGGTG	274
9	Ta.10024.1.S1_s_at	GGGAAGGAAGGCAACGTGATCC	TCCGTTGCTAGTGGGAAGGATCAC	336
10	Ta.Affx.11970.2.S1_at	CTATACCATTGCCACAACCCCCAAT	AGCCCAAGGCTCCTCCTGCA	187
11	Ta.8260.1.A1_at	CAGCTCACCGCCCAGGAGG	CATTCAACTGAGATAACAAGCCGCC	323
12	Ta.13989.1.S1_at	CCCAGGGCATGGATCCGG	GATACGCACACACGTGCACCGT	286
13	TaAffx.16307.1.S1_at	CGGACGAGGCCGACCAGTC	CACCAAGGACGGATACAGAGCTGAT	335
14	Ta.14043.1.S1_at	GTGAATTGCCCTACCAATTATCCGTG	TGTGGTAGCTCATCATTCATGCCG	203
15	TaAffx.87249.1.S1_at	CCCAGTGACGAAGACATGCTCTTCT	ATGTGAGCCCAAACTGAGTCCCAA	326
16	TaAffx.58843.1.S1_at	GCATTTGGTGCCTATGATATCCCG	CCCGAGCATTGTCGATGGGT	373
17	Ta18725.1.S1_at	CGTGCTGGTGTCACCTGACGAG	CTATCTGAGGCGCTGAAGCCTCTG	207
18	TaAffx.25097.1.S1_at	TGATGGACAGACCTCTCCTTCTGCAT	TTAGAAGCCACTTCACATGGACTCTAACC	268
19	TaAffx.24409.1.S1_x_at	AGCCTACGGAGGCTCAAACTGAAG	GCCGAATTCGGCTCTTTGTGC	263
20	TaAffx.20193.1.A1_at	AGACGTGCCGTGAGGTCCTGAA	TTAGAGGTCTACAGCACCGAGTGAAGC	410
21	TaAffx.60258.2.S1_s_at	AGGATCGCCCCAGCCAAGTC	CAAGTGAGCGAGCCGCCG	399
22	TaAffx.38287.1.S1_at	CCAGCCTGAAGAGCATATGACTCCTG	GCTTCTGCGCCAACAGGGTC	286
23	Ta.12761.1.A1_at	AAGGCTGCAGTGGCACAAGACC	AATACAGCCAAATCTTCTCACGCTCTG	463
24	TaAffx.84212.1.S1_at	GTAACATCGGTGAGCCAGGTCCC	GTGGCAGTGTCACCCTGCATACC	300
25	Ta.25933.1.S1_at	GTGTTTATTGGCCTGCTCTGTGCTG	TCGAGTGGTGAACCTGTGTGCATAC	242
26	Ta.6748.2.A1_a_at	CTTTCAAGACTCGTAGCACCGGGAT	AGCTCCACTCCATTGATCATCAGCA	279
27	Ta.16123.1.A1_at	TTCAACACCAACAAGGCGCCA	ATCTAAGCGCTTCGCGTTTCCAG	145
28	Ta.6748.1.A1_at	CTGCATTTTGTGGTTCATGAATGAATG	TGCATGCAGTCTAATCTCGAGTCGAG	218
29	TaAffx.44789.1.S1_at	GTCTTACCACCTTATCAGGGTGAAGGG	CAGCACCGCAGGAACTGCTTCT	320
30	Ta.16224.1.S1_at	GCTCCACGGGAGGTCAAAACTTG	AAGTACTGGACGGCTGCCTGCA	258
31	Ta.18287.1.S1_at	TTCCCATTGGATACATACACTCCAG	GGATTCTTAGATCATGCATCAGCA	245

Appendix 2.3 – *K*-means clustering of 1094 meiotically expressed novel transcripts

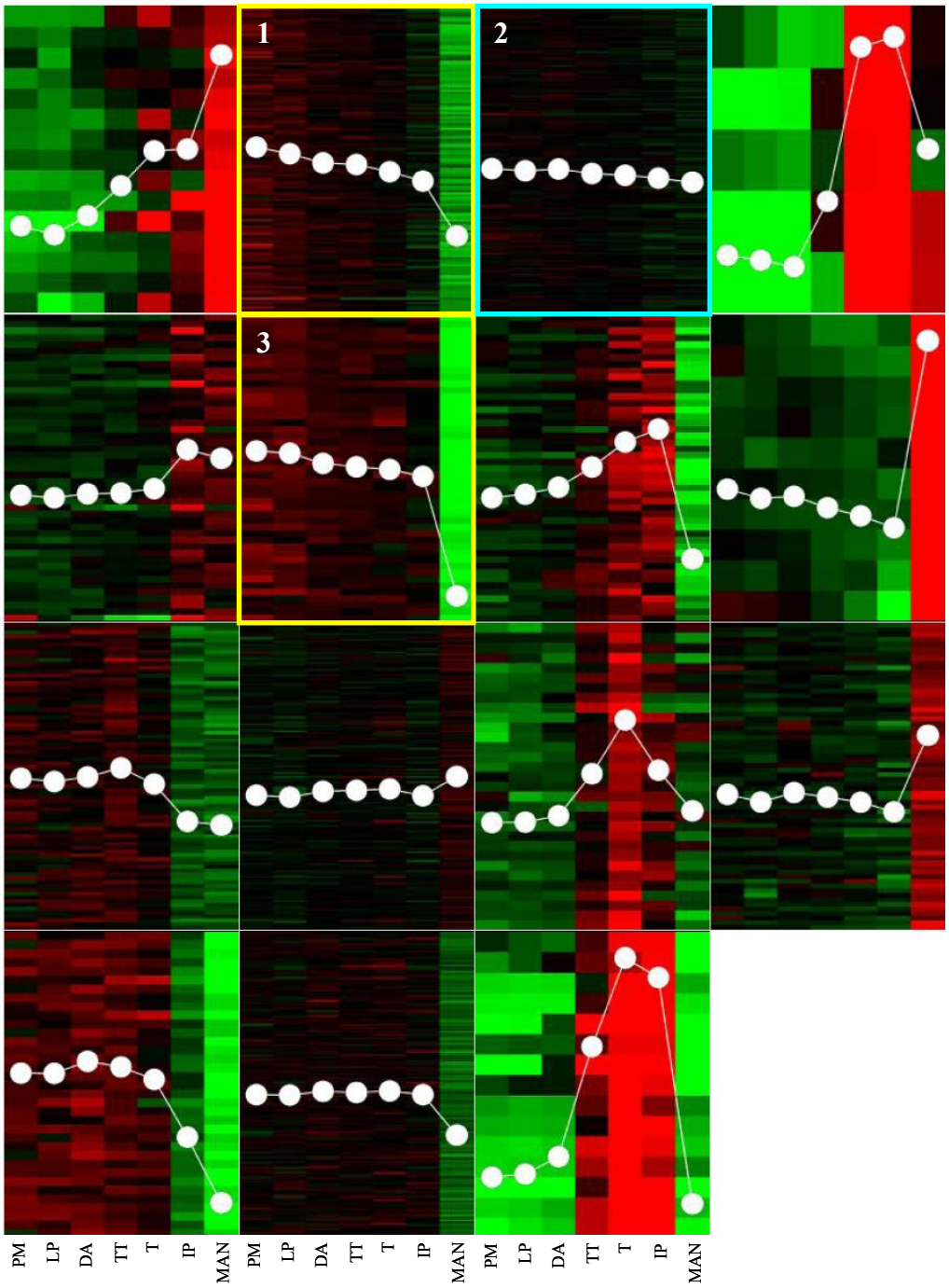
K-means clustering (Acuity 4.0) was conducted using 10, 15, 20, 25 and 30 partitions on the 1094 transcripts found to be meiotically regulated by Crismani and colleagues (2006). Highlighted profiles (boxed) were those which had higher expression during the first three stages of pre-meiosis (PM), leptotene to pachytene (LP) and diplotene to anaphase I (DA) when compared to telophase I to telophase II (TT), tetrads (T), immature pollen (IP) and mature anthers (MAN). Boxes highlighted in yellow represent the profiles with average expression levels above 9 (log base 2, RMA-normalised), considered to be abundantly expressed. Those in blue have average expression levels below 9 during the first three stages (log base 2, RMA-normalised) and are considered to be moderately expressed. The *k*-means cluster using 20 partitions returned 3 groups (containing 167 transcripts) with average expression levels above 9 during the first three stages, the most of any of the partitions, and thus was selected for further analysis. Tables represent the average absolute expression for each of the profiles highlighted; values are log base 2, RMA-normalised.

***K* = 10**



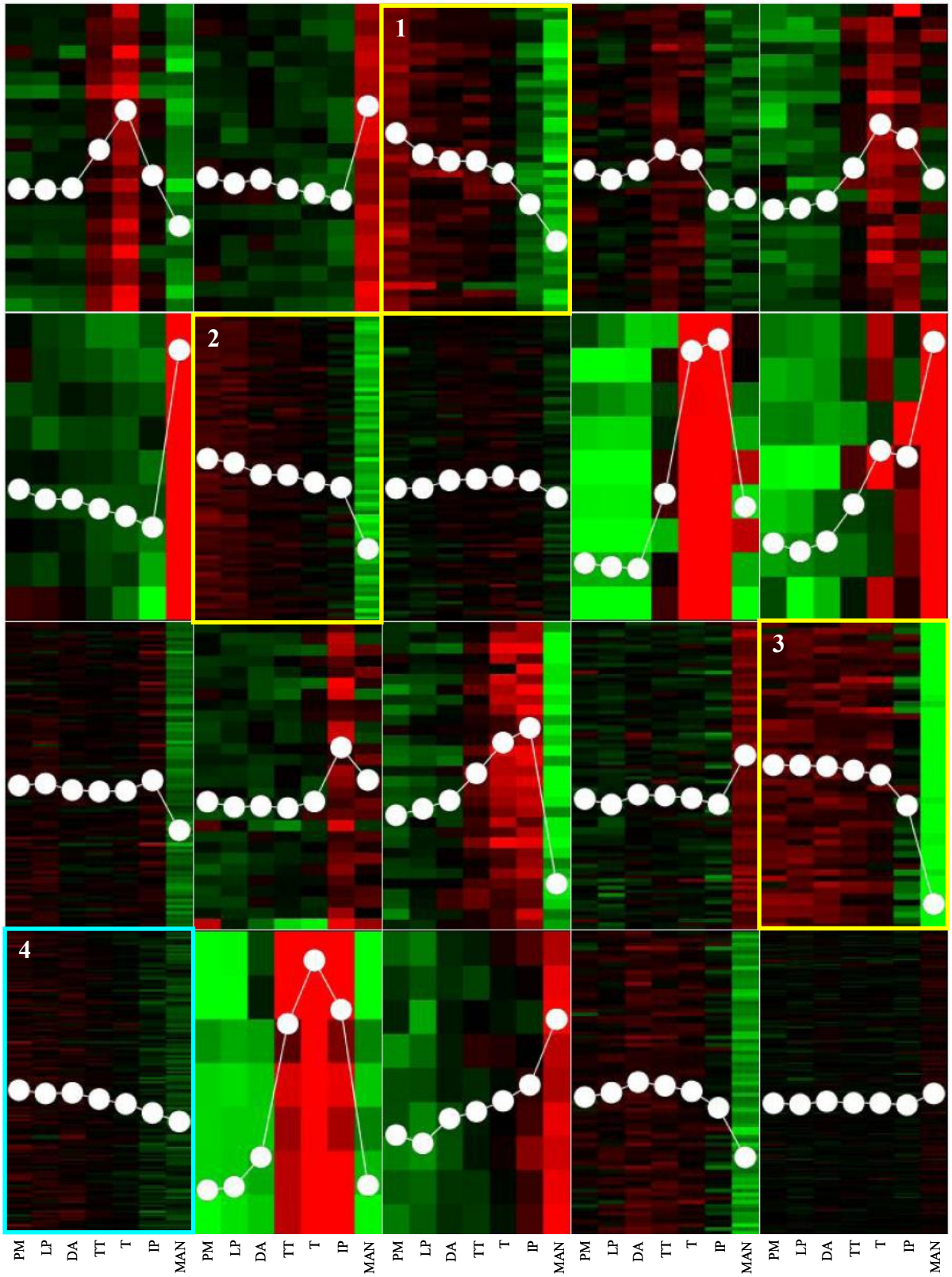
	PM	LP	DA	TT	T	IP	MAN
Profile 1	10.52	10.62	10.71	10.52	10.28	8.90	5.77
Profile 2	10.10	9.69	9.36	9.33	9.04	8.59	6.24
Profile 3	8.87	8.80	8.76	8.60	8.56	8.51	7.86

***K* = 15**



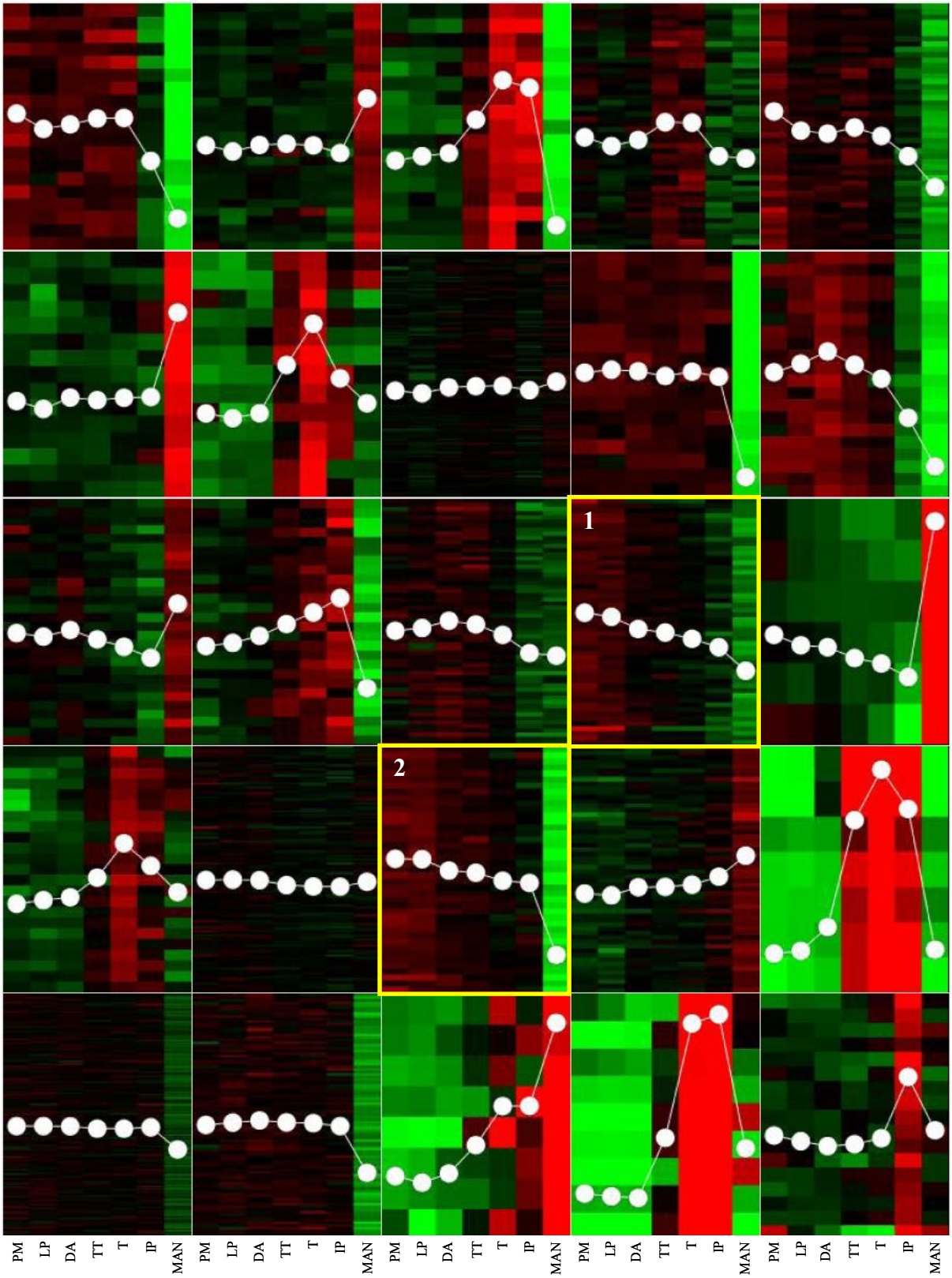
	PM	LP	DA	TT	T	IP	MAN
Profile 1	9.80	9.59	9.28	9.23	8.99	8.68	6.88
Profile 2	8.19	8.12	8.18	8.03	7.97	7.87	7.74
Profile 3	10.84	10.76	10.42	10.31	10.22	9.99	6.06

***K* = 20**



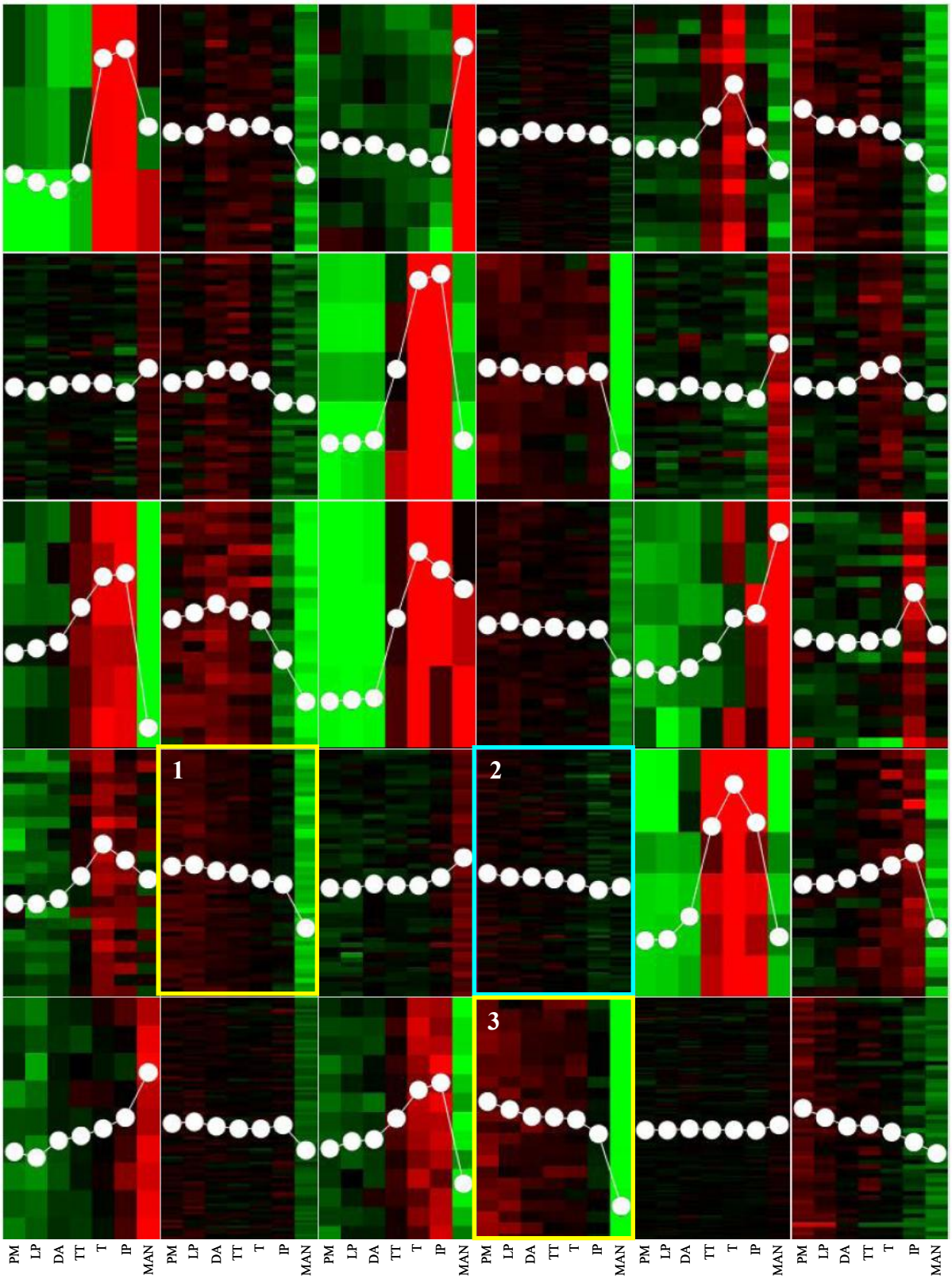
	PM	LP	DA	TT	T	IP	MAN
Profile 1	9.93	9.18	8.93	8.93	8.50	7.40	6.07
Profile 2	9.82	9.68	9.25	9.24	8.98	8.80	6.60
Profile 3	10.65	10.64	10.61	10.47	10.30	9.21	5.69
Profile 4	8.87	8.74	8.78	8.55	8.38	8.06	7.75

***K* = 25**



	PM	LP	DA	TT	T	IP	MAN
Profile 1	9.33	9.15	8.61	8.45	8.17	7.80	6.75
Profile 2	10.06	10.05	9.54	9.44	9.09	8.98	5.78

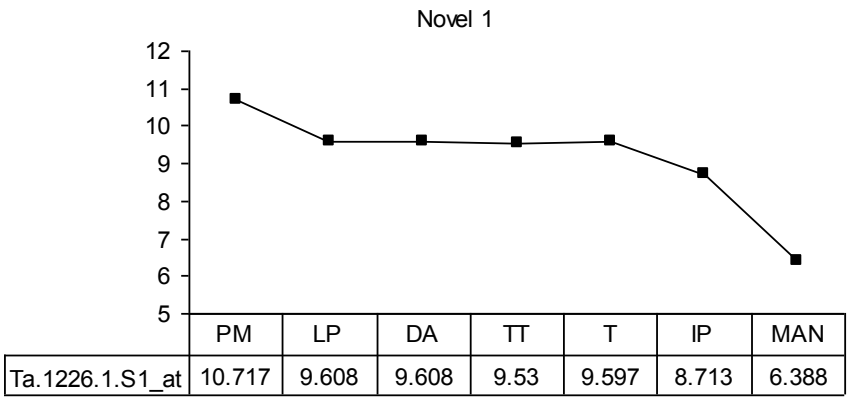
***K* = 30**



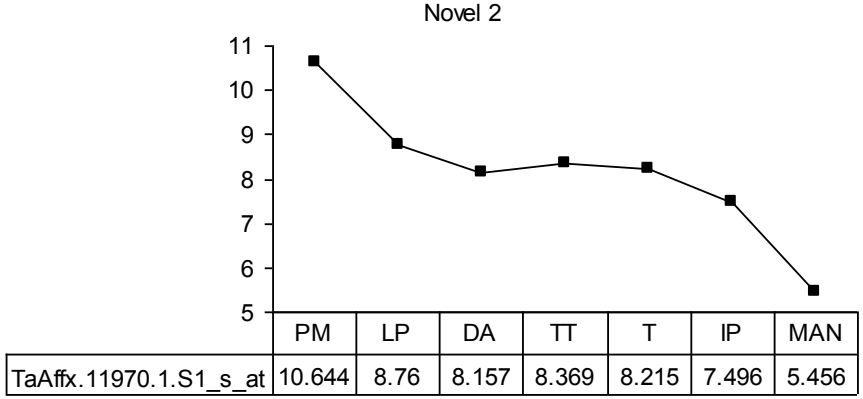
	PM	LP	DA	TT	T	IP	MAN
Profile 1	10.18	10.26	9.94	9.81	9.53	9.23	6.96
Profile 2	8.48	8.30	8.30	8.19	7.99	7.63	7.79
Profile 3	10.90	10.50	10.13	10.10	9.98	9.23	5.47

Appendix 2.4 – Expression Profiles of 39 novel transcripts chosen for further analysis

Expression profiles of the 39 transcripts chosen for further analysis in section 2.3.1 are shown. Expression values on the y axis are log base 2, RMA-normalised. Expression values >10 are considered to be highly expressed, values between 6 and 9 are considered to be moderately expressed, while values below 5 are considered to be expressed at a basal level. Stages of anther development are displayed on the x axis: pre-meiosis (PM), leptotene to pachytene (LP), diplotene to anaphase I (DA), telophase I to telophase II (TT), tetrads (T), immature pollen (IP) and mature anthers (MAN).

