FROM INTIMATE CHROMOSOME ASSOCIATIONS TO WILD SEX IN WHEAT (*TRITICUM AESTIVUM*)

by

Ryan Whitford

B.Ag.Sc. Hons

University of Western Australia

Perth, Western Australia

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Plant Science

Waite Agricultural Research Institute

Adelaide University

2001

Date _____

Abstract

FROM INTIMATE CHROMOSOME ASSOCIATIONS TO WILD SEX IN WHEAT (*TRITICUM AESTIVUM*)

by Ryan Whitford

Meiosis occupies only a short period of the life cycle of higher plants but it is a crucial process ensuring the correct passage and maintenance of genetic information from parent to offspring. The ability of allohexaploid bread wheat (*Triticum aestivum* L.) to act as a diploid, although it consists of three closely related genomes, at meiosis makes it an ideal organism for the study of the mechanisms which ensure correct homologous chromosome pairing at early prophase of meiosis I. Wheat contains several diploidising systems, which act to restrict chromosome pairing to homologous chromosomes. One of these genes, and the one which this study addresses, is termed *Ph2* (*P*airing *H*omeologous), a dominant gene that is located on the short arm of chromosome 3D (3DS). *Ph2* is believed to be one of the principal genes responsible for the diploid like behaviour of *T. aestivum*, acting to prevent pairing between homeologous chromosomes of different genomes and also in inter-specific and inter-generic hybrids.

Several approaches have been taken in this study to isolate and analyse genes expressed at early meiosis which may play a role in homologous chromosome pairing.

The first approach was the structural analysis of *W*heat *M*eiosis 1 (*WM1*) gene family in relation to its localation within the region deleted in the *ph2a* wheat mutant. This work has resulted in the isolation of a cluster of nine, predominantly type Ia plasma membrane-anchored leucine rich repeat-like receptor genes that are predominantly expressed in floral tissue at early meiosis. Importantly, these genes have been identified, sequenced and located in the region of the *Ph2* gene.

The second approach in the systematic analysis of potential meiosis genes was an attempt at the functional analysis of the *Wheat Meiosis 5* (*WM5*) gene through biolistic transformation of wheat. Successful transformation of four wheat genotypes showed that antisense and overexpression of the *WM5* gene showed no significant observable phenotype in the T_0 and T_1 generations as compared to the controls. Analysis of GFP and GUS reporter genes across all 26 independent transgenic wheat lines suggested that the *WM5* gene is under complex transcriptional control and further research with respect to promoter analyses needs to be performed to elucidate this complex control.

The final approach undertaken in identifying the genes controlling chromosome pairing and recombination was comparative mapping among the grasses. This approach used the available rice genome sequence to identify those genes that may represent the *Ph2* gene. Comparative mapping of the rice region homologous to that deleted in the *ph2a* mutant highlighted two clear candidate genes that are likely to explain the *Ph2* phenotype. These genes encode a rice Asy1 (*Arabidopsis* asynaptic mutant 1) protein homologue and a mismatch repair protein (*Arabidopsis* AtMSH7) homologue.

Through the structural, functional and comparative analysis of genes within the grasses, this study aimed to increase our awareness of grass genome structure and draw attention to those genes likely to control the process of chromosome pairing and recombination during meiosis.

STATEMENT

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 contains no material previously published or written by any other person, except where due reference is made in the text.

I consent to this thesis being made available for photocopying and loan.

Ryan Mark Whitford

October 2001

TABLE OF CONTENTS

Statement iv

1 CHAPTER 5

Literature review 5
1.1 General introduction 5
1.2 Mitosis and meiosis 6
1.3 Model organisms for the study of meiosis 9
1.3.1 Chromosome pairing 9
1.3.2 Homologous recombination 10
1.3.2.1 Double strand break repair 11
1.3.2.2 Mismatch repair system 13
1.4 Meiosis in allohexaploid wheat 15
1.4.1 Genome organisation 15
1.4.2 Chromosome pairing and recombination 17
1.4.2.1 Suppressors of homeologous chromosome pairing 18
1.4.2.1.1 The mechanistic action of *Ph* genes 19
1.4.2.1.2 Pairing Homeologous gene 1 (*Ph1*) 20
1.4.2.1.3 Pairing Homeologous gene 2 (*Ph2*) 21
1.4.2.2 Promoters of homeologous chromosome pairing 22
1.4.2.3 Molecular models of chromosome pairing in wheat 23
1.4.2.3.1 Spatial distribution of chromosomes 23
1.4.2.3.2 Timing of chromosome interactions 24
1.4.2.3.3 Stringency of both synapsis and homologous exchanges. . 24
1.4.3 Improvement of wheat breeding by the control of the *Ph* genes ... 25
1.5 Aims of this study 26