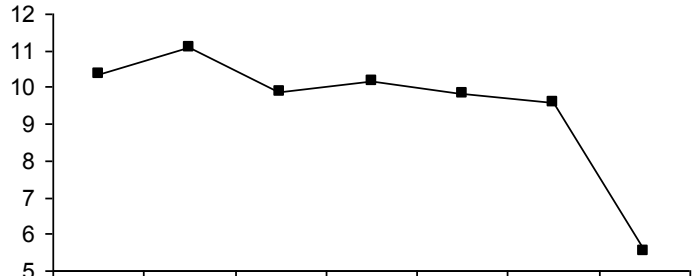


	PM	LP	DA	TT	T	IP	MAN
Ta.1226.1.S1_at	10.717	9.608	9.608	9.53	9.597	8.713	6.388



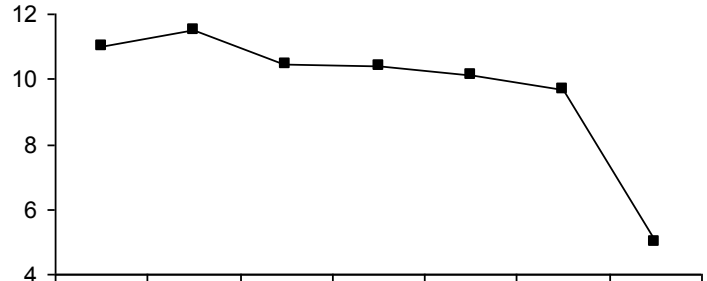
	PM	LP	DA	TT	T	IP	MAN
TaAffx.11970.1.S1_s_at	10.644	8.76	8.157	8.369	8.215	7.496	5.456

Novel 3



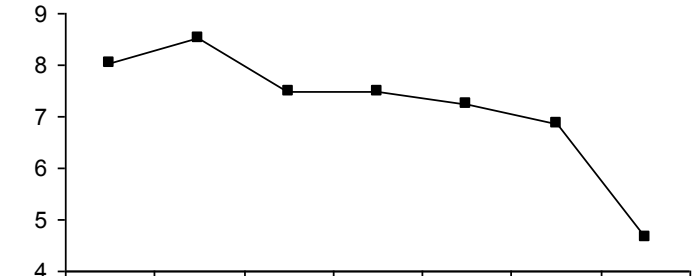
	PM	LP	DA	TT	T	IP	MAN
Ta.2952.1.S1_at	10.382	11.076	9.852	10.149	9.808	9.563	5.541

Novel 4



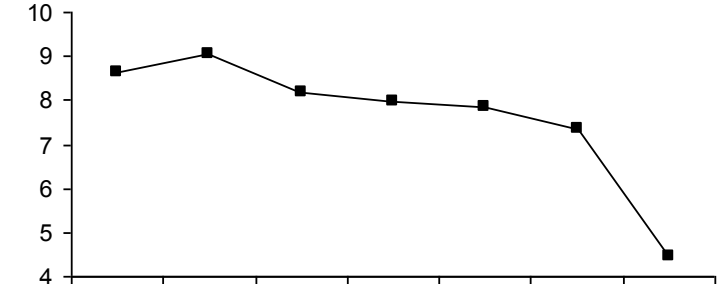
	PM	LP	DA	TT	T	IP	MAN
Ta.9089.1.A1_at	10.981	11.487	10.472	10.407	10.138	9.701	4.988

Novel 5



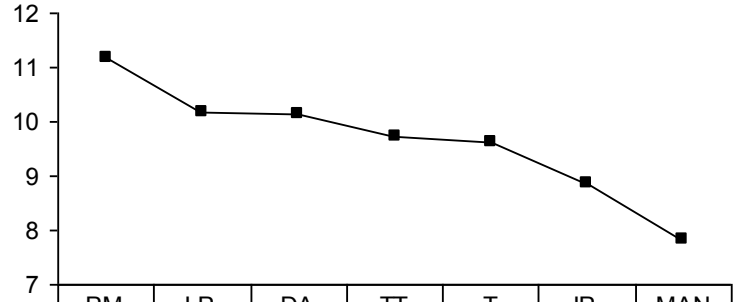
	PM	LP	DA	TT	T	IP	MAN
Ta.24603.1.A1_at	8.036	8.504	7.497	7.498	7.248	6.868	4.672

Novel 6



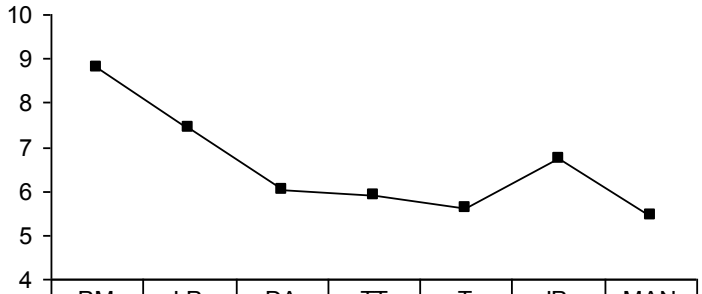
	PM	LP	DA	TT	T	IP	MAN
Ta.25984.1.S1_at	8.655	9.063	8.184	7.968	7.857	7.372	4.461

Novel 7



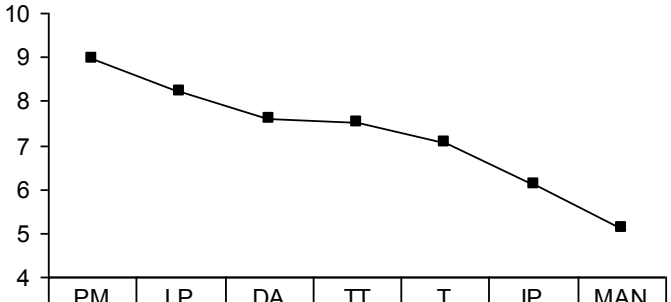
	PM	LP	DA	TT	T	IP	MAN
Ta.3397.1.S1_at	11.168	10.168	10.144	9.727	9.616	8.868	7.835

Novel 8



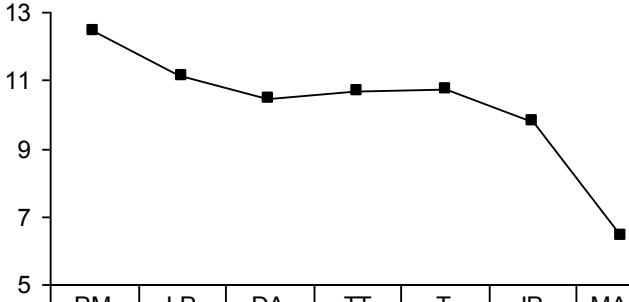
	PM	LP	DA	TT	T	IP	MAN
Ta.10020.1.S1_at	8.805	7.427	6.028	5.916	5.6	6.737	5.458

Novel 9



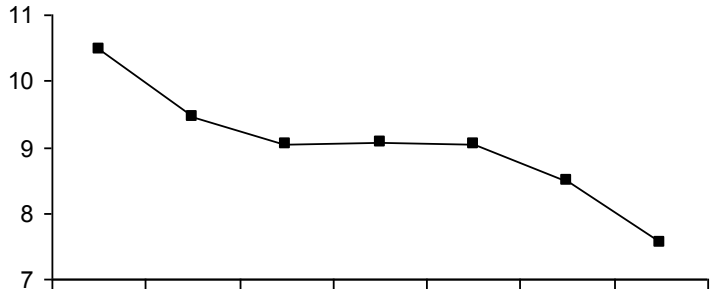
	PM	LP	DA	TT	T	IP	MAN
Ta.10024.1.S1_s_at	8.964	8.22	7.608	7.514	7.061	6.099	5.102

Novel 10



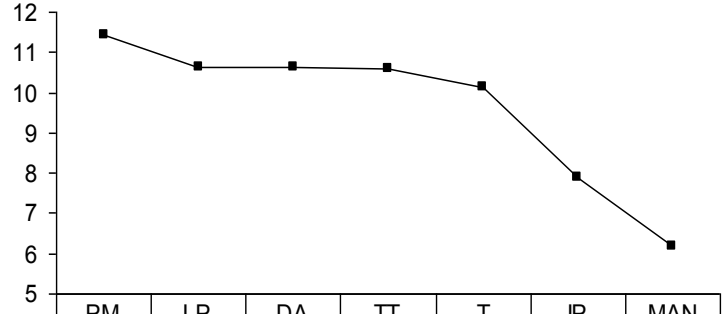
	PM	LP	DA	TT	T	IP	MAN
Ta.Affx.11970.2.S1_at	12.429	11.151	10.457	10.656	10.735	9.824	6.43

Novel 11



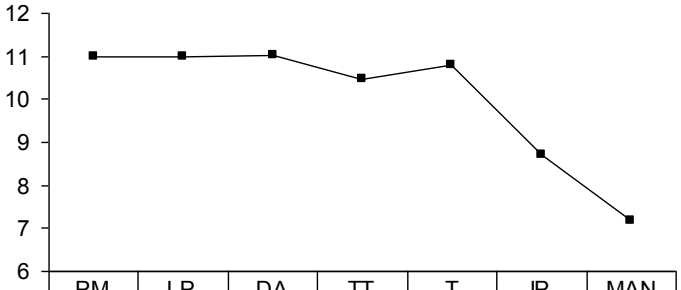
	PM	LP	DA	TT	T	lP	MAN
Ta.8260.1.A1_at	10.474	9.458	9.039	9.057	9.03	8.483	7.552

Novel 12



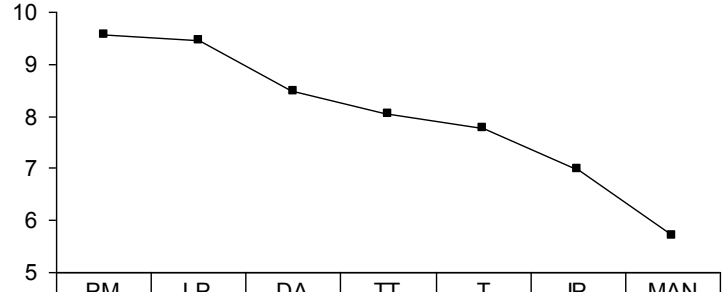
	PM	LP	DA	TT	T	lP	MAN
Ta.13989.1.S1_at	11.432	10.633	10.644	10.603	10.124	7.891	6.179

Novel 13



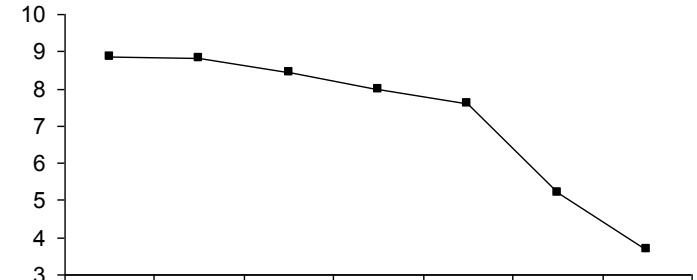
	PM	LP	DA	TT	T	lP	MAN
TaAffx.16307.1.S1_at	10.979	10.985	11.022	10.477	10.806	8.699	7.187

Novel 14



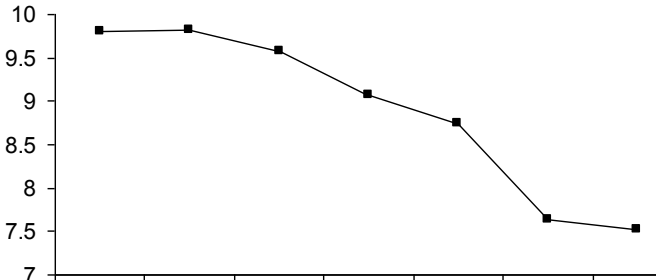
	PM	LP	DA	TT	T	lP	MAN
Ta.14043.1.S1_at	9.576	9.451	8.487	8.049	7.775	6.978	5.714

Novel 15



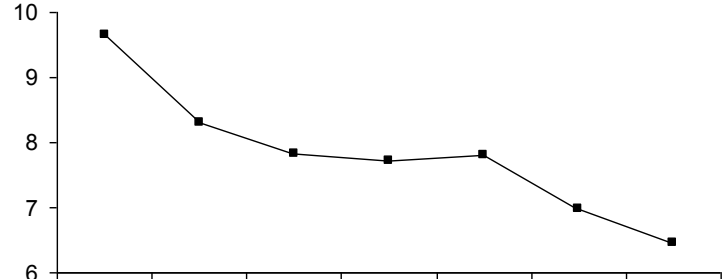
	PM	LP	DA	TT	T	lP	MAN
TaAffx.87249.1.S1_at	8.858	8.837	8.438	7.971	7.591	5.225	3.691

Novel 16



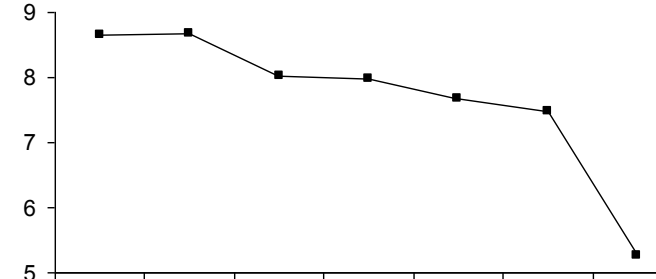
	PM	LP	DA	TT	T	lP	MAN
TaAffx.58843.1.S1_at	9.81	9.827	9.568	9.067	8.751	7.632	7.528

Novel 17



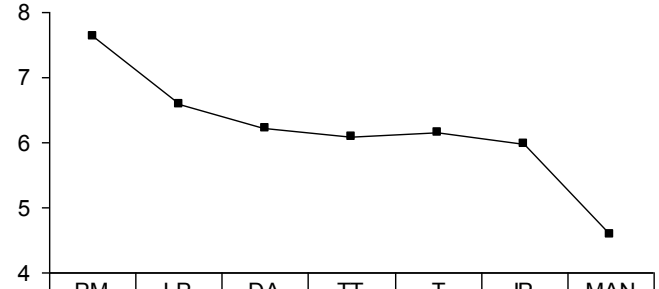
	PM	LP	DA	TT	T	lP	MAN
Ta.18725.1.S1_at	9.66	8.31	7.825	7.722	7.804	6.989	6.449

Novel 18



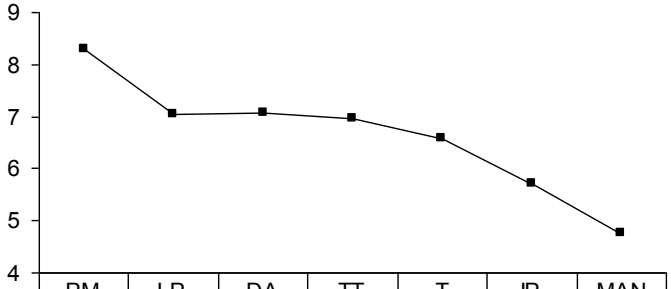
	PM	LP	DA	TT	T	lP	MAN
TaAffx.25097.1.S1_at	8.659	8.674	8.021	7.974	7.679	7.48	5.27

Novel 19



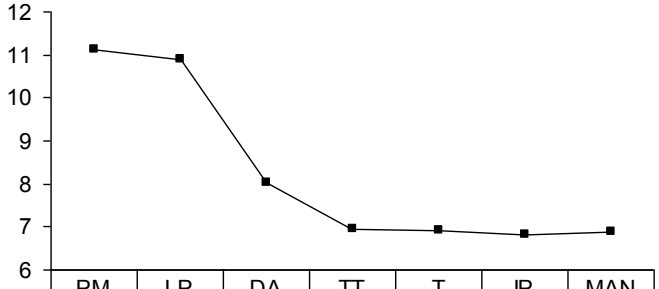
	PM	LP	DA	TT	T	IP	MAN
TaAffx.24409.1.S1_x_at	7.641	6.592	6.218	6.08	6.152	5.981	4.583

Novel 20



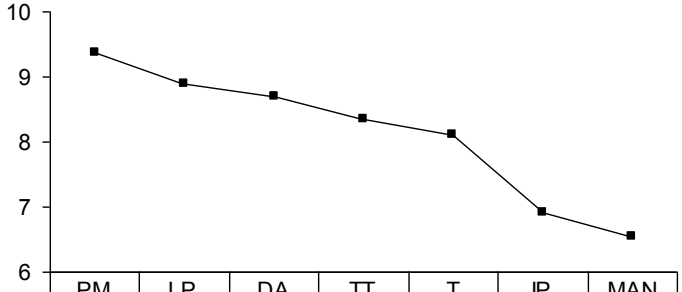
	PM	LP	DA	TT	T	IP	MAN
TaAffx.20193.1.A1_at	8.288	7.043	7.072	6.956	6.572	5.721	4.764

Novel 21



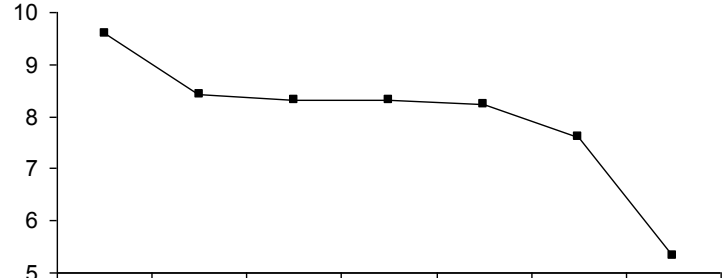
	PM	LP	DA	TT	T	IP	MAN
TaAffx.60258.2.S1_s_at	11.112	10.896	8.024	6.959	6.928	6.817	6.88

Novel 22



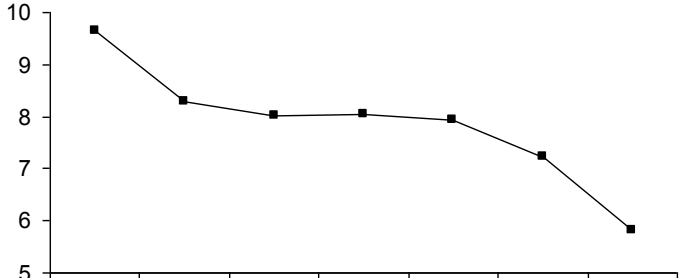
	PM	LP	DA	TT	T	IP	MAN
TaAffx.38287.1.S1_at	9.359	8.883	8.695	8.345	8.104	6.909	6.537

Novel 23



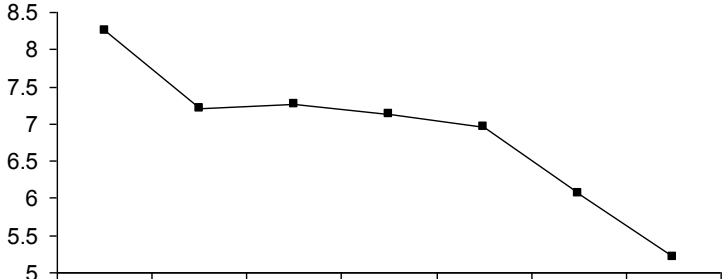
	PM	LP	DA	TT	T	IP	MAN
Ta.12761.1.A1_at	9.595	8.436	8.325	8.319	8.241	7.613	5.329

Novel 24



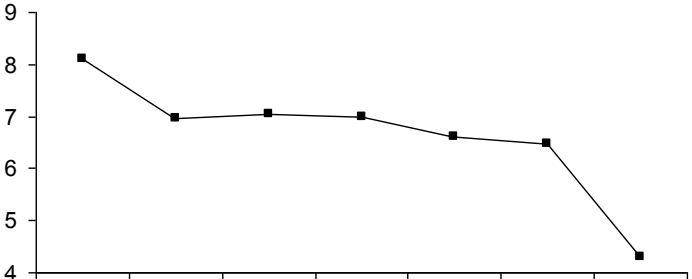
	PM	LP	DA	TT	T	IP	MAN
TaAffx.84212.1.S1_at	9.656	8.275	8.018	8.042	7.929	7.223	5.819

Novel 25

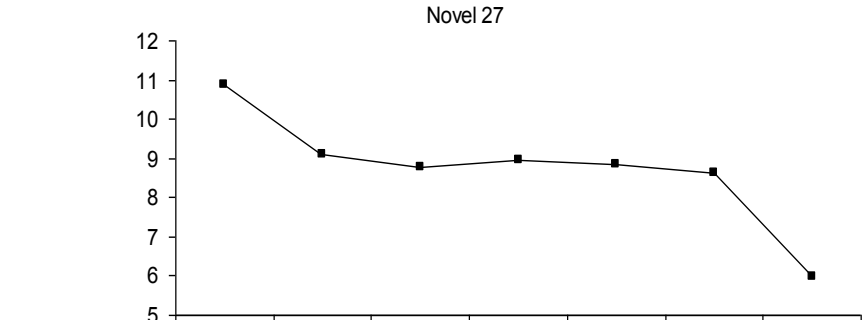


	PM	LP	DA	TT	T	IP	MAN
Ta.25933.1.S1_at	8.251	7.205	7.27	7.131	6.967	6.066	5.217

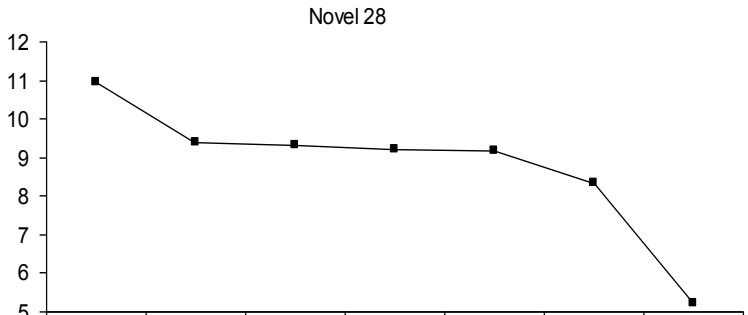
Novel 26



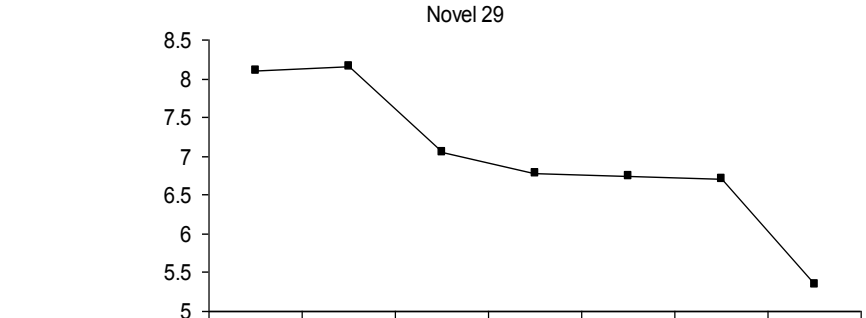
	PM	LP	DA	TT	T	IP	MAN
Ta.6748.2.A1_a_at	8.099	6.956	7.043	7	6.605	6.47	4.293



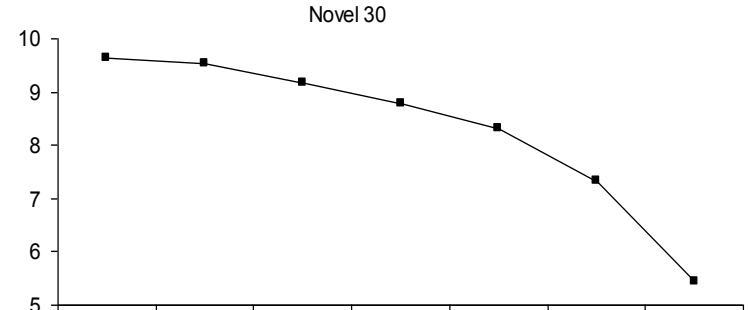
	PM	LP	DA	TT	T	IP	MAN
Ta.16123.1.A1_at	10.88	9.088	8.773	8.956	8.861	8.623	5.971



	PM	LP	DA	TT	T	IP	MAN
Ta.6748.1.A1_at	10.961	9.394	9.311	9.209	9.176	8.349	5.23

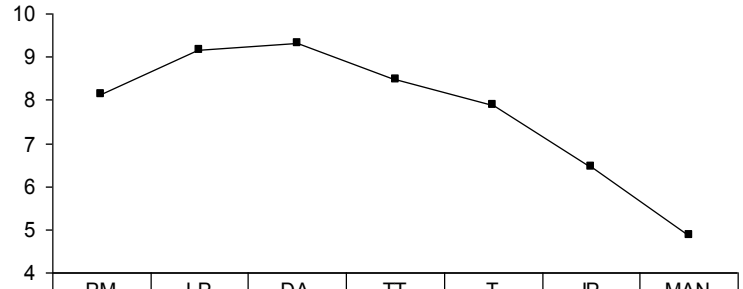


	PM	LP	DA	TT	T	IP	MAN
TaAffx.44789.1.S1_at	8.093	8.158	7.043	6.778	6.739	6.712	5.351



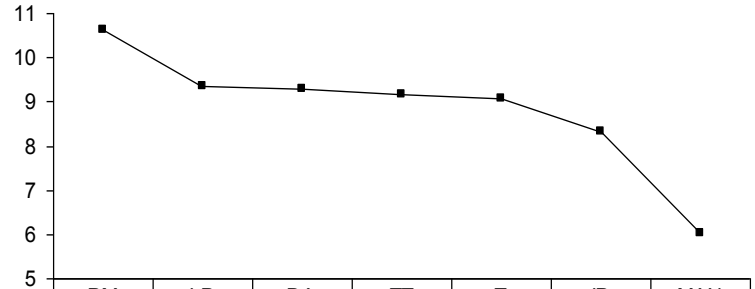
	PM	LP	DA	TT	T	IP	MAN
Ta.16224.1.S1_at	9.633	9.532	9.17	8.794	8.308	7.331	5.434

Novel 31



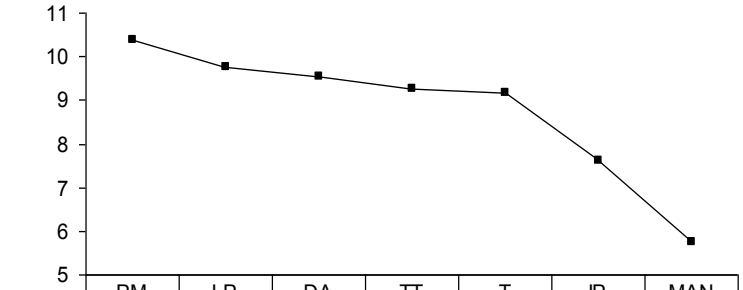
	PM	LP	DA	TT	T	IP	MAN
Ta.18287.1.S1_at	8.147	9.151	9.313	8.488	7.891	6.463	4.855

Novel 32



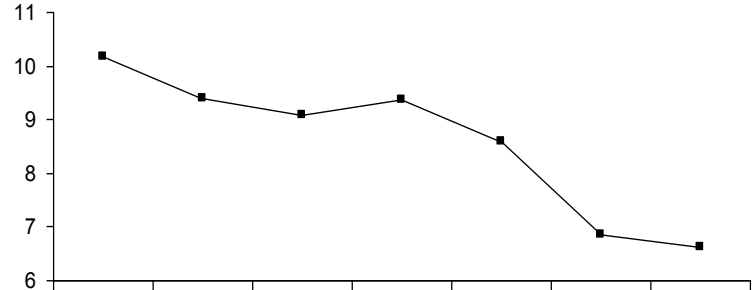
	PM	LP	DA	TT	T	IP	MAN
Ta.9286.1.S1_at	10.615	9.358	9.286	9.165	9.06	8.318	6.031

Novel 33



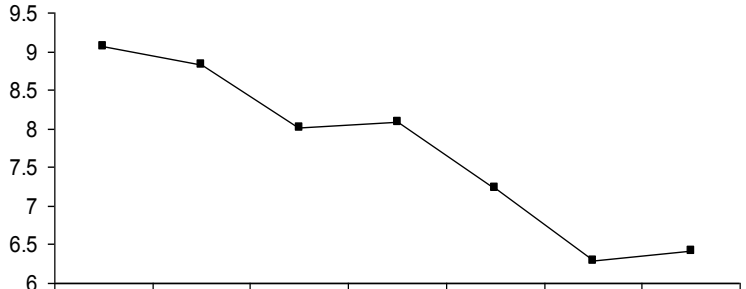
	PM	LP	DA	TT	T	IP	MAN
TaAffx.65731.1.A1_at	10.382	9.77	9.526	9.245	9.166	7.613	5.744

Novel 34



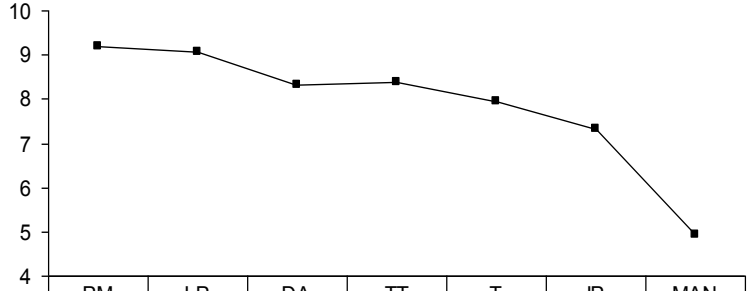
	PM	LP	DA	TT	T	IP	MAN
Ta.1673.1.S1_at	10.168	9.384	9.086	9.375	8.588	6.858	6.616

Novel 35



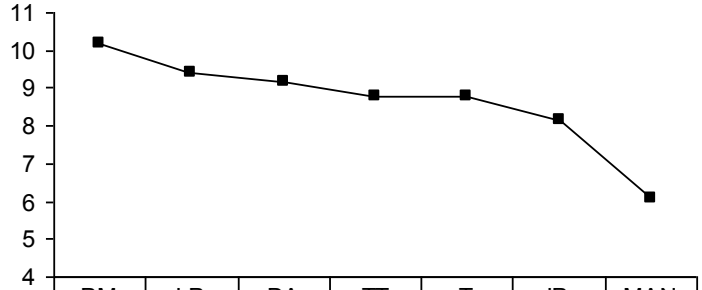
	PM	LP	DA	TT	T	IP	MAN
Ta.11783.1.A1_at	9.071	8.82	8.019	8.092	7.227	6.285	6.423

Novel 36



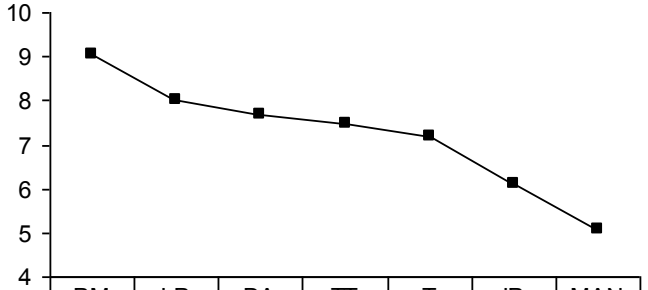
	PM	LP	DA	TT	T	IP	MAN
Ta.3093.2.A1_at	9.18	9.076	8.326	8.392	7.953	7.328	4.943

Novel 37

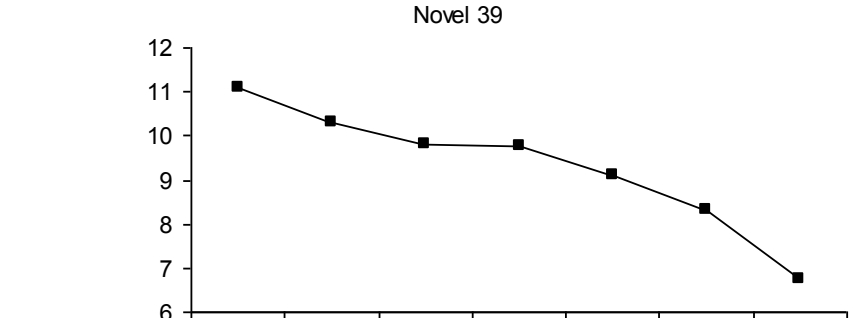


	PM	LP	DA	TT	T	IP	MAN
Ta.25584.1.S1_at	10.2	9.387	9.184	8.794	8.8	8.159	6.053

Novel 38



	PM	LP	DA	TT	T	IP	MAN
TaAffx.124106.1.S1_at	9.042	8.02	7.697	7.471	7.178	6.102	5.059

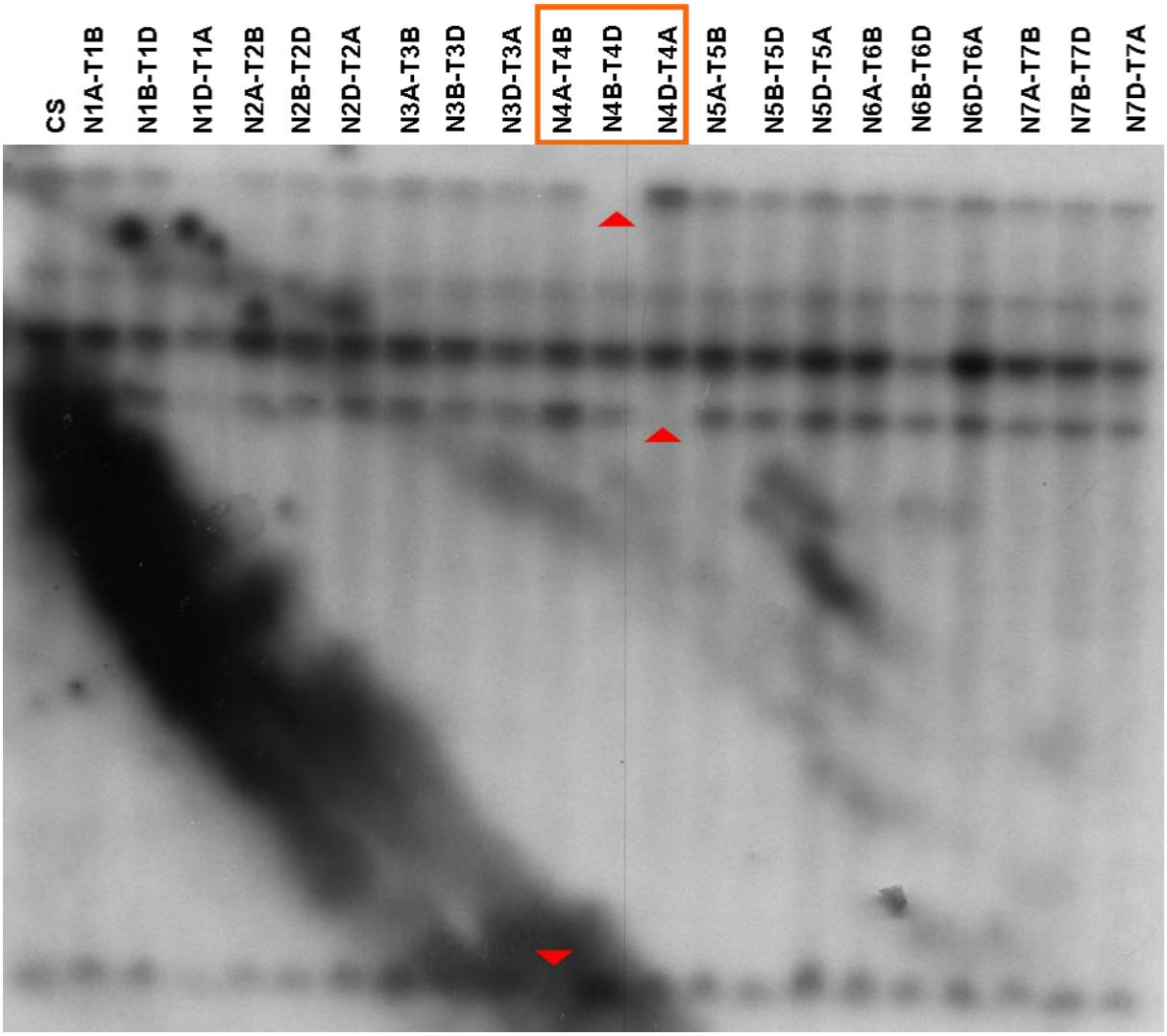


	PM	LP	DA	TT	T	IP	MAN
Ta.9807.1.A1_at	11.094	10.287	9.801	9.755	9.088	8.32	6.728

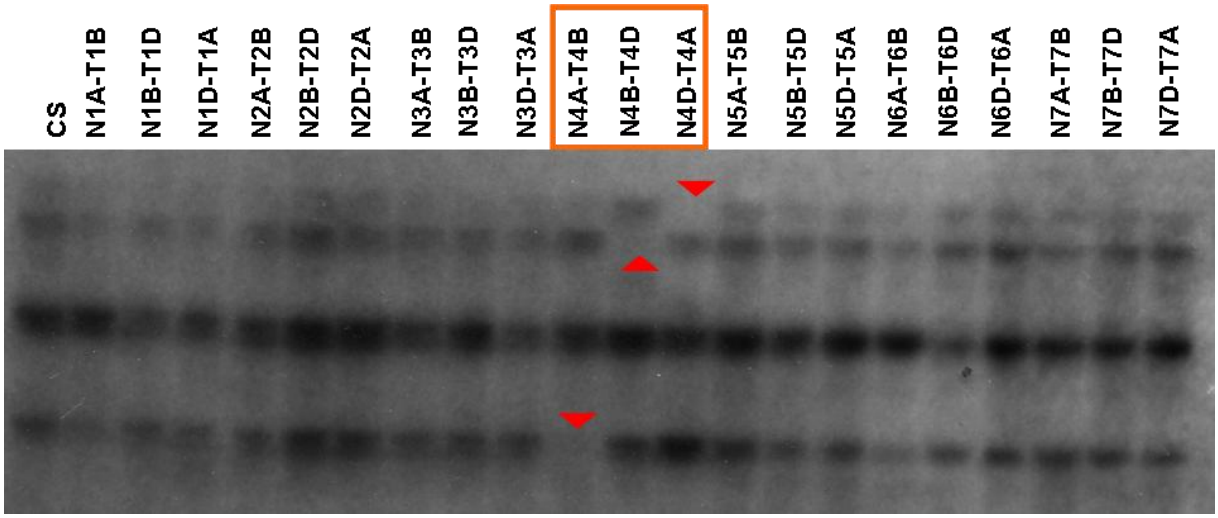
Appendix 2.5 – Chromosome locations of novel transcripts

Chromosome locations of the novel transcripts was determined via Southern autoradiography, using membranes which contained digested genomic DNA (with one of the following enzymes: *Bam*HI, *Dra*I, *Eco*RI, *Eco*RV or *Xba*I) of Chinese Spring (CS) and 21 CS derivatives which are nullisomic for a particular chromosome and tetra compensated with a homoeologous chromosome. For example, the absence of the two copies of chromosome 1A is compensated with the presence of another 2 copies of 1B chromosomes. The loss of a hybridisation signal on two or three of the chromosomes of a particular group was sufficient evidence to predict the chromosome location of that transcript. Arrows represent the loss of a hybridisation signal, while the box highlights the chromosome group location. Selections of the Southern analysis are shown here, with the remaining autorads not reproducible via scanning. Meiotically expressed novels were predominantly located on chromosome groups 2, 5 and 7.

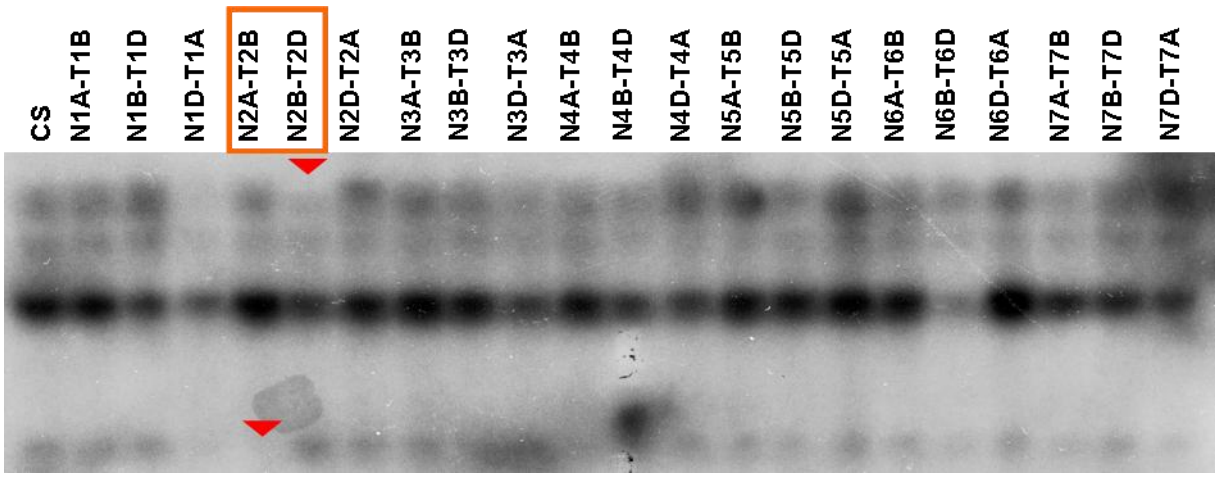
MN02 – *Eco*RI – Group 4



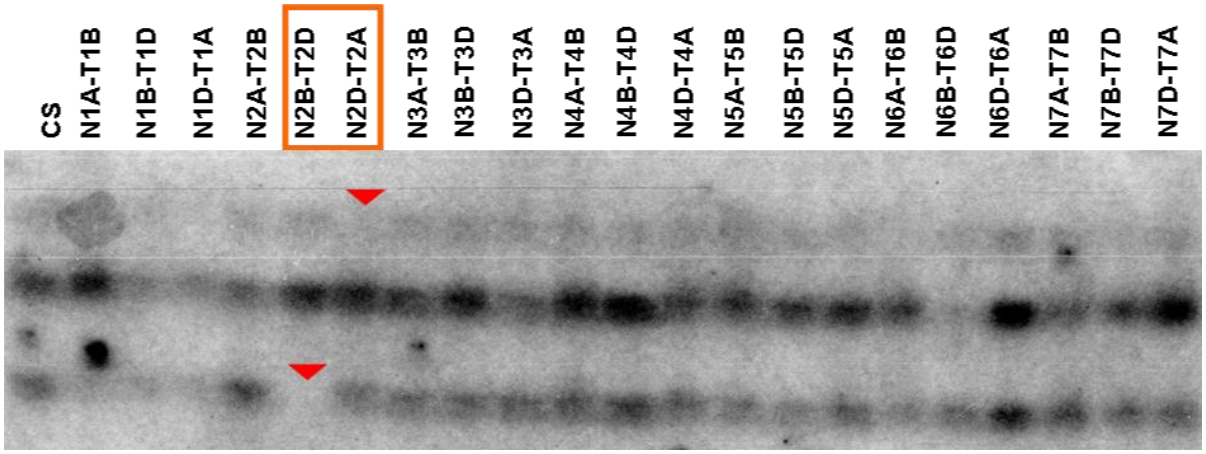
MN10 – *Xba*I – Group 4



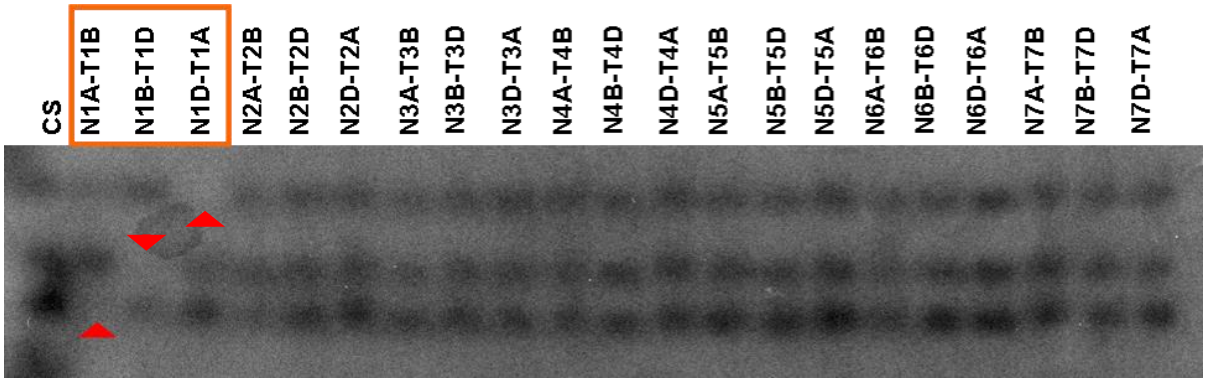
MN18 – *Eco*RI - Group 2



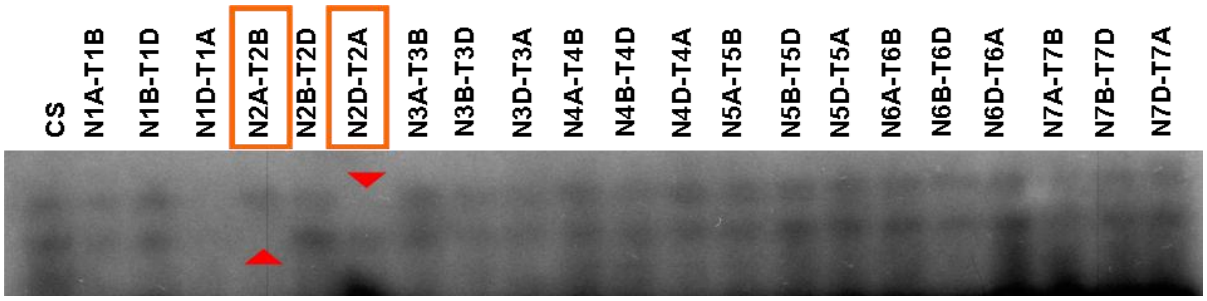
MN19 – *Dra*I – Group 2



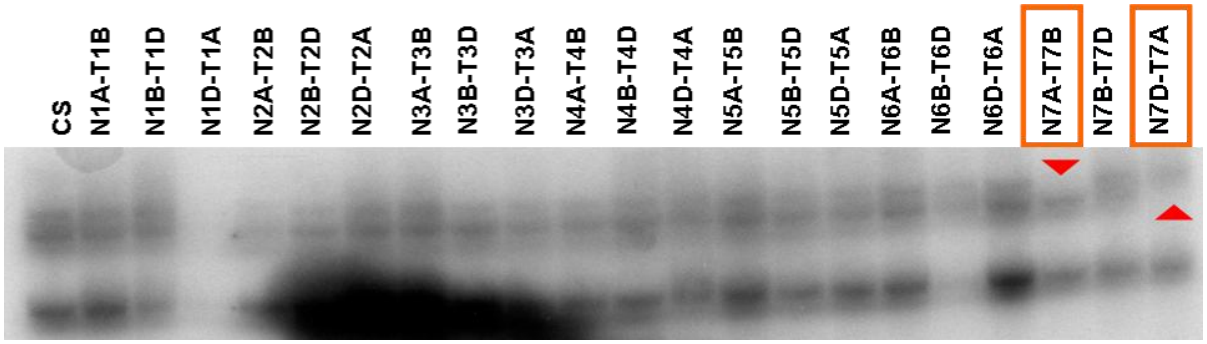
MN22 – *Dra*I – Group 1



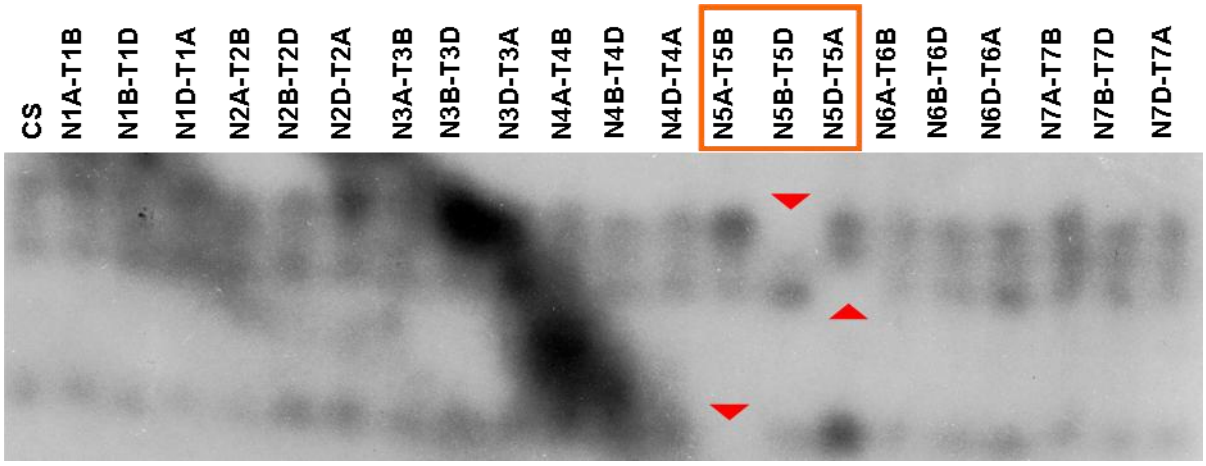
MN25 – *Eco*RV – Group 2



MN26 – *Xba*I – Group 7



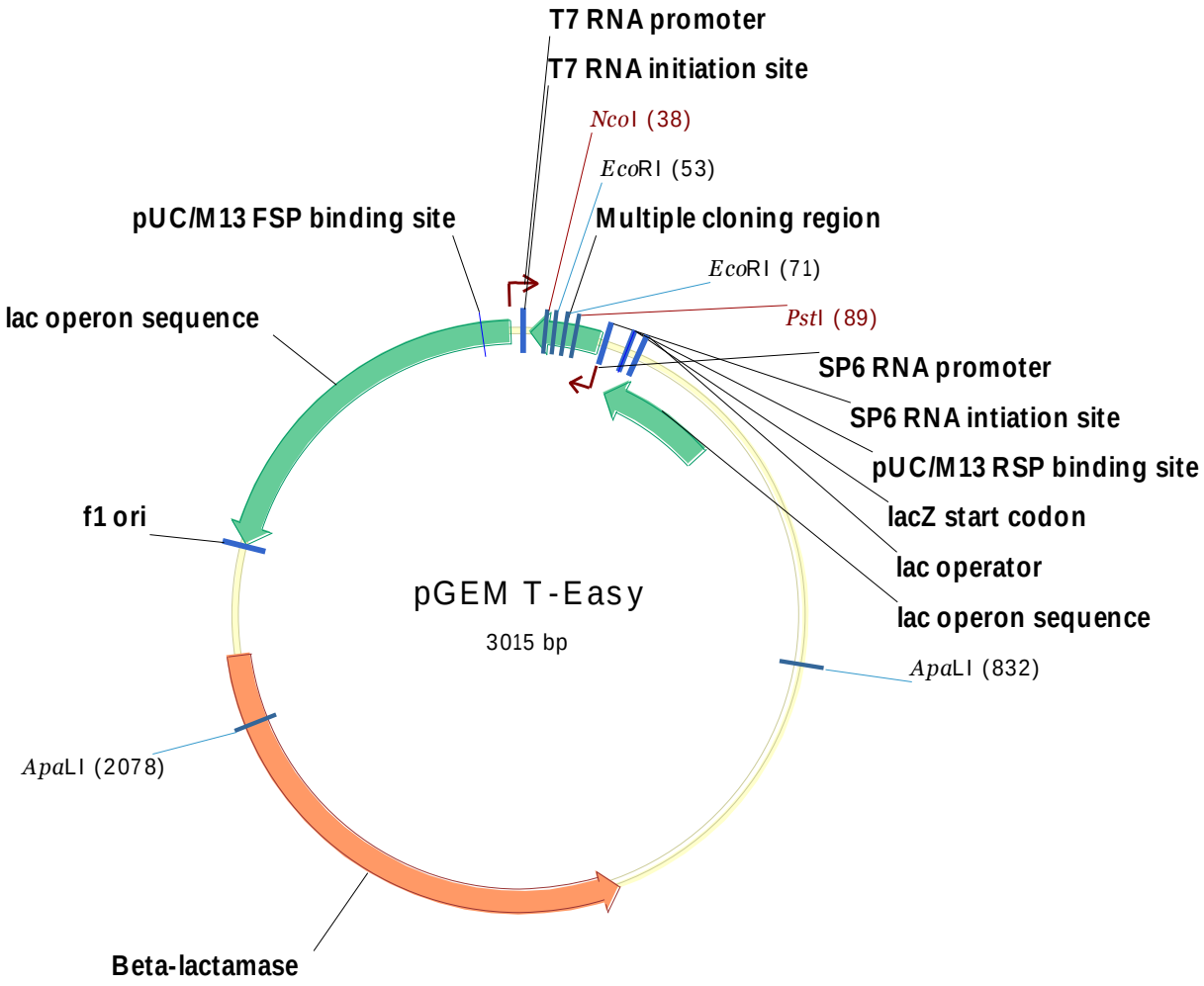
MN29 – *Dra*I – Group 5



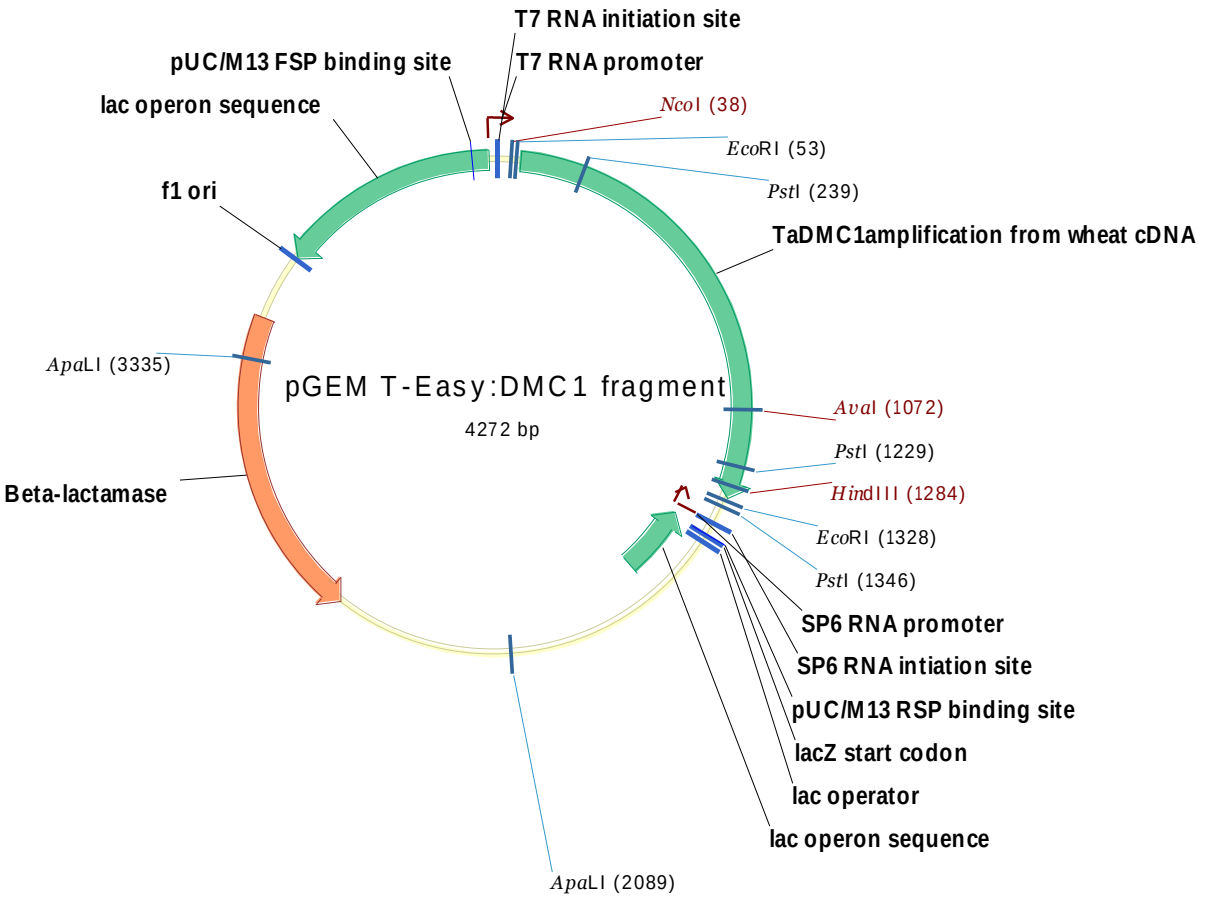
Appendix 3.1 - pGEM®-T Easy vectors used for PCR isolation of *TaDMC1*, *TaHOP2* and *TaMND1*

The pGEM®-T Easy Vector (i) (Invitrogen, AUS) was used for the isolation of PCR fragments and sequence conformation of *TaDMC1* (*ii*), *TaHOP2* (*iii*) and *TaMND1* (*iv*).

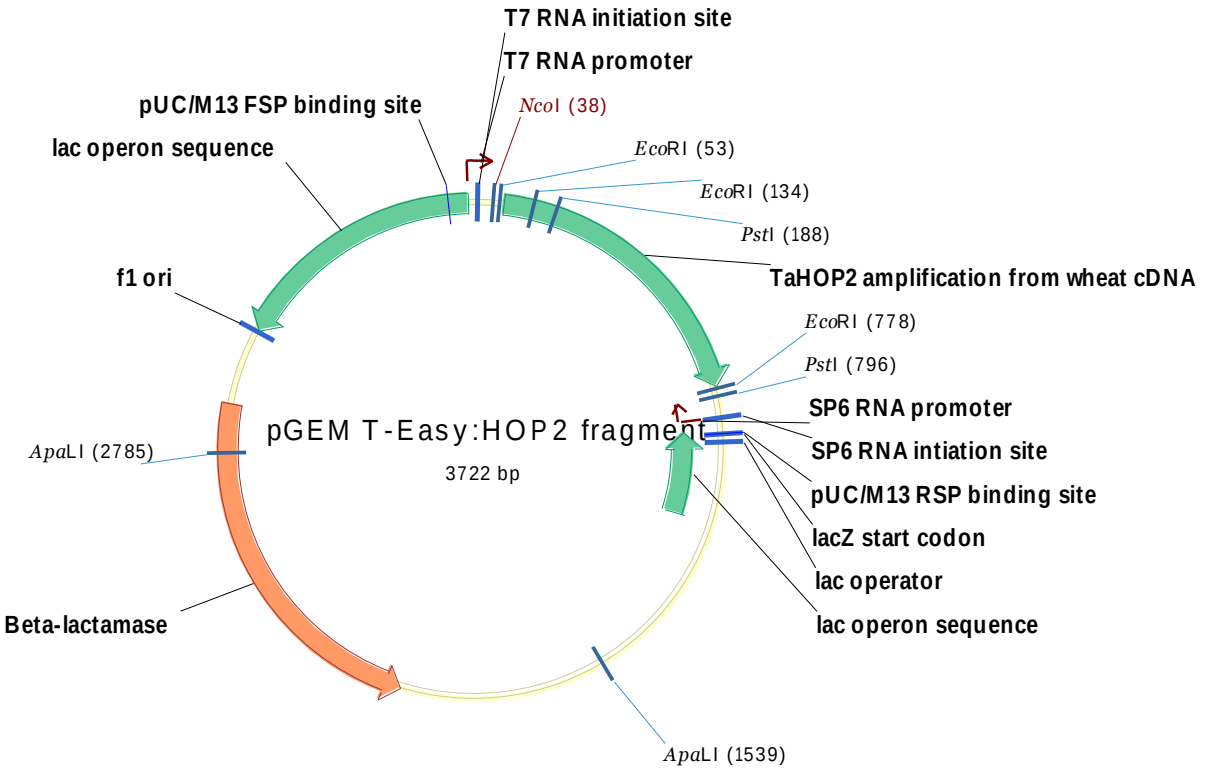
***i* - pGEM®-T Easy Vector**



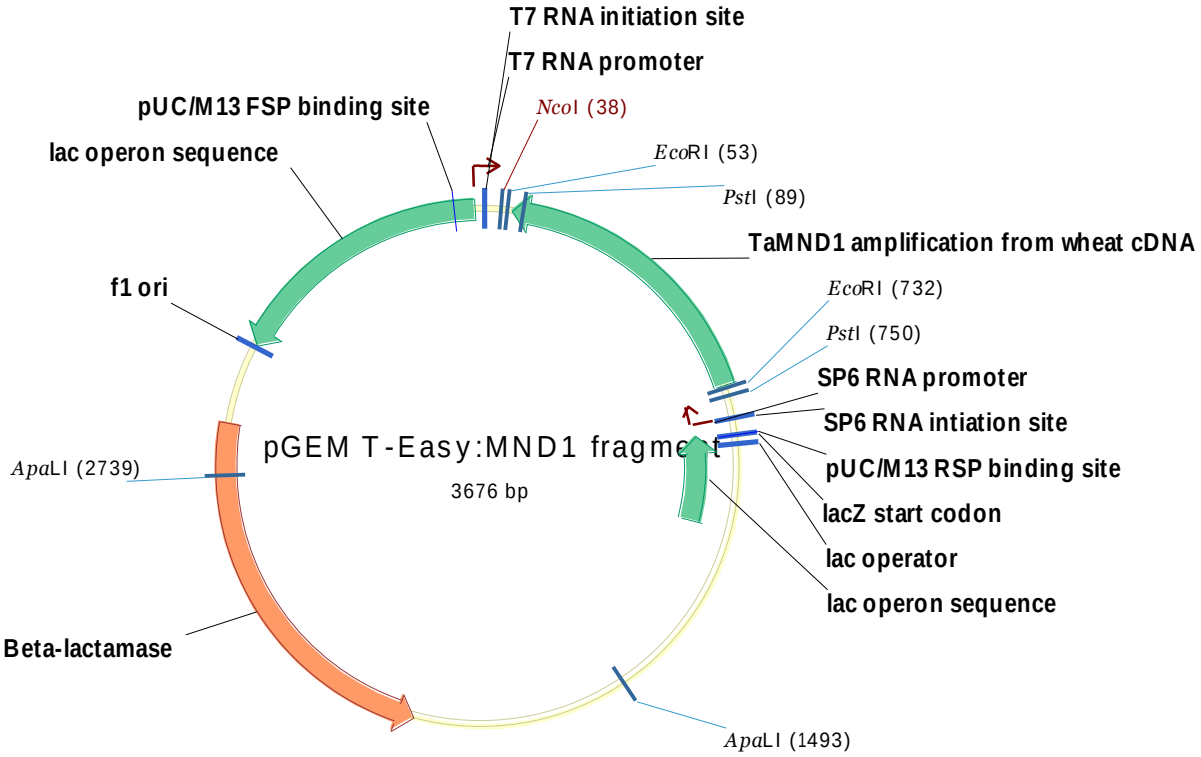
ii – TaDMC1:pGEM®-T Easy Vector



iii – TaHOP2:pGEM®-T Easy Vector



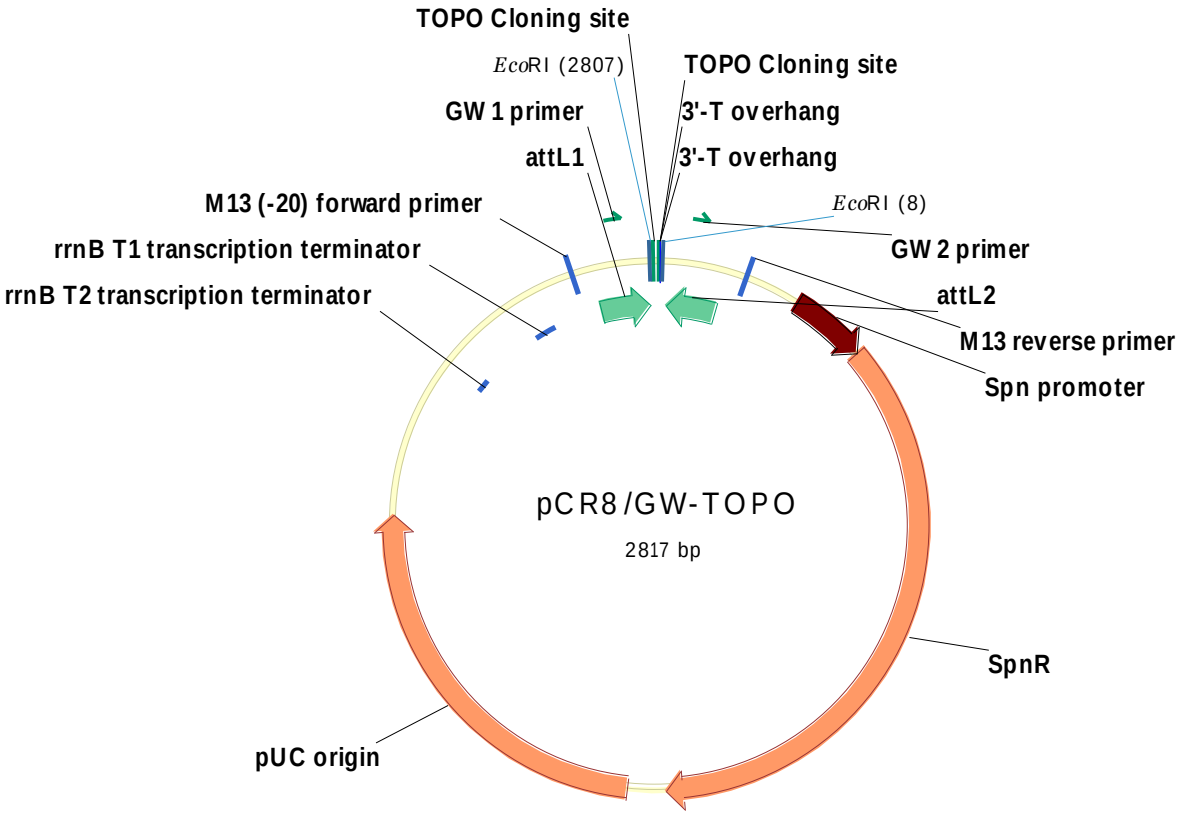
***iv – TaMND1*:pGEM®-T Easy Vector**



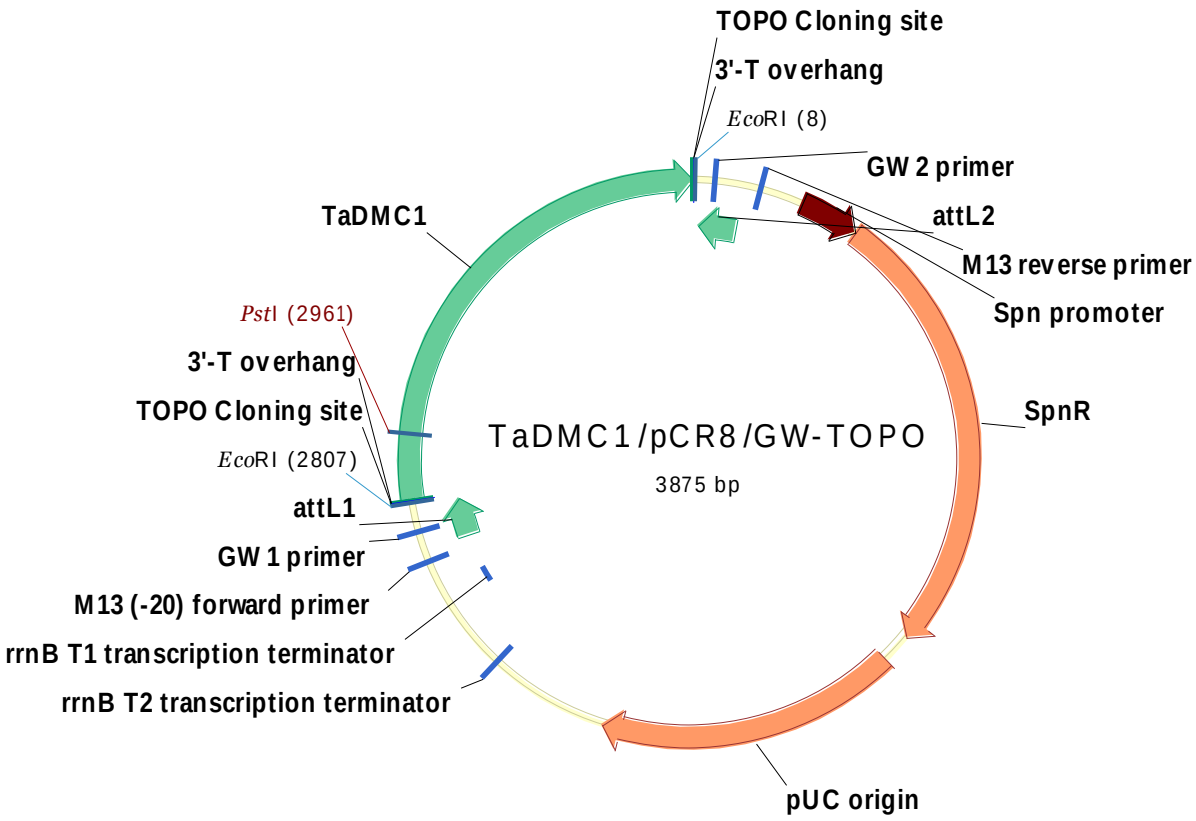
Appendix 3.2 - pCR8®/GW/TOPO® vector used as an intermediate vector for *TaDMC1*, *TaHOP2* and *TaMND1*

The pCR8®/GW/TOPO® vector (*i*) (Invitrogen, AUS) was used as an intermediate vector for cloning the open reading frames of *TaDMC1* (*ii*), *TaHOP2* (*iii*) and *TaMND1* (*iv*) before transferring into the protein expression vector pDEST17®.

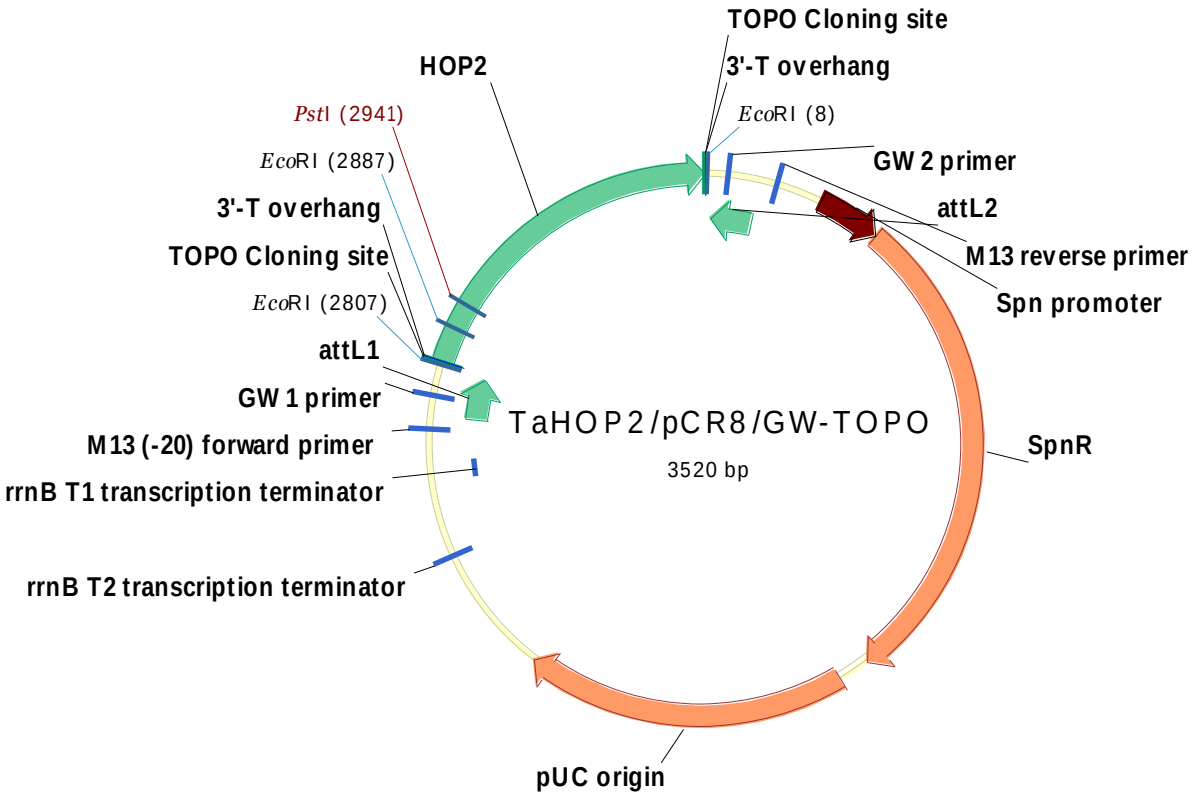
i - pCR8®/GW/TOPO® Vector



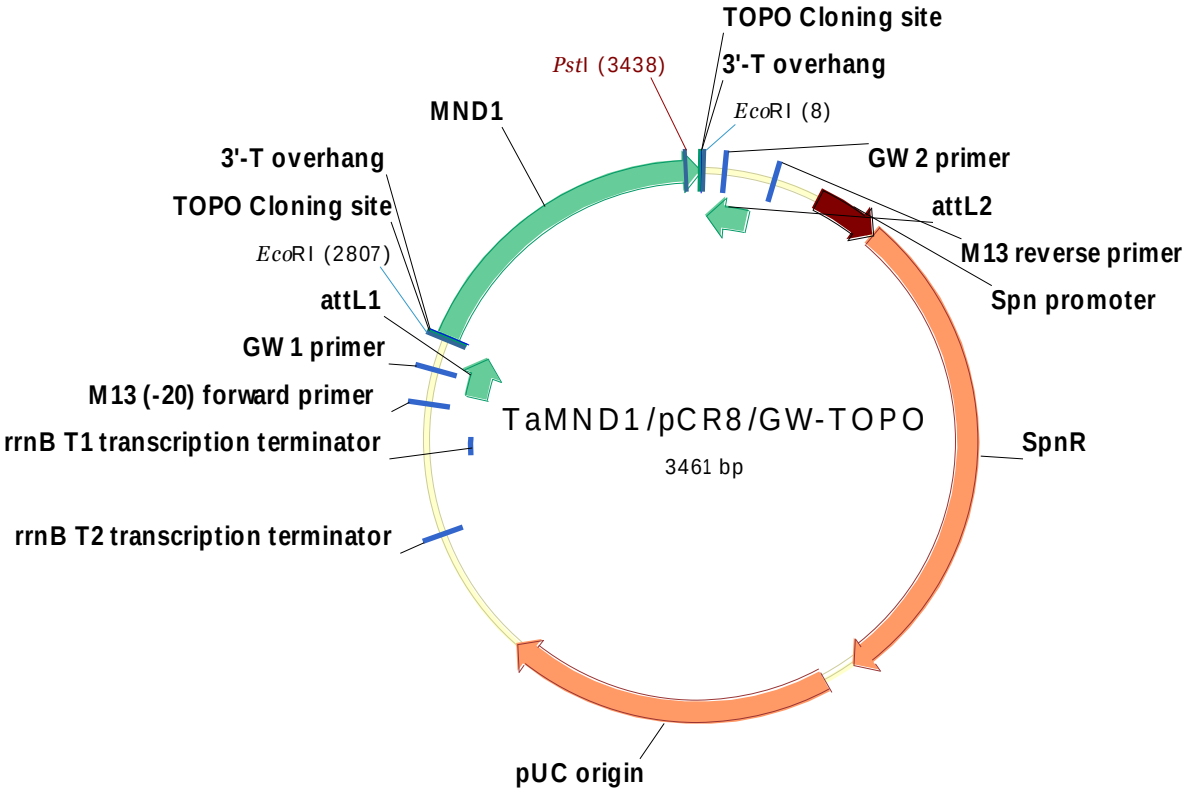
ii – TaDMC1:pCR8®/GW/TOPO® Vector



iii – TaHOP2:pCR8®/GW/TOPO® Vector



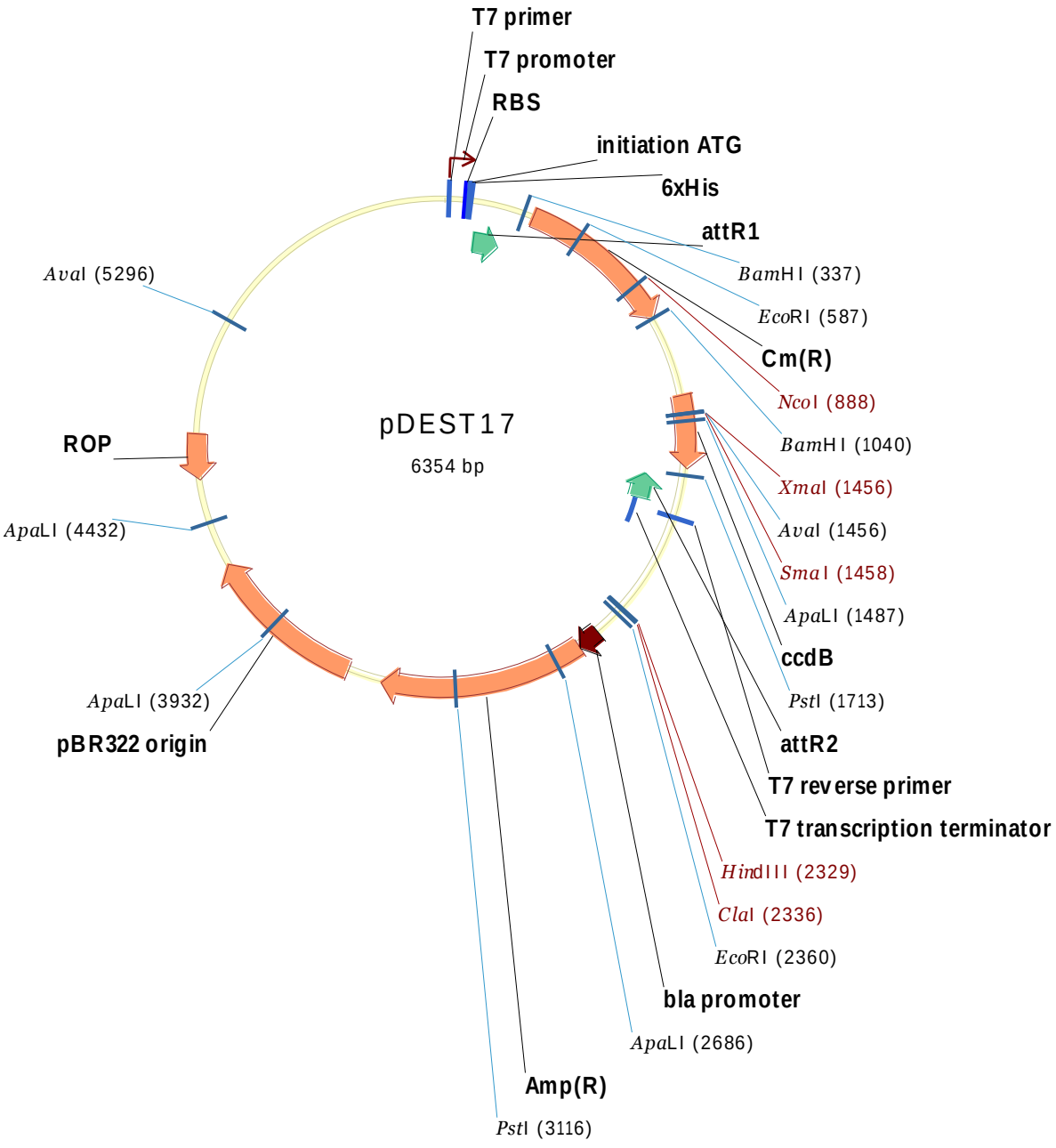
***iv – TaMND1*:pCR8®/GW/TOPO® Vector**



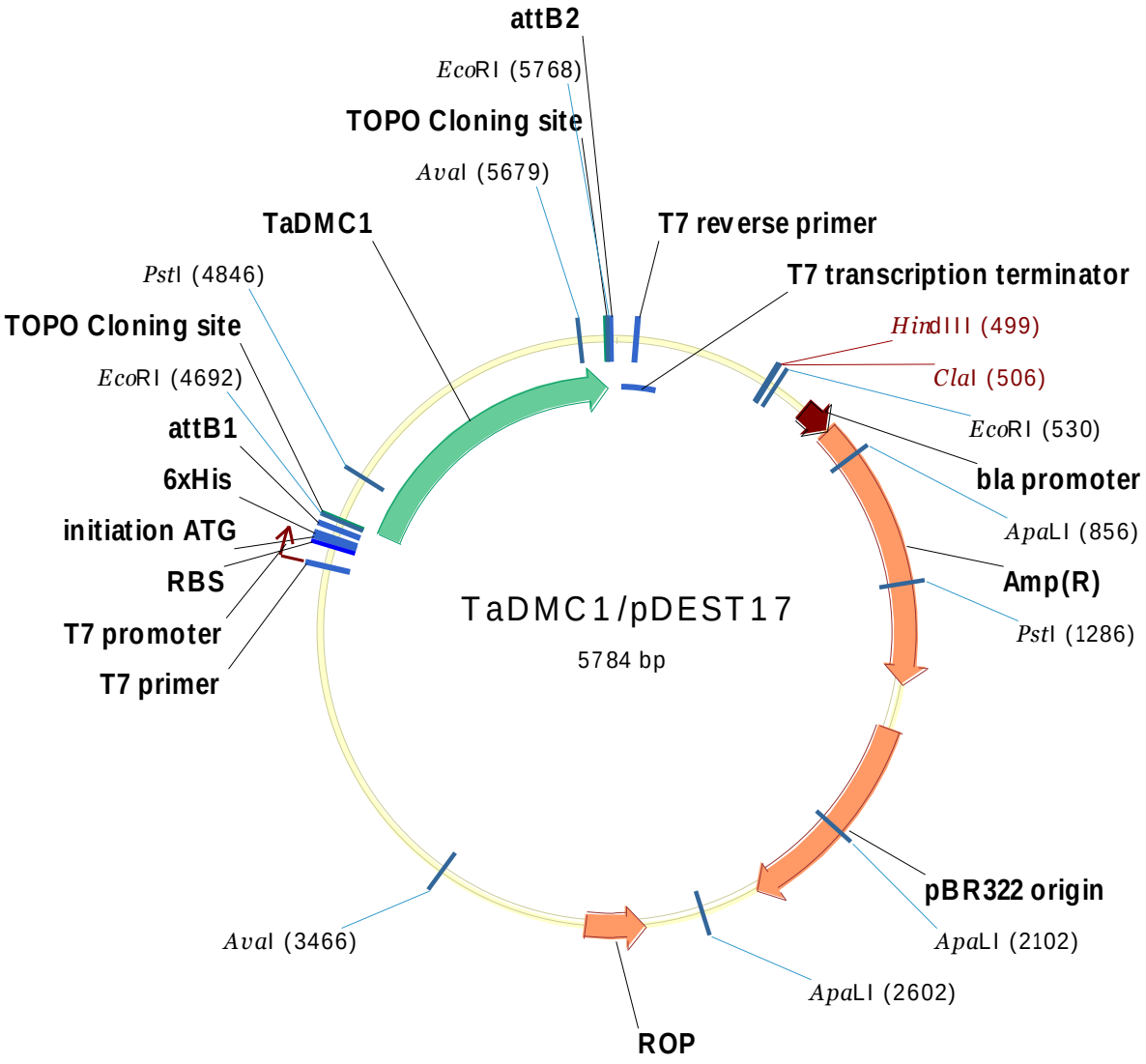
Appendix 3.3 - pDEST17® protein expression vectors of *TaDMC1*, *TaHOP2* and *TaMND1*

The pDEST17® vector (*i*) (Invitrogen, AUS) was used for heterologous protein expression of the meiotic candidates *TaDMC1* (*ii*), *TaHOP2* (*iii*) and *TaMND1* (*iv*).

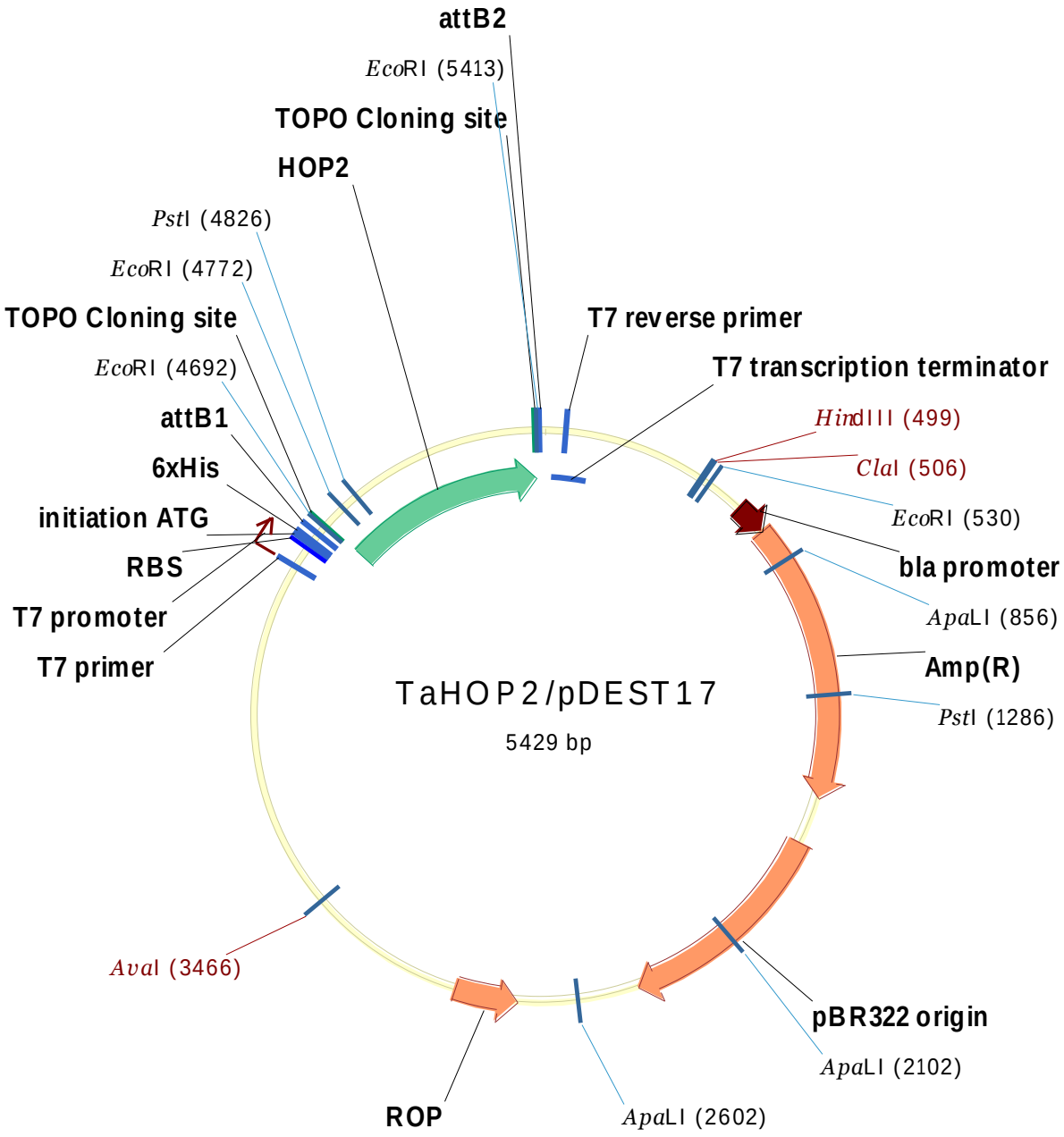
i – pDEST17® Vector



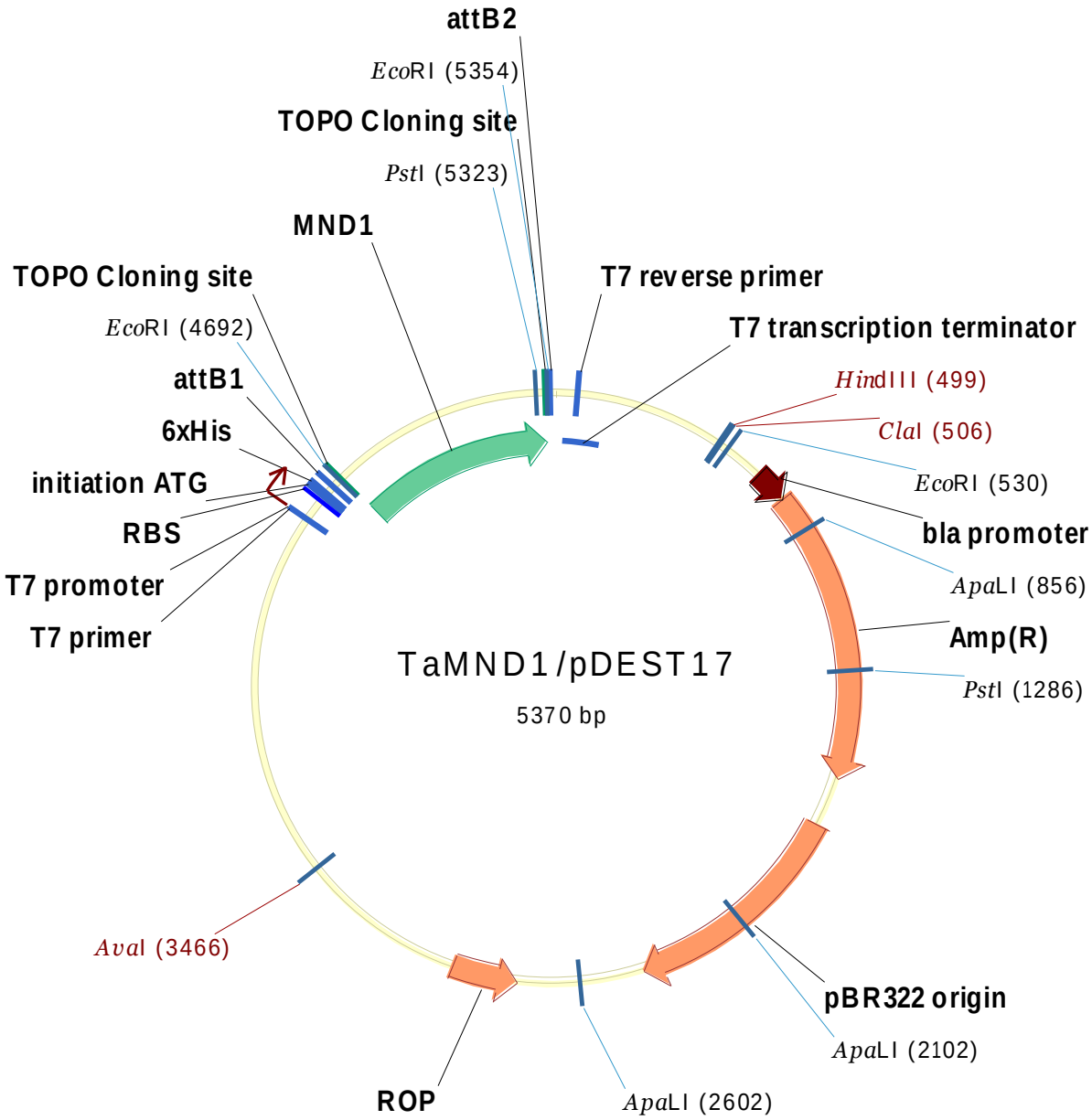
ii – TaDMC1:pCR8®/GW/TOPO® Vector



iii – TaHOP2:pCR8®/GW/TOPO® Vector



iv – TaMND1:pCR8®/GW/TOPO® Vector



Appendix 3.4 – Confirmed open reading frames of *TaDMC1*, *TaHOP2* and *TaMND1*

Open reading frames for *TaDMC1* (1032 bp), *TaHOP2* (681 bp) and *TaMND1* (621 bp), but not including the stop codons.

TaDMC1

ATGGCGCCGTCCAAGCAGTACGACGAGGGCGGGCAGCTCCAGCTCATGGAGGCC
GACCGGGTGGAGGAGGAGGAGGAGTGCTTCGAGTCCATCGACAAGTTGATCTCG
CAGGGAATAAACTCAGGAGACGTGAAGAAGCTGCAGGATGCGGGGATCTACACT
TGCAATGGGCTGATGATGCACACCAAGAAGAGCCTTACAGGGATTAAGGGCTTG
TCTGAAGCGAAGGTTGATAAGATCTGTGAGGCTGCTGAAAAACTTCTGAGTCAGG
GTTTCATGACAGGAAGTGATCTCCTTATTAAGCGAAAGTCTGTTGTCCGGATTAC
CACTGGGAGCCAAACGCTTGATGAGCTGCTTGGAGGAGGGATTGAAACACTCTGT
ATCACAGAGGCATTTGGAGAGTTCCGGTCAGGGAAGACCCAGTTGGCTCATACTC
TTTGTGTCTCCACTCAGCTTCCACTCCACATGCATGGTGGGAACGGGAAGGTTGC
CTACATTGACACTGAGGGAACATTCCGGCCTGAACGCATTGTGCCAATTGCTGAG
AGATTTGGGATGGATGCCAATGCTGTTCTTGACAATATCATATATGCTCGTGCCTA
CACCTATGAGCACCAGTACAACTTACTCCTGGGCCTTGCTGCCAAGATGGCTGAA
GAGCCTTTCAGGCTTCTGATCGTGGATTCTGTGATTGCGCTGTTCCGTGTTGATTT
CAGTGGTAGGGGTGAACTTGCAGAGCGTCAGCAAAAGCTGGCACAAATGCTGTC
CCGCCTTACAAAGATTGCTGAGGAGTTCAATGTTGCAGTGTACATCACCAACCAA
GTGATTGCGGACCCAGGTGGTGGTATGTTCATCACTGACCCCAAAAAGCCGGCGG
GAGGCCACGTGCTGGCGCATGCAGCCACCATCCGGTTGATGCTGAGGAAAGGCA
AAGGCGAGCAGCGTGTCTGCAAGATCTTTGACGCCCCTAACCTTCCCGAGGGAGA
AGCTGTTTTCCAGATCACAACAGGCGGATTGATGGATGTGAAAGAC

TaHOP2

ATGCCGCCTAAATCGGACAGCGTCGAAGGAATCGTCCTCAACTTCGTCAACGAGC
AAAACAGGCCATTGAATTCTCAGAATGTGGCTGATGCACTACAAAAGTTTAGTCT
TAAGAAGACTGCAGTGCAAAAAGGACTGGATGCATTGGCTGACAGTGGTCAAAT
ATCCTTCAAGGAGTATGGCAAGCAGAAAATCTACCTTGCTCGGCAAGACCAGTTT
GACATTCCAAATGGGGAGGAACTTGAGGAGATGAAGAAAGCAAATAGCAAGCTA
CAGGAAGAACTCGCAGAACAAAAGAAAATTACTAGTGAGGTCGAATCTGAGATA
AAAAGTCTACAGTCAAACTTGACACTGGAAGAGATTAGGTCAAAGGAGGCCAGA
TTACAAAGTGAGGTCCAGGAAATGGAAGAAAAATTAAAGAAATTGCAGAGTGGT
GTGATCTTGGTGAAACCTGAAGACAAGAAGATAATTGAAGAGACCTTCTCTGAG
AAAGCTAACCAGTGGAGAAAGCGAAAGAGGATGTTCAAGGAGCTATGGGATAAT
ATCACTGAGAATAGCCCCAAAGACCAGAAGGAATTTAAGGAAGAACTTGGTCTT
GAGTATGATGAAGATGTGGGTGTGAACTTCCAACCGTACAGTGAAATGCTGGCG
AGTCTCAACAAAAGGAGAAAAGTTTCTCGC

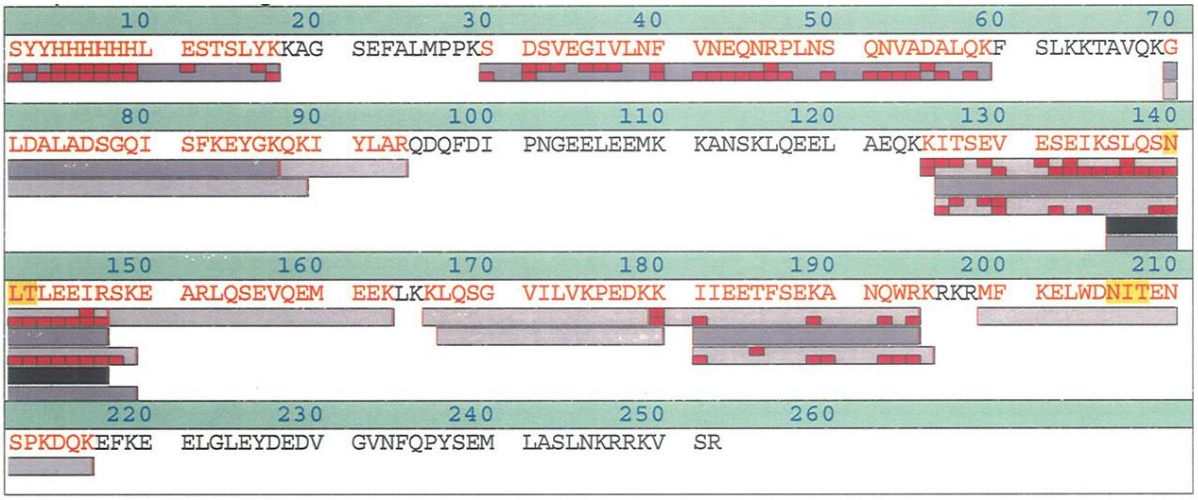
TaMND1

ATGTCGAAGAAGAGGGGCCTTTCCTTGGAGGAGAAGTGGGAGCAAATGCTTCAA
ATATTTTACGAGAGTCAAGACTTCTATCTGCTTAAAGAGCTTGAGAAGATGGGTC
CTAAAAAAGGAGTGATCAGCCAGTCTGTTAAGGATGTTGTGCAAAGCCTGGTGG
ATGATGATCTTGTCTTGAGAGACAAAATAGGAACTTCTGTGTACTTTTGGAGTCTT
CCCAGTTGTGCCGGAAATCAGCTGAGGACTACCTACAACAAGCTGGAGTCTGATC
TTTCAAACTCTAAAAAACGTTACATGGAGCTTCTTGAGCAGAGAGATGACTTGAA
AAGAGGCAGGGAAGACACTGATGAGAGAGAAGACGCTTTGGAGGAGCTGAAGG
CTGTAGAGCTACGCCATAAGAAGTTAAAGGAAGAACTAGCTGCCTATGCTGATA
GTGATCCATCTGCACTAGAGGCGATGAAGGATGCTACTGAGGTTGCCCATTCGGC
AGCTAGCAGATGGACAGACAACATCTTCACTTTGCAACAATGGTGTTCAACTACA
TTCCCGCAAGCAAAAGAACAGCTTGAACACATGTACAAGGAGGTGGGCATAACT
GAAGACTTTAAGTATCTGCAG

Appendix 3.5 – Protein confirmation of *Ta*HOP2

The peptide mass fingerprint is shown with the regions covered by the MS data indicated with the grey bars. Colour intensity corresponds to relative intensity. Red boxes show the amino acids sequenced by MS/MS analysis of the parent ion, where the upper series of boxes represent the *b*-ions, and the lower the *y*-ions.

Intensity Coverage: 33.4%
Sequence Coverage MS: 63.1%
Sequence Coverage MS/MS: 34.9%



Appendix 4.1 – Sequence of *TaDMC1* and *TaASY1* genes and the *TaHORMA* domain used for yeast two-hybrid analysis

The sequences below correspond to the *TaDMC1* and *TaASY1* genes and the *TaHORMA* domain isolated for creating the yeast two-hybrid bait vectors.

TaDMC1

ATGGCGCCGTCCAAGCAGTACGACGAGGGCGGGCAGCTCCAGCTCATGGAGGCCGACCGGGT
GGAGGAGGAGGAGGAGTGCTTCGAGTCCATCGACAAGTTGATCTCGCAGGGAATAAACTCAG
GAGACGTGAAGAAGCTGCAGGATGCGGGGATCTACACTTGCAATGGGCTGATGATGCACACC
AAGAAGAGCCTTACAGGGATTAAGGGCTTGTCTGAAGCGAAGGTTGATAAGATCTGTGAGGC
TGCTGAAAAACTTCTGAGTCAGGGTTTCATGACAGGAAGTGATCTCCTTATTAAGCGAAAGT
CTGTTGTCCGGATTACCACTGGGAGCCAAACGCTTGATGAGCTGCTTGGAGGAGGGATTGAA
ACACTCTGTATCACAGAGGCATTTGGAGAGTTCCGGTCAGGGAAGACCCAGTTGGCTCATAC
TCTTTGTGTCTCCACTCAGCTTCCACTCCACATGCATGGTGGGAACGGGAAGGTTGCCTACA
TTGACACTGAGGGAACATTCCGGCCTGAACGCATTGTGCCAATTGCTGAGAGATTTGGGATG
GATGCCAATGCTGTTCTTGACAATATCATATATGCTCGTGCCTACACCTATGAGCACCAGTA
CAACTTACTCCTGGGCCTTGCTGCCAAGATGGCTGAAGAGCCTTTCAGGCTTCTGATCGTGG
ATTCTGTGATTGCGCTGTTCCGTGTTGATTTCAGTGGTAGGGGTGAACTTGCAGAGCGTCAG
CAAAAGCTGGCACAAATGCTGTCCCGCCTTACAAAGATTGCTGAGGAGTTCAATGTTGCAGT
GTACATCACCAACCAAGTGATTGCGGACCCAGGTGGTGGTATGTTCATCACTGACCCCAAAA
AGCCGGCGGGAGGCCACGTGCTGGCGCATGCAGCCACCATCCGGTTGATGCTGAGGAAAGGC
AAAGGCGAGCAGCGTGTCTGCAAGATCTTTGACGCCCCTAACCTTCCCGAGGGAGAAGCTGT
TTTCCAGATCACAACAGGCGGATTGATGGATGTGAAAGACTGA

TaASY1

ATGGTGATGGCTCAGAAGACGAAGGAGGCGGAGATCACGGAGCAGGACTCGCTGCTTCTAACA
AGGAATTTGCTCCGGATTGCTATATACAACATCAGCTACATCAGAGGCCTATTCCCTGAGAAG
TACTTCAATGATAAGTCTGTTCCGGCACTAGAGATGAAGATTAAGAAGCTGATGCCCATGGAT
GCTGAATCCAGGAGGTTGATTGATTGGATGGAGAAAGGTGTCTATGATGCCTTACAAAAGAAA
TATCTCAAGACCCTTCTCTTCTGTATATGTGAGAAGGAGGAAGGCCCAATGATTGAAGAGTAT
GCCTTCTCATTTAGCTACCCCAACGCAAACGGGGAGGAAGTTGCAATGAACATGAGTCGCACA
GGGAGCAAAAAGAATAGTGCCACATTCAAGTCAAATGCAGCAGAAGTCACTCCTGATCAGATG
AGGAGCTCTGCTTGTAAGATGATCAGAACGCTAGTTTCACTTATGAGGACCTTGGACCAAATG
CCAGAGGAGCGAACCATTCTAATGAAGCTGCTATACTATGATGATGCTACACCTGAGGATTAC
GAGCCTCCCTTCTTTAAGGGTTGTGCTGAGAATGAGGCCGTAAATATATGGAACAAGAACCCC
TTGAAGATGGAAGTGGGGAATGTCAATAGCAAGCATCTTGTGTTAGCTTTGAAGGTTAAGAGT
GTCCTTGATCCATGTGATGCTAATGATGCTAACAGTGATGATGACAAGATGAGCGTGGGTCAT
GAGTCAGACCAAGATGACTTTACGGACACCGAGGTTCACCCATCTGAAGTGGATCGTTACGTC
ATTGCTCCTAATGATGGAAATGGCAAAGGTCAAAGTGGTACAAACTCAGATGATGAAACTCAA

GATGCTGCTCATGAGGAAGAGCTAACAGCTCAAGTAAGAGCATGGGTATGCTCAAGAAACATG
GGTACTGTTAATGCTTCAGATGTCCTTTCCAACTACCCTGACATATCATTGGAAATGGTGGAA
GATATTTTGGAGAGGCTACTTAAAGATGGTTTACTTTCGAGGGCAGGCAAGGATGGTTATGCT
GTTAACAAGATTACTGATCCCAAAACACCCTACATAAAGGAAGAGGTTGCCATGCACAATGTT
TCACCTACTGAAGGAACCAAGAACAACAGTGGAGATCTGATGTATATGAAGGCATTATACCAC
GCACTTCCAATGGATTATGTGACTATAGCTAAGCTTCAGGGCAAGCTTGATGGCGAAGCCAAC
CAGAGCACAGTCCGAAAGTTGATGGACAAAATGGTGCAAGATGGATACATTAAGAATTCAGGC
AACAGAAGATTAGGCAAAGCTGTCATTCACTCTGAAGTCACCAACAGAAAGCTCCTTGAGATA
AAGAAGATACTGGAAGTTGATATCACTGAAGACATGGCAATTGATACCAACGCAAGGCCTGCT
GAGTTTGATCGCAGAGATCACCAAACGGCTGACCAGGAAATGAAAGATGGCTCGACAAACGGC
CGCTTCCAGTCAGTTGGATCTGATCTTACCCGCACACGGGAGCTGCCGGAGCAGCAGCAGACT
AACAAGGACCCAAGCAGGACTCCCACAAGCAATCGCGAGTCGGCCACGTCCCTGGAGAGTGGG
GTGCTCGGGCAGAGGATCAGGAAGTCTCTGGCTGGCGAAGAGTCGATGTGCATGCCGGACAAG
CGGACCAGGAAGGCCAGCATGGTGAAGGAGCCGATCCTCCAGCAGGTCAAGCGCCAGAAGTCC

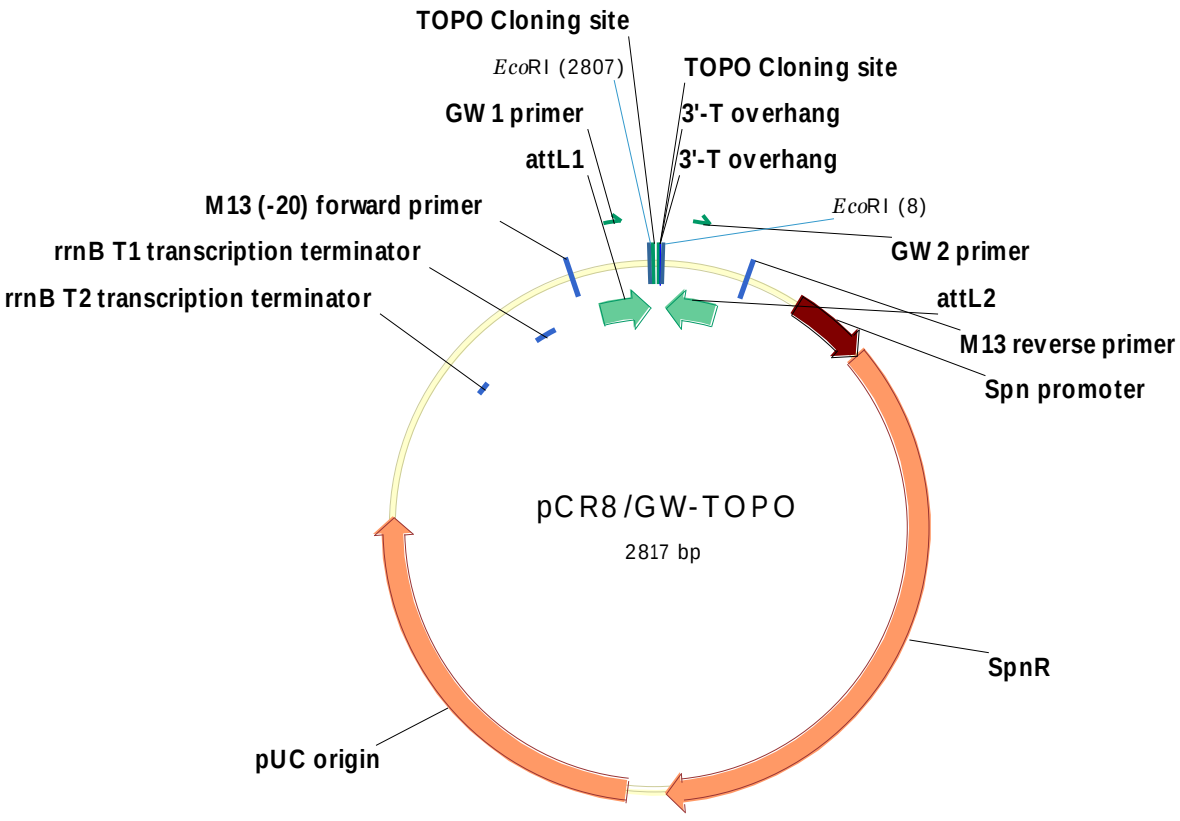
TaHORMA

GGAGATCACGGAGCAGGACTCGCTGCTTCTAACAAGGAATTTGCTCCGGATTGCTATATACA
ACATCAGCTACATCAGAGGCCTATTCCCTGAGAAGTACTTCAATGATAAGTCTGTTCCGGCA
CTAGAGATGAAGATTAAGAAGCTGATGCCCATGGATGCTGAATCCAGGAGGTTGATTGATTG
GATGGAGAAAGGTGTCTATGATGCCTTACAAAAGAAATATCTCAAGACCCTTCTCTTCTGTA
TATGTGAGAAGGAGGAAGGCCCAATGATTGAAGAGTATGCCTTCTCATTTAGCTACCCCAAC
GCAAACGGGGAGGAAGTTGCAATGAACATGAGTCGCACAGGGAGCAAAAAGAATAGTGCCAC
ATTCAAGTCAAATGCAGCAGAAGTCACTCCTGATCAGATGAGGAGCTCTGCTTGTAAGATGA
TCAGAACGCTAGTTTCACTTATGAGGACCTTGGACCAAATGCCAGAGGAGCGAACCATTCTA
ATGAAGCTGCTATACTATGATGATGCTACACCTGAGGATTACGAGCCTCCCTTCTTTAAGGG
TTGTGCTGAGAATGAGGCCGTAAATATATGGAACAAGAACCCCTTGAAGATGGAAGTGGGGA
ATGTCAATAGCAAGCATCTTGTGTTAGCTTTGAAGGTTAAGAGTGTCCTTGATCCATGTGAT
GCTAATGATGCTAACAG

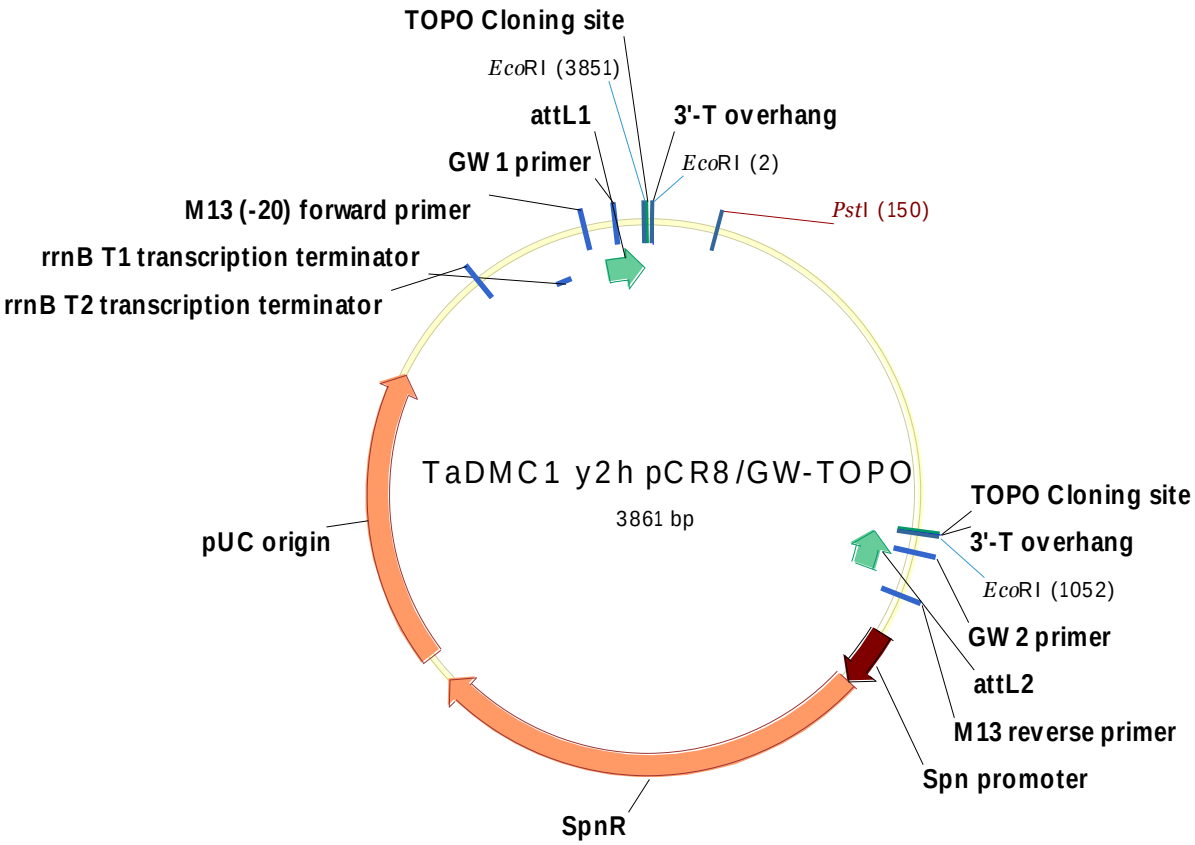
Appendix 4.2 – Vectors for the isolation of *TaDMC1* and *TaASY1* genes and the *TaHORMA* domain

The pCR8®/GW/TOPO® vector (*i*) (Invitrogen, AUS) was used as an intermediate vector for cloning the open reading frames of *TaDMC1* (*ii*) for the future placement of this fragment into the pGBKT7 bait vector. *TaASY1* and the *TaHORMA* fragments were placed into the pGEM®-T Easy Vector (*iii*) (Invitrogen, AUS) for the formation of a *TaASY1*:pGEM®-T Easy Vector (*iv*) and a *TaHORMA*:pGEM®-T Easy Vector (*v*), allowing for subsequent sequence confirmation and cloning of the *TaASY1* and the *TaHORMA* fragment into the yeast two-hybrid bait vector.

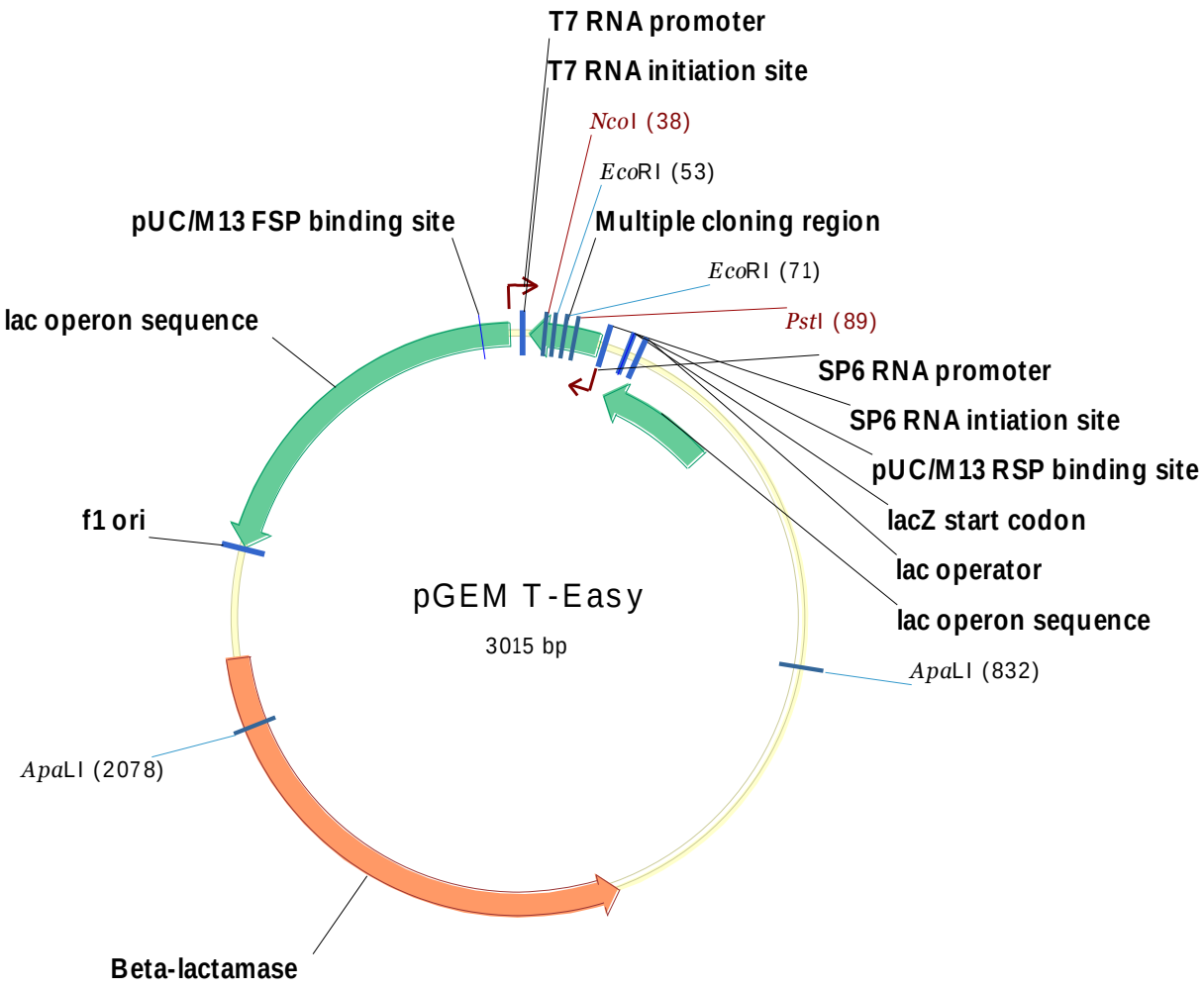
***i* - pCR8®/GW/TOPO® Vector**



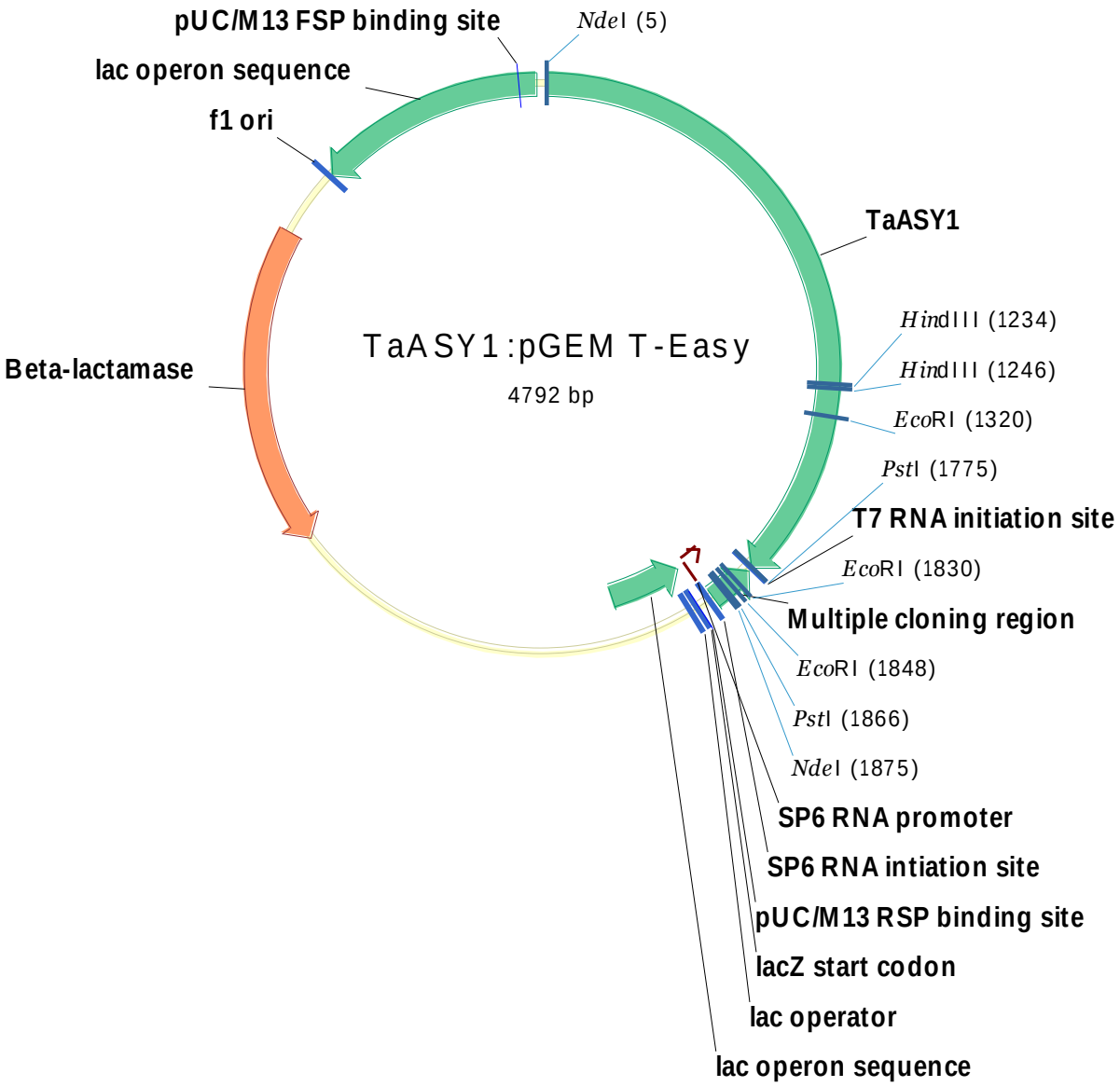
ii – TaDMC1:pCR8®/GW/TOPO® Vector



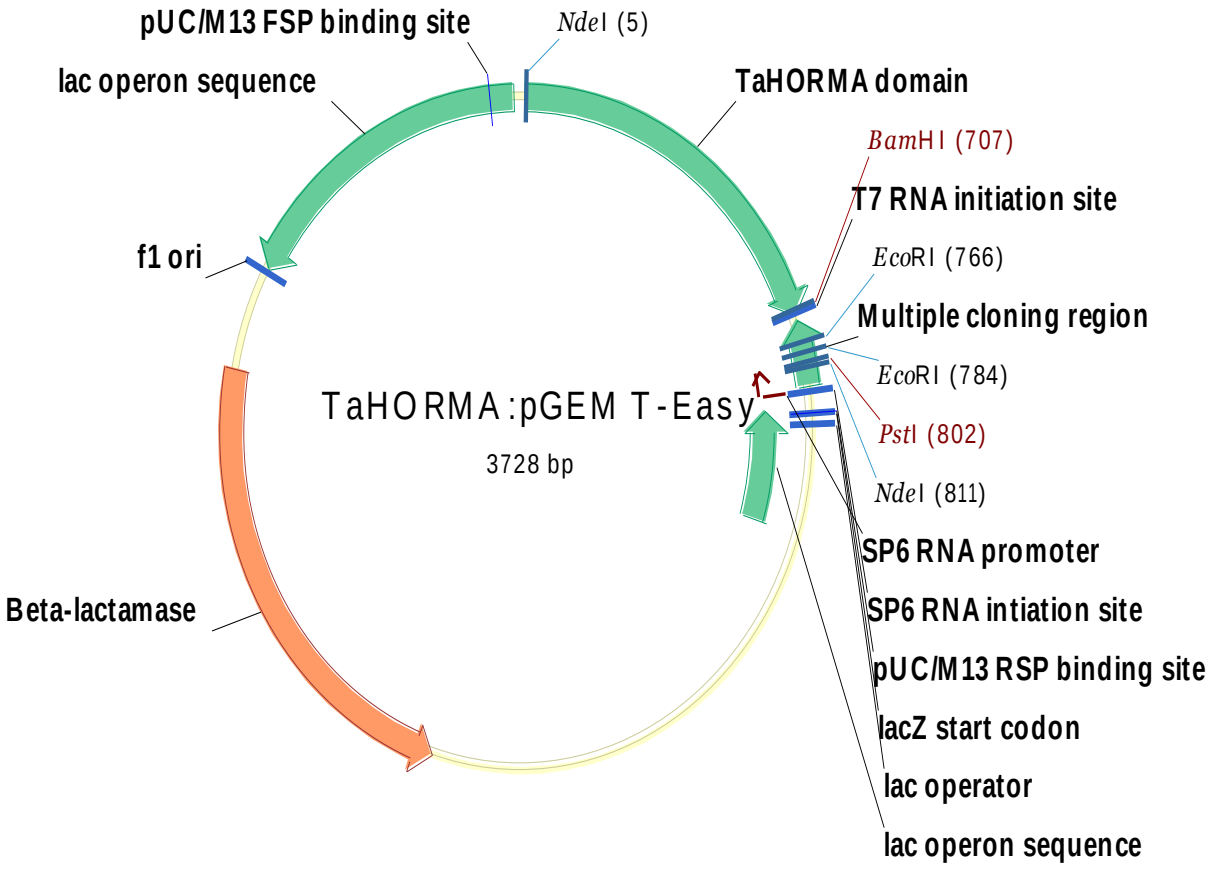
***iii* - pGEM®-T Easy Vector**



***iv – TaASY1*:pGEM®-T Easy Vector**

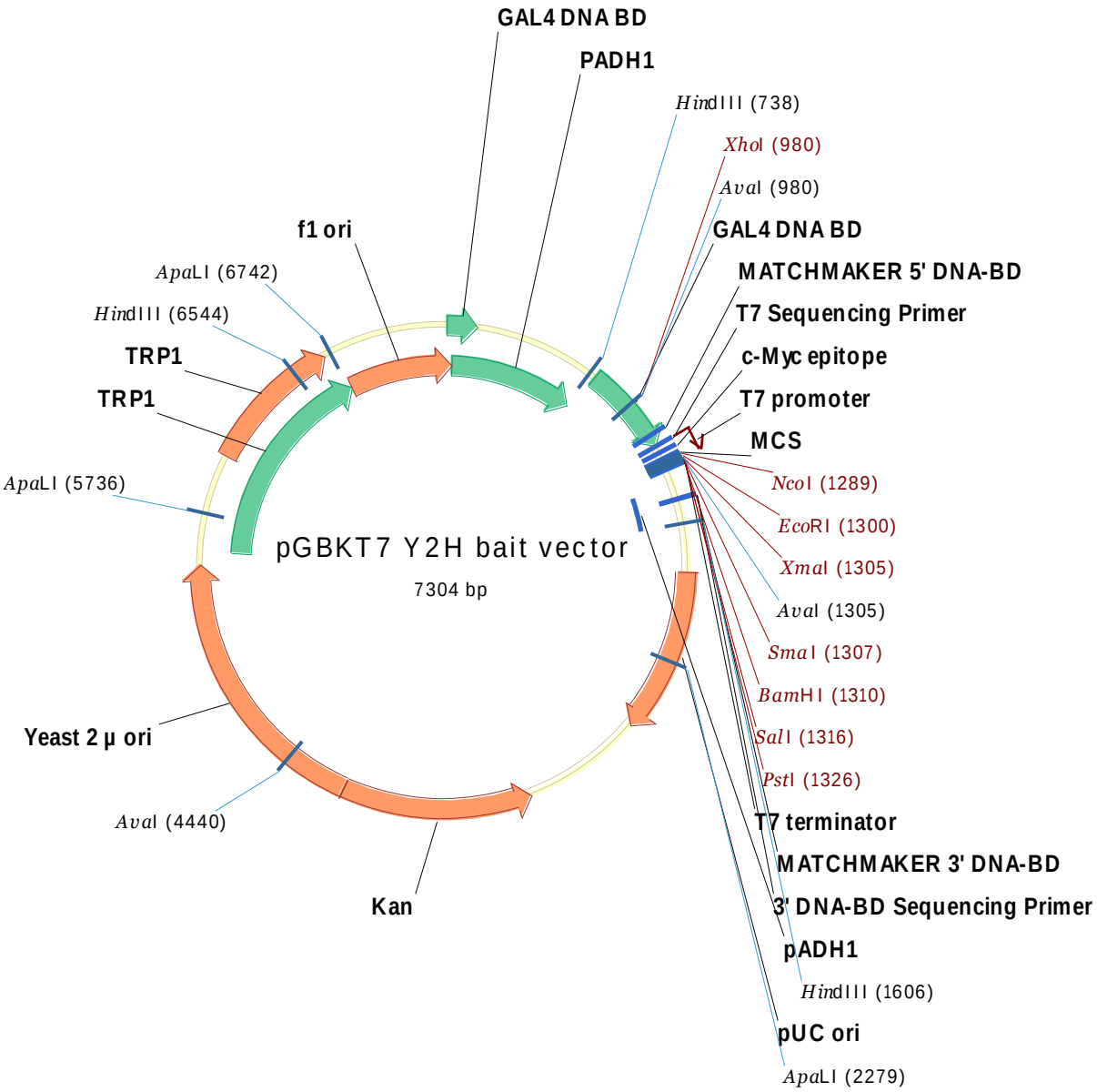


***v – TaHORMA*:pGEM®-T Easy Vector**



Appendix 4.3 – pGBKT7 bait vector for yeast two-hybrid analysis

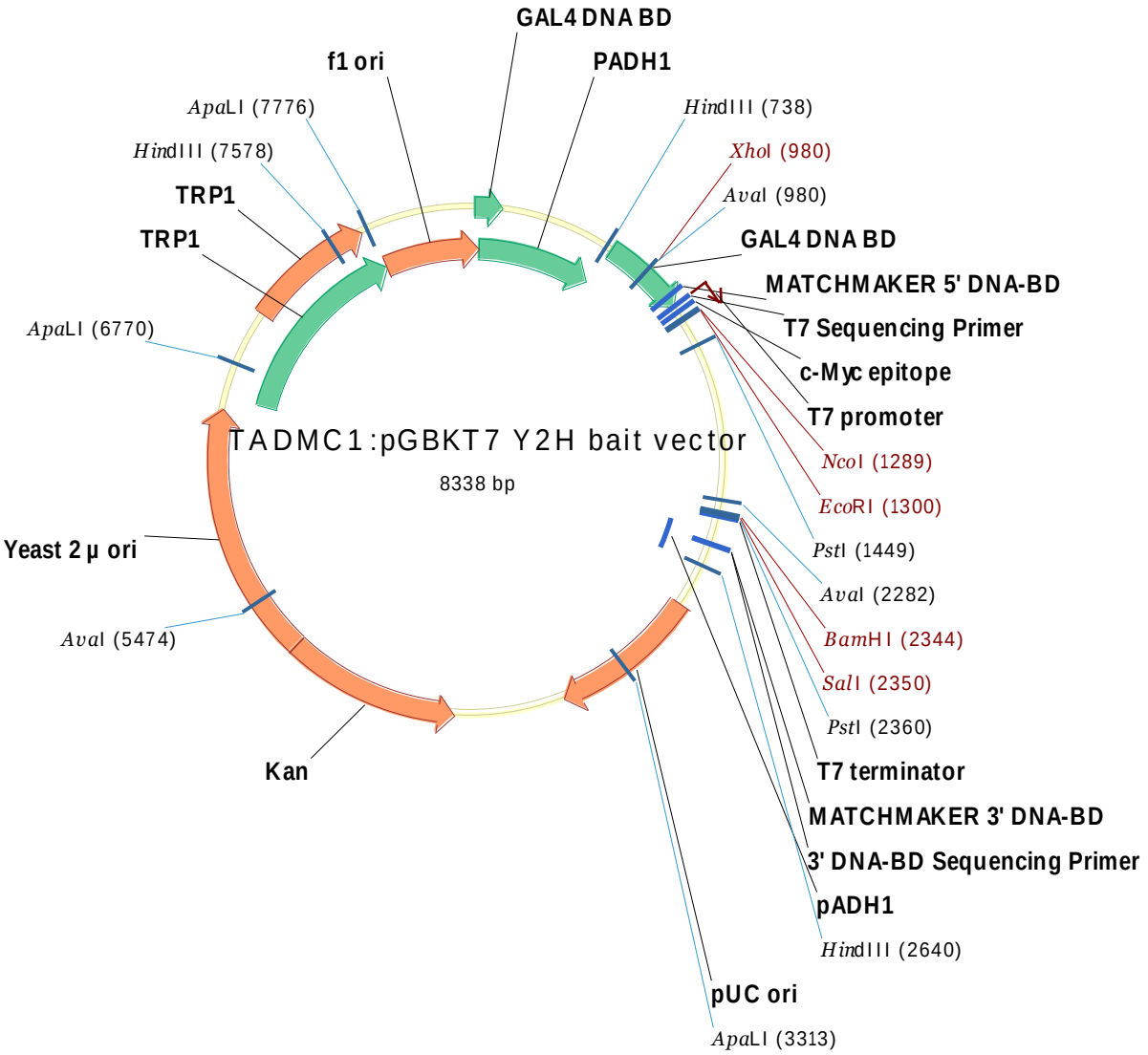
The pGBKT7 (Clontech, USA) bait vector was used in the yeast two-hybrid system.



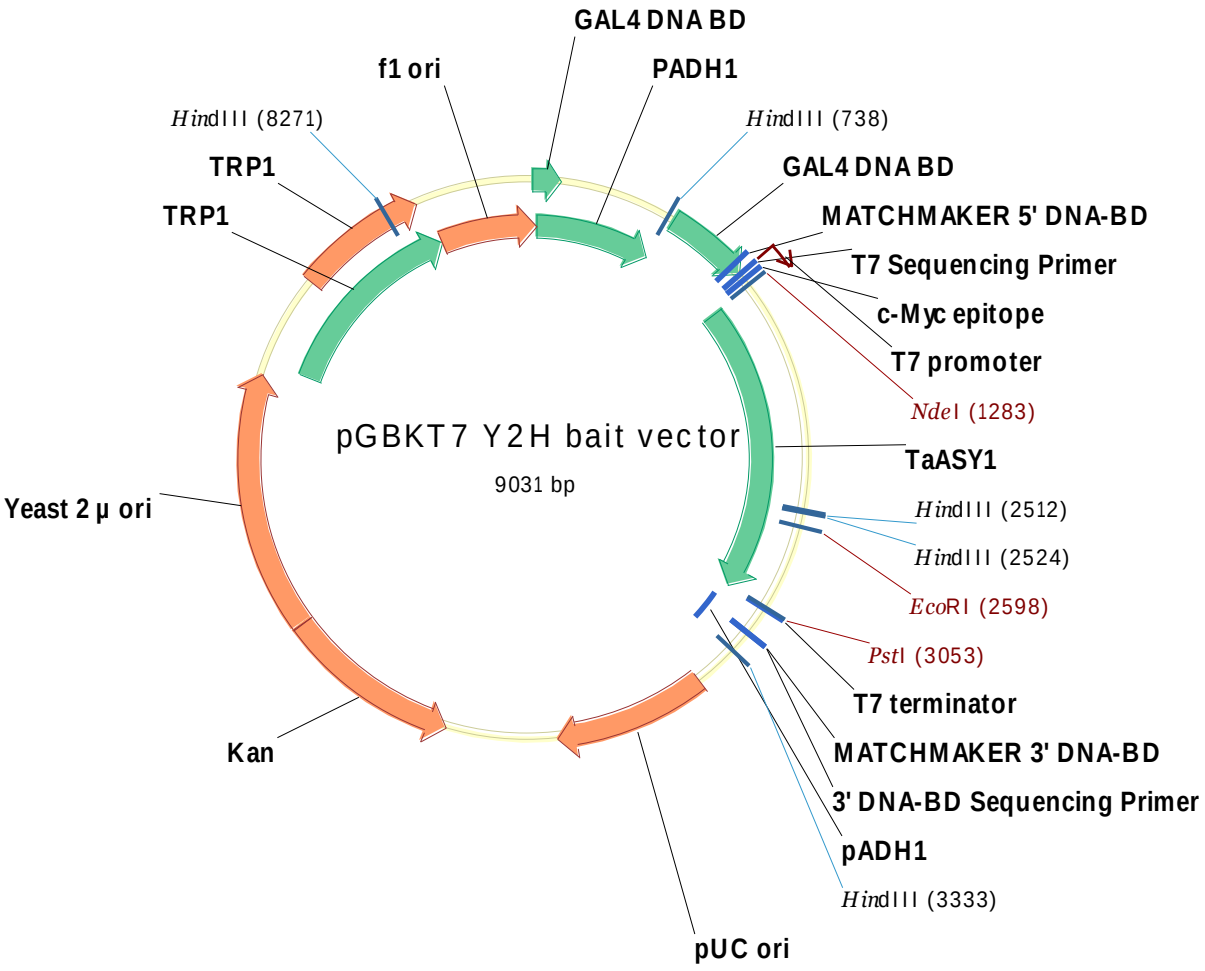
Appendix 4.4 - *TaDMC1, TaASY1 and TaHORMA* pGBKT7 bait vectors

TaDMC1:pGBKT7 (*i*), *TaASY1*:pGBKT7 (*ii*) and *TaHORMA*:pGBKT7 (*iii*) bait vectors were constructed and used in a yeast two-hybrid screen.

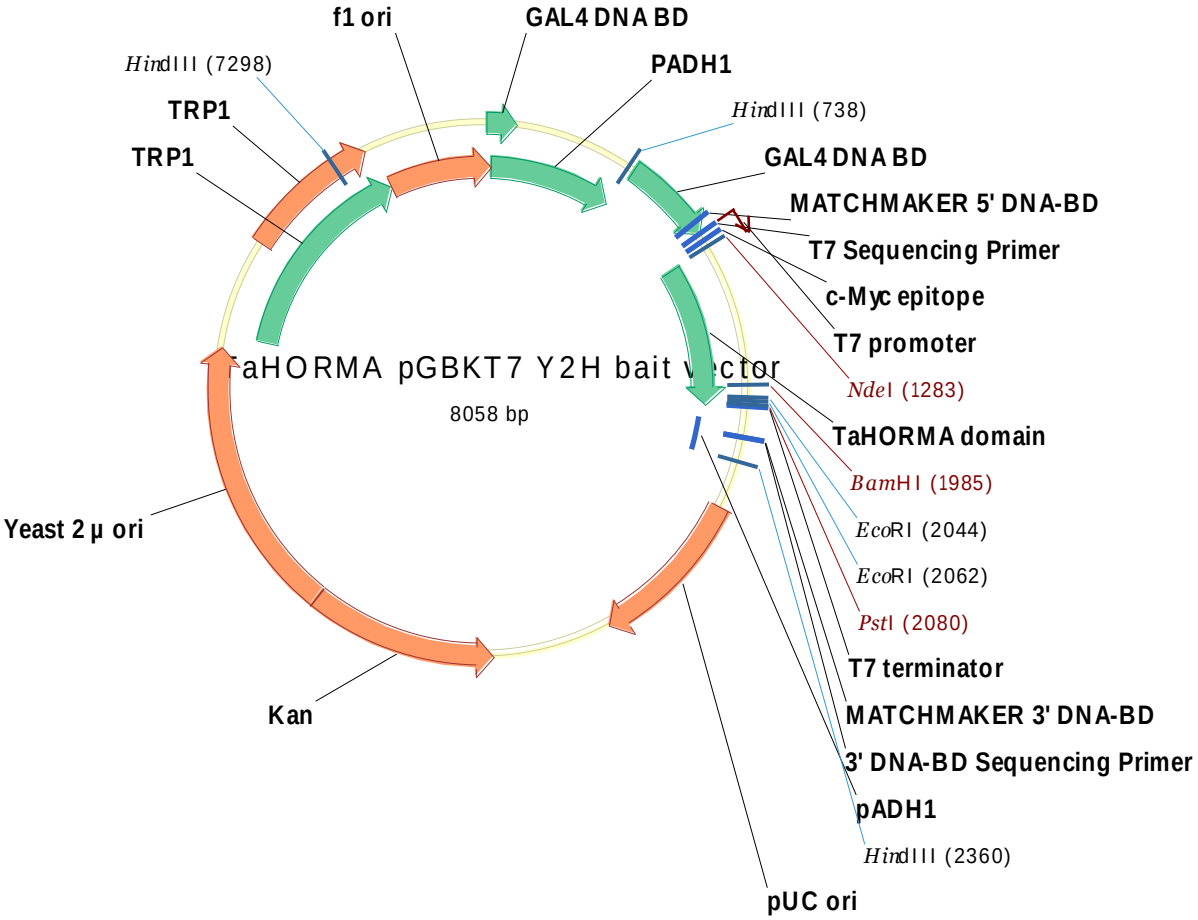
i- TaDMC1:pGBKT7



***ii- TaASY1*:pGBKT7**

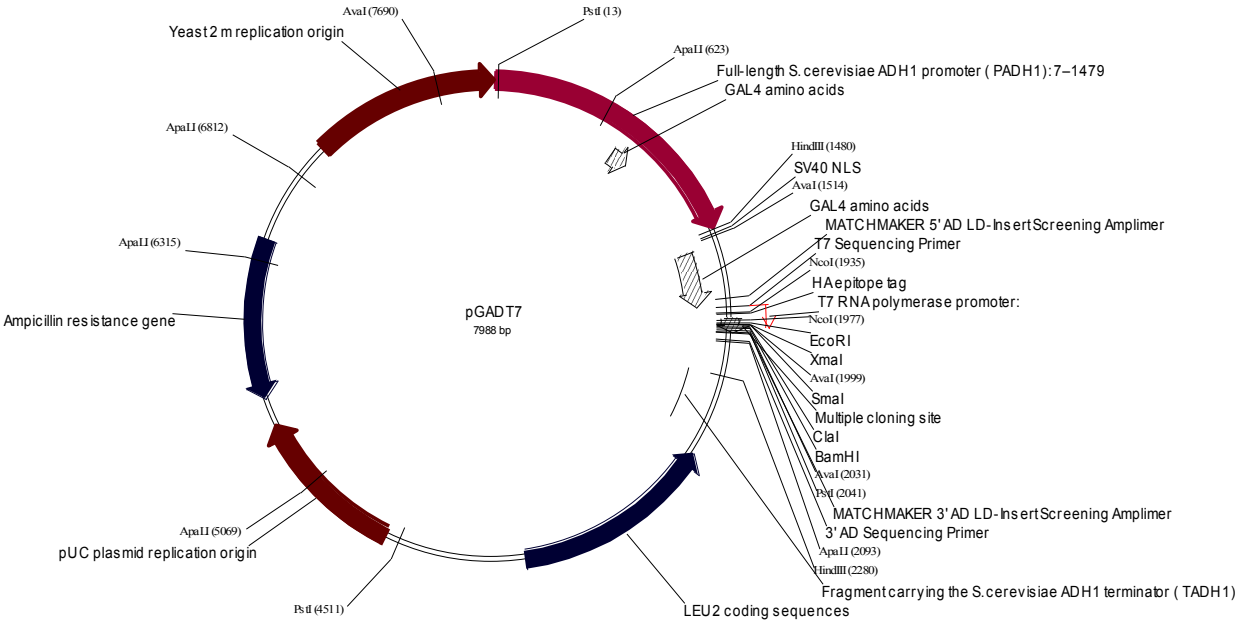


iii- TaHORMA:pGBKT7



Appendix 4.5 – pGADT7 prey/library vector for yeast two-hybrid analysis

pGADT7 prey/library vector into which the *TaHORMA* interacting fragments were re-ligated, for the formation of the *TaHORMA* interacting prey vectors for yeast two-hybrid confirmation.



Appendix 4.6 – Sequences of *Ta*HORMA interactors

*Ta*HORMA interactor sequences obtained through sequence analysis.

Mixed #1

GGGCACGCCGGACAAGCGGACCAGGAAGACCAGCATGGTGAAGGAGCCGATCCTCCAGCAGG
TCAAGCGCCAGAAGTCCTAGTTCAGTTGGGGCGGCACGGAATCGCAAGGAGCCTCAGCGCAA
CTAACCTCGTAGGCCAGGTGAAGAAGCAGATACCCCCTAGTTCGGTACGTCAGTTCGTTCAG
GTGCTTGAAACTTGTGAAGATAGCGACAGAGCACCCCCCAGGCAGATGTAGATCCTTGTCCC
TAGCCGTATCCTTGGTAGCAAAACGTAAACGAGAGAGACTTGTCCCCCCTGAGACGTGTTGT
TGTTCATGATGTTTGTTTCACTGGAGACTGTACTGTGAGGCCACTTTTGCCTGAATGATGAA
ACGTAACAATGTTGTATACAATGATGATCATCGTTGGTTGCTTC

Mixed #2

CAAGGACAGGGACTCGTGGCCGGAGCTGCAGGACCACGATGCCAACAACAGCAGCAACGCCT
TNTCGGCCGCCAAGCTGTCCATCTCCATCCCGGTGACCAGCTCCGACTTCTCCACCACCGCC
GGCTCCCGCTCGCCCAACGGTATATACTCCCGGTGAAGGTGAACGTCGTCCGCCGGCCGGCC
TGATTTTTGCTGATTTTTGCCGAAGCTTTTTGGTTTGGCCCACGACGGACGTCCTCAAATCA
TCACAGATGAGTGAACCGACCGAATGCGTCTGTGAGCTGACCGCCGCCTGCTTGCTCGCTCG
TTTTGTATGGATGCTGCCTTAACTATCATCGCCGTGCAGCACGAACGAAGCACTACTTGTTT
ATTTTTCTTCCCGTGATTTCACAGCGCTTTGCTTAATTTTCTTTCTTTGTTTTGCCGTGTGT
CGGCCGGAACTGTACGAAGTTTTCTATAGCCTCGATGGTCCAACTGTTTCCCGCGCAAAAGA

Mixed #3

TCTAACATTACCCAAACTGATATTATCACAATGTTAGCCGGTGACAATGGACGGAAAGAGTA
ACCATCTAGGTACAAGCAGGGGCACCCCGCGGAAAAGGACATACCCCCTCCACCACTAGGGC
ATCCCCGCCGGCGGCGGCGCCAACAGGGAGCGCTATCCGTTATAGCAGGACAATGACAGGAA
AAAGGTAAAACAACATCACGATGTGAAGTGAAACTTGGGATAATTCTAGTGGAAATGGATAG
GCTCCTTATGACACCACCGCACACAGCTCAGGTCCGTTCTCGCACAGCAGCTCCCGACTTGT
CCGCCGGANCAGCCTTGACAAAATGCGTCTTCTTCACTCGTCATCGTCCTCGTCCTCATCGC
TGCCCTCATCGTCATCGTCGTCATTGATCTCCGACTTGGACTTGTCAAATTCCTCCTCATCT
TCCTCCTCTACCTCCTTGGCCGCAGCCTTCTTTGGGGCAGCAGCAGCAGCGCTCTCGCCCTT
GTTGTAGGCAGCAATGGCCTTGTTGTACTCGCCCTTGAGCTTGTTGGCCTTGGCCACATAGG
GGGCCTTTTCCGATTCGCTCAAGCTCTTCCACCTCTCACCGGCCGCTTTGCCGACGGCGGCG
ACGGACTTGTTCTTGGGGTTCTTCTGCTTGAACTCCTCGCGAAACACGCCCATGAAGACGAA
GAAGGCGGATGGCGCCCTCTTGGGCTTGTTGGGGTCCTTGCCGGCCTTGCCCTTCTTCCCCT
TGGCGGCCGGCTTCTCCGCCCCCTTGCTCTTCACCGCCAACTTGGCGTCGGCCCTGGCGGCG
CCCTTGGATTTGGCCCCCTTCATGGCTGGTTTCCTGGGGGGAGGGGGCGGGGGAGAGAGGGC
GCTCCTGGTCTGGGTTGGGGAAAGGAGGAGGAGACGCAGACAAGGCTGGGCTCC

Mixed #4

AGACCAGAGCGCGATCTGCTCGCTGGCCAAGACGGAGAGGGAGCCGCTGTCCTTCGGCGGCG
GCGGCGGCTTCGAGGACGACGAGTCGGCGGTGAAGCAGGAGAACCAGACGCTGCGGCCCTTC
TTCGACGAGTGGCCCAAGGACAGGGACTCGTGGCCGGAGCTGCAGGACCATGACTCCAACCA
CAACAGCAACGCCTTCTCGGCCACCAAGCTGTCCATCTCCATCCCGGTGACCAGCTCCGACT
TTTCCACCACCGCCGGCTCCCGCTCGCCCCCCGGTATATACTCCCGGTGAAGGTGAAGGGGG
TCGGCCGGTTTGATTTTTGCTGATTTGCCGAAGTCACGACGGGCGTCCTCAAATCATCACAG
ATGAGTGAACCGACCGAATGCGTGCTTGCTCATTTTGTATGGAATGCTGCCTTAACTATCAC
CGCCGTGCAGCACGAACAAGGCCCTACTTGTTTATTTTTTTTCCCGTGATTTCACAGCGCTT
TGCTTAATTTTCTTTTTTGGTTTTGCCGTGTGTCGGCCGGAACTGTACGAAGTTTTCTGTAG
CCTCGATGGTTTAACTGTTTCCCGCGCAATTACTGTTGA

Oligo-dT #5

GCAGCCTTCCAGTTCCAGGATTTAAATCCAAGCAACGATAATATAAACCTAAGGTTTGCTAT
GCATTGCTTACTAATAGTGATTACTACAAGTGAAGGCAGTAATAAGCGAGACACACACACAC
AACACAAACTGTCTTCATAGTGAGCTTGATCTCCCAAATGACTTCCTCTGCAGGAAAACTTC
CTTCACCTTGCCGTTCACCACAAGCTTGCACAGCCCATTGAACGCATCAGCATCAAAGTTGT
CATCAAAATCTTTGCAACTTTCAGCACTGATGGACTGGTCGTTGTCAGCAATCCCAATAACC
TCTAAATAATGGGATTCACGGCCTTCCGAGGCACCTCTGATGGTCAACTGATGTCCATCAGT
CGACTGCGCAACAAGAACTCCACCTTGAGTGTGTTGGACCTTAAGTACTGTACGCACTCTTT
GCCCAACAAACATCTTGAGAGTCTCTCCATTGACGATTGCTGAACGAAATGACGCATCCATG
TCCATGTCGAGAGAGGAGATGAGCCTCGATCAATCTATCTACTCCTGCACACCCCACCCCGG
ACGCAGACGGAGATGGAAGCCTTCGAAGTGGCGCAGCGGAGGACTGCCCGGCTCTCGGCCGC
GCTGGGGTTCTGCGCTTCTGCTGGATTCGGAACGGGTGAGCCGGGCGGCGCCC

Oligo-dT #6

GCAGAACGCCGCCGACAAGAAGGACGAAAAGGCCGTCAACGCTGCCCGCCGCTCCGGCGCCG
AGATCGACACCACCAAGAAGTACAACGCTGGAACGAACAAGGCTGCATCTAGCGGAACTTCC
CTCAACACCAAGCGGCTCGATTGCGACTCGAGAACCTTTCGTTCATCTCCAAAAATATCTTC
CTCCTCTTTCCTGCGCCTCCATTATCCCTTGCGTGGTTTTAGTCTGTGAACTCGAGAGACCG
AGCGAGGAAAGTTGCCGTGTGAGACCTGAGAGTTAGGAGATATGGCTGGGATTGGTCCTATC
AGGCAGGACTGGGAGCCGATAGTGGTGCGGAAGAAGGCGCAGAACGCCGCCGACAAGAAGGA
CGAAAAGGCCGTCAACGCTGCCCGCCGCTCCGGCGCCGAGATCGACACCACCAAGAAGCACA
ACGCTGGAACGAACAAGGCTGCATCTAGCGGAACTTCCCTCAACACCAAGCGGCTCGACGAC
GACACCGAGAACCTTTCCCATGAGCGTGTTTCAAGTGACCTCAAGAAAAACCTTATGCAAGC
AAGGCTGGATAAGAAGATGA

Oligo-dT #9a

CTACCCCCACCATCAACATACATGGGTTGCTGCCAAAACAGTTCTTCAAACTTGATACAACA
AACAGGCAAAGAATCGCAGGGCTGATGATTTCGTTACAGAGGGATCCACTGATTCTTTCCCT
TTTTTTTCCCTTTCGATAACACGCAGGAGGGATCCACCTGTTCTTTTCAGCCATGGTACAAA
ATAATTCCTTACTCCTGCTCTGGCTGCGCTGATGGGACGAGCGGCTGATGTTCATCAGCACC
ACTTGATATTCTGTCCATGCCACACAACTCGGGGGATTTTATAAGCACAAATGGTGATTTGT
GTAAAGCAGCGAAGGCCTTGTTTCCTTCACCGTTGCTCCCTAAACAAATGGTATCACCCCGC
CCGTGCCAGACACCACACTGGACATTATTCATCTCTAAAGTTCTCGCCTCCTGGCACAGGGT
ACAACAAGGACATATCANATGGAGAACACAGTCATCTAGACTGCCATCAGTCCCCCGGATAT
TGAACTGTTTCCTGATTCTTCTTCGGAAGTAACCGGTGTAGATTGCTATCAAGAGAACAGAA
CCAAGTCCCATATAGAGATATACATGATGCCTTGTTACGGAAAAGGCAACTAAGCTGACCAG
AACAGCAACTAATGAAATGAAGTAAGCCATTGCCTGAAGGAAGCAAGATCCGAGATTGGCTC
TCCGCGTGTTCTTGCCAAACCTATAGCATGGGCAACAGGCGGCCTCGAGCGCGATTCGCCGG
TCCTCCAAGCAATCAATCACGACGTCCCCCTCCCACATCCTCTGCACCCCTCCTCCGCCGGC
GCTCTCGGCGCCACCCGAGCACGCCGCCACCGCCTCCATGCCGGCGCCCTTCCTCCTCCCCG

Oligo-dT #9b

TTCGACGAGTGGCCCAAGGACAGGGACTCGTGGCCGGAGCTGCAGGACCATGACTCCAACCA
CAACAGCAACGCCTTCTCGGCCACCAAGCTGTCCATCTCCATCCCGGTGACCAGCTCCGACT
TTTCCACCACCGCCGGCTCCCGCTCGCCCCCCGGTATATACTCCCGGTGAAGGTGAAGGGGG
TCGGCCGGTTTGATTTTTGCTGATTTGCCGAAGTCACGACGGGCGTCCTCAAATCATCACAG
ATGAGTGAACCGACCGAATGCGTGCTTGCTCATTTTGTATGGAATGCTGCCTTAACTATCAC
CGCCGTGCAGCACGAACAAGGCCCTACTTGTTTATTTTTTTTCCCGTGATTTCACAGCGCTT
TGCTTAATTTTCTTTTTTGGTTTTGCCGTGTGTCGG

Appendix 4.7 – Open reading frames of *TaRPA3 (TaHIP3)* and *TaHIP4*

Open reading frames of *TaRPA3 (TaHIP3)* and *TaHIP4*.

TaRPA3 (TaHIP3)

ATGGATACTTCAGCTCCTTCGCCATTTGTCAATGGAGAGACTCTGAAGATGTTTCTTGGGCG
ACGAGTGCGCACAGTGGTTCAAGTCCAACACAATGAAGGTGGAGTTCTTCTCGGGCTGTCCA
CTGATGGGCATCAGTTGACTATCAGAGGTGCTTCTGGTGCCCCTGAACCACCACACTACATC
GAGGTTATTGGGATCGCTGACAGCAGCCTGTCCATCCGTGCTGAATCTTGCACTGATTTTGG
TGAAAACTTTGATGGTGTGGCATTCAACGGGCTATGCAAGCTTGCGAATGACAAGTACAATT
ACCTGTTCCTG

TaHIP4

ATGGCCGCGGCGGCGGAGGAGGCGTCCACGGCGCCGAGGAGCCTCCCCGAGGAGGAGGAGGA
GGCGATGGCGGTGCTGGACTTCGACATGCTCTGCGCGTCCGTCGCCATGTCGGCGGAGAGGA
GGAAGGGCGCCGGCATGGGGGAGGCGGTGGCGTGCGCGGGCTCGGGCGGCGGCGAGGGGGCC
GCCGGCGGAGGCGTGCAGAGGATGTGGGAGGGGGACGTGGTAATTGATTGCTTGGAGGACCG
GCGAATCGCGCTCGAGGCCGCCTGTTGCCCATGCTATAGGTTTGGCAAGAACATGCGGAGAG
CCAATCTCGGATCTTGTTTCCTTCAGGCAATGGCTTACTTCATTTCATTAGTTGCTGTTCTG
GTCAGCTTAATTGCCTTTTCTGTAACAAGGCATCATGTATATCTATATGTGGGACTTGGTTC
TGTTCTCTTGATAGCAATCTACACCGGTTACTTCCGAAGAAGAATCAGGAAACTGTTCAATA
TCCGGGGGACTGATGGCAGTCTAGATGACTGTGTTCTCCATCTGATATGTCCTTGTTGTACC
CTGTGCCAGGAGGCAAGAACTTTAGAGATGAATAATGTCCAGTGTGGTGTCTGGCACGGGCG
GGGTGATACCATTTGTTTAGGAAGCAATGGTGAAGGAAACAAGGCCTTCGCTGCTCTACACA
AATCACCATTTGTGCTTATAAAATCCCCCGAGTTGTGTGGCATGGACAGAATATCAAGTGGT
GCTGATGAACATCAGCCGCTCGTCCCATCAGCGCAACCAGAGCAGGAG

Appendix 5.1 – Oligonucleotides for genotyping Arabidopsis T-DNA plants

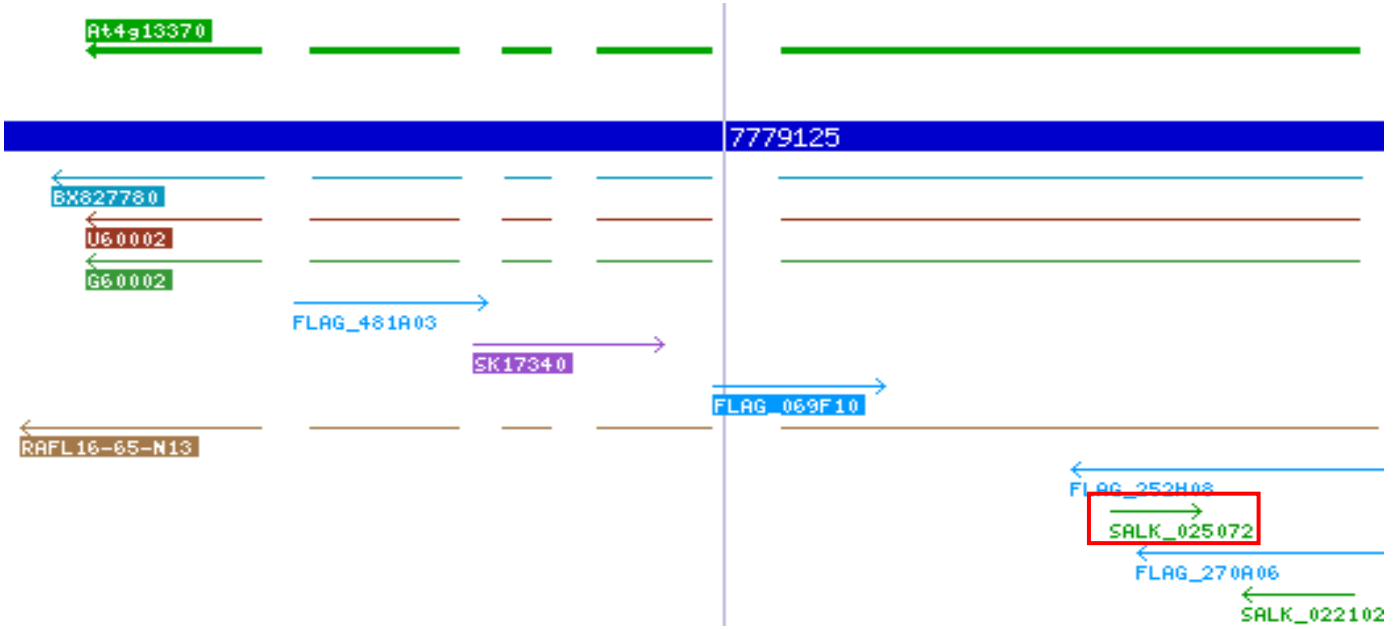
Oligonucleotides designated for each SALK line by T-DNA Express: T-DNA primer design (http://signal.salk.edu/tdnaprimers.2.html) from Salk Institute Genomic Analysis Laboratory (SIGnAL). For detecting presence of a T-DNA insert, RP oligonucleotides are used in conjunction with the LBb1.3 (Lb) insert oligonucleotide.

Novel	Arabidopsis Homologue	T-DNA SALK line	Oligonucleotide 5′-3′		Expected Size (bp) of LP+RP	Expected Sized (bp) Lb+RP
2	At4g13370	SALK_025072	LP	AGTACGCAAAATTGCTTGGTG	1005	492-792
			RP	TTACAAGTCACCGGAATCGTC		
9	At5g46880	SALK_032734	LP	TTCCAAAATAAAATTTCCCGC	1098	569-869
			RP	TAGCATAAGCAAGTGGCATTG		
11	At4g02800	SALK_108010	LP	TTCATGCTGGTGAAAATTTTG	1131	474-774
			RP	TTCGTGTTTCCTCTTCTTCCC		
18	At3g56870	SALK_058310	LP	CCATTCTCGAATCGATACAGC	990	438-738
			RP	ATGTCACTGGACGAATCAAGC		
24	At5g37630	SALK_018039	LP	ACGAAACGACGTTGTTTTGTC	989	436-736
			RP	TTATCTGACACGCCCTGAATC		
25	At3g57060	SALK_044796	LP	GTAGAGGTTCCACTTTTGCCC	1151	602-902
			RP	AACATAACCCCTTTGGTCCAC		
26	At5g02370	SALK_106055	LP	CGATGCAGGTCTGTTAAAAGC	1196	548-848
			RP	CTTAGCACTTTCTGCAATGGC		
29	At5g56740	SALK_051832	LP	CCGTATAAAACCAATCCAACG	1128	542-842
			RP	TTGAATTGGTCTTGCATTTGG		

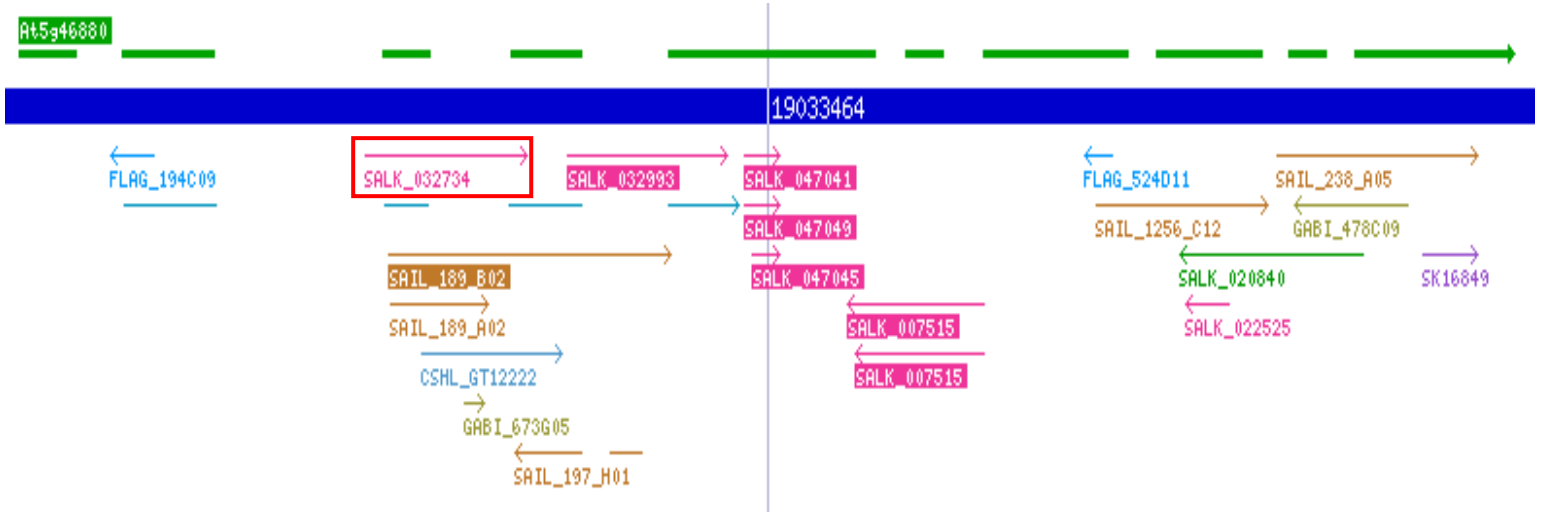
Appendix 5.2 – Identified T-DNA mutant lines for novel candidates

Identification of Arabidopsis homologue for wheat novel meiotic candidates using T-DNA Express: Arabidopsis Gene Mapping Tool (http://signal.salk.edu/cgi-bin/tdnaexpress) from Salk Institute Genomic Analysis Laboratory (SIGnAL). Mutant lines with T-DNA insertions which reside as close to or within the first exon were preferentially chosen (boxed) for increased likelihood of obtaining a knock-out of the gene of interest (Krysan *et al.*, 1999). Blue box represents the Arabidopsis chromosome, dark green arrow above the chromosome box represents the direction of the exons of the gene candidate and arrows below the chromosome box represent the orientation of the T-DNA insert within the chromosome, beginning from the left border of the insert and into the flanking plant DNA sequence. Figure derived from T-DNA express (**http://signal.salk.edu/cgi-bin/tdnaexpress**).

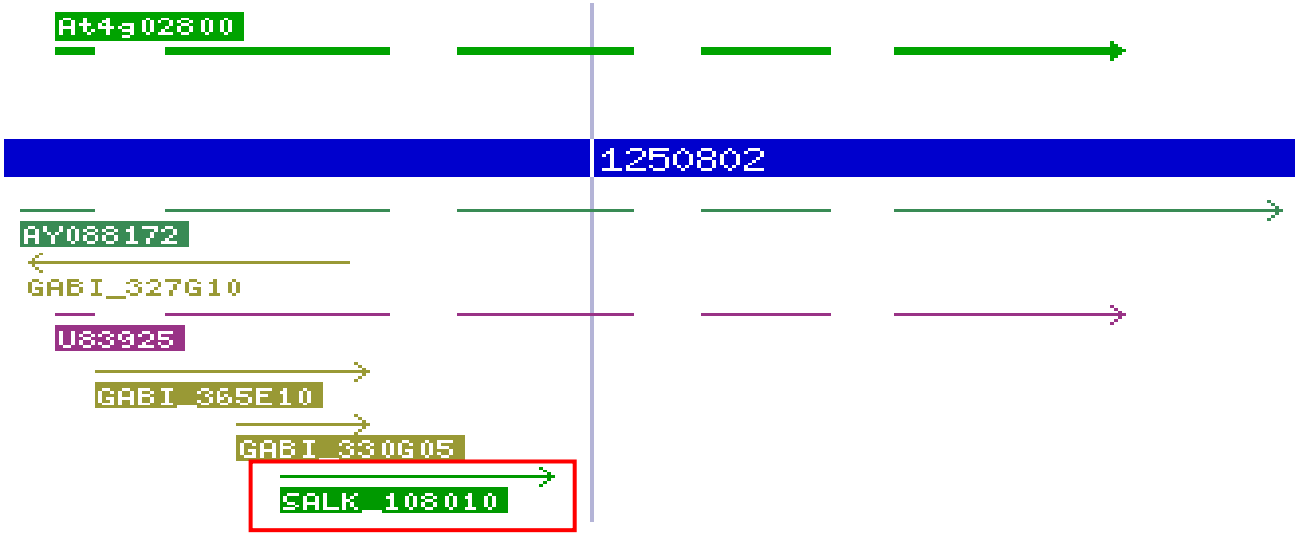
MN02 – DUF936 domain containing protein – At4g13370



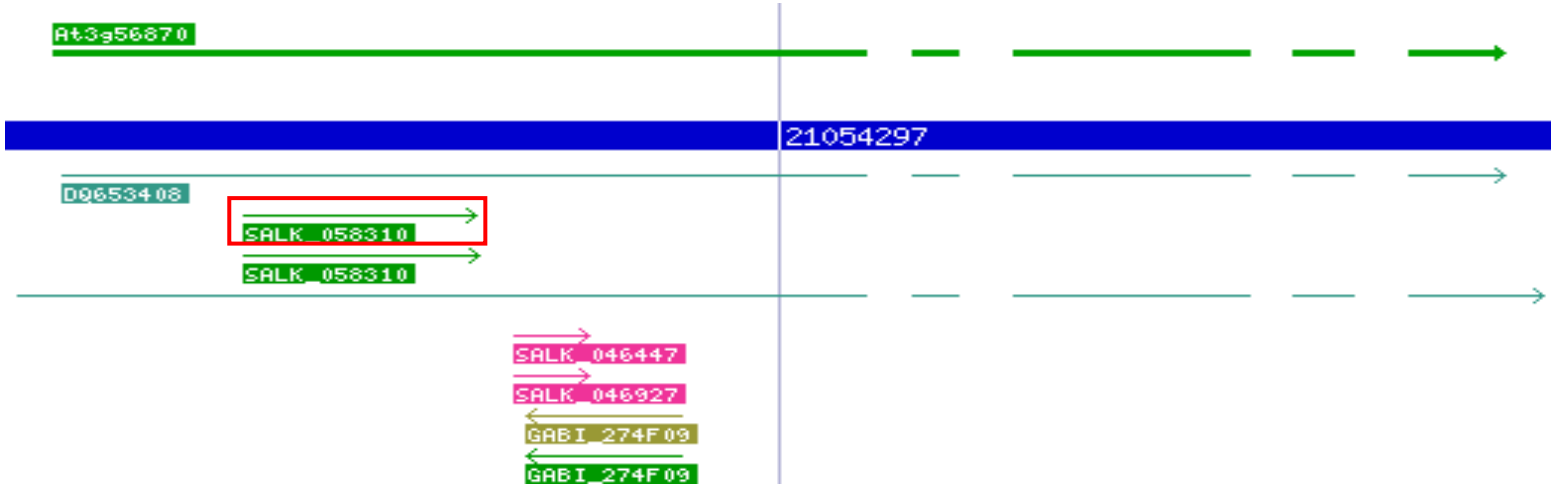
MN09 – Homeobox-7 Transcription Factor (HB7) – At5g46880



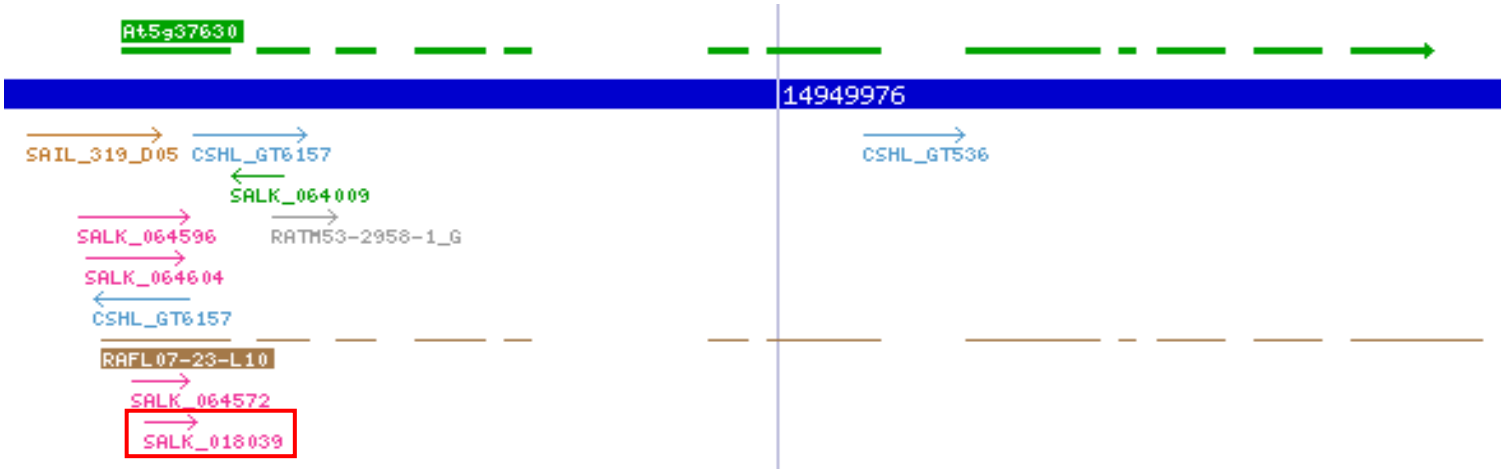
MN11 – Unknown Protein – At4g02800



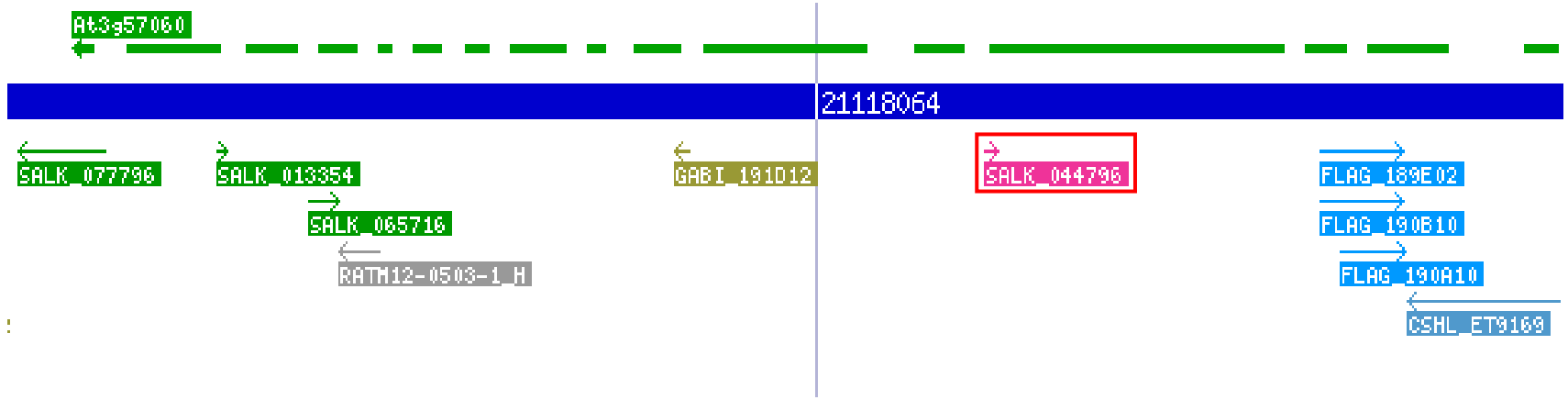
MN18 – Unknown Protein – At3g56870



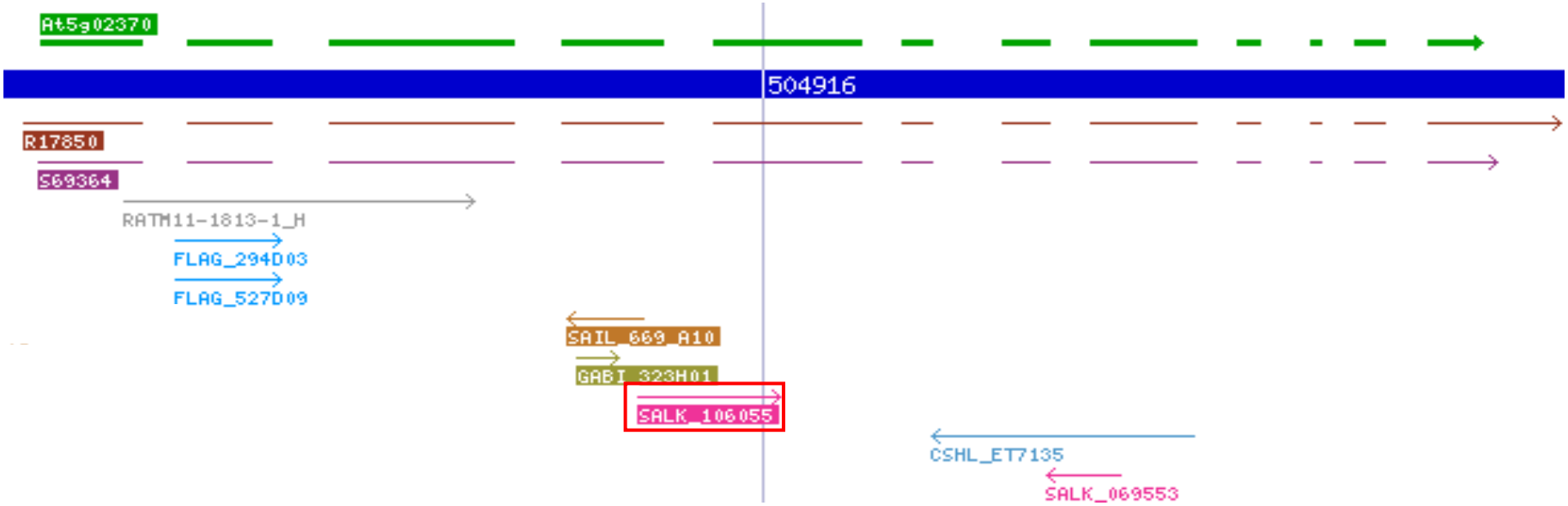
MN24 – Putative Chromosome Condensation Protein – At5g37630



MN25 – Putative Condensin subunit 1 Protein – At3g57060



MN26 – Putative Kinesin Motor Protein – At5g02370



MN29 – Putative Histone Acetyltransferase B (HATB) Protein – At5g56740