2 CHAPTER 27

General Materials and Methods 27
2.1 Materials 27
2.1.1 Chemicals 27
2.1.2 Enzymes 28
2.1.3 Molecular weight markers and cloning vectors 28
2.1.4 Agaroses 29
2.1.5 Kits 29
2.2 Methods 29
2.2.1 Plant growth conditions 29
2.2.2 Growth of bacteria 29
2.2.3 Transformation of *E. coli* with plasmids by electroporation 30
2.2.4 Mini-preparation of plasmid DNA 31
2.2.4.1 PCR amplification of cloned insert DNA 31
2.2.4.2 Recovery of insert DNA from agarose gels using glass-milk and Qiagen gel extraction kit 32
2.2.5 Large scale preparations of plasmid DNA 32
2.2.6 Phenol:chloroform extraction and ethanol precipitation of DNA 33
2.2.7 Agarose gel electrophoresis 34

2.2.8 Polyacrylamide gel electrophoresis 34
2.2.9 Cereal genomic DNA preparation 35
2.2.9.1 Small-scale genomic DNA preparation 35
2.2.9.2 Medium scale genomic DNA preparation 36
2.2.10 DNA restriction, electrophoresis and Southern transfer 36
2.2.11 Preparation of ^{32}P-labelled DNA based probes 37
2.2.12 Hybridisation and autoradiography 37
2.2.13 Total plant RNA isolation 38
2.2.14 Northern blot hybridisation 38
2.2.15 Purification of DNA clone inserts 39

3 CHAPTER 41

Structural analysis of the *WM1* gene family 41
3.1 Introduction 41
3.2 Materials and methods. 42
3.2.1 Genetic stocks 42
3.2.2 DNA sequence and partial genomic clones of *WM1* genes 42
3.2.3 Amplification of *WM1.1*, *WM1.2* and *WM1.3* specific sequences 43
3.2.4 Nullisomic tetrasomic determined chromosome localisation 43
3.2.5 Pulse field gel electrophoresis 43
3.2.5.1 Protoplasts 45
3.2.5.2 Restriction enzyme digestion of HMW DNA 46
3.2.5.3 PFGE separation, transfer and hybridisation 46
3.2.6 BAC library screening 48
3.2.6.1 Screening of a barley and wheat bacterial artificial chromosome (BAC) library 48
3.2.6.1.1 Purification of DNA probe template 48
3.2.6.1.2 BAC filter preparation 48
3.2.6.1.3 Growth of BAC clones 49
3.2.6.1.4 Mini-preparation of BAC plasmid DNA 49
3.2.6.1.5 Confirmation of BAC clones 50
3.2.6.1.6 Southern fingerprint analysis of BAC clones 50
3.2.6.1.7 Preparation of ^{32}P-labelled total BAC DNA probes 51
3.2.6.1.8 Pulsed field gel electrophoresis 51
3.2.7 BAC fingerprinting 51
3.2.8 BAC contig assembly 52
3.2.9 BAC DNA sequencing 52
3.2.9.1 BAC subcloning 52
3.2.9.2 Removal of subcloned BAC vector 55
3.2.9.3 DNA sequencing 56
3.2.9.4 DNA sequence assembly 56
3.2.9.5 Sequence analysis 56
3.2.9.5.1 BLAST on non-redundant and EST databases 56
3.2.9.5.2 Gene prediction 56
3.2.9.5.2.1 Signal prediction 57
3.2.9.5.2.2 Promoter prediction 57
3.2.9.5.2.3 *cis*-acting regulatory element binding site prediction 57
3.2.9.5.3 Identification of repetitive elements and sequence charactersistics 57
3.2.9.5.4 Matrix attachment region prediction 58
3.2.10 Structural analysis of the WM1 predicted peptides 58

3.2.10.1 Transmembrane domains 58
3.2.10.2 Leucine zipper and leucine rich repeats 58
3.2.10.3 Signal peptides 58
3.2.10.4 Predicted subcellular localisation 59
3.2.11 Genes homologous to *WM1* genes 59
3.2.12 Northerns of *WM1* genes 59
3.2.13 RT-PCR of *WM1* genes 59
3.2.13.1 Gene specific primers 60
3.3 Results 60
3.3.1 Nullisomic tetrasomic determined chromosome localisation 60
3.3.2 Pulse field gel electrophoresis 63
3.3.3 BAC library screening 67
3.3.3.1 Southern fingerprint analysis of BAC clones 67
3.3.3.2 Pulsed field gel electrophoresis 71
3.3.4 BAC fingerprinting 71
3.3.5 BAC contig assembly 71
3.3.6 BAC sequence analysis 73
3.3.6.1 *WM1* genes 76
3.3.6.1.1 Prediction of transcription regulatory elements 76
3.3.6.1.2 Prediction of poly-adenylation signals 80
3.3.6.2 WM1 predicted primary peptide structure 80
3.3.6.2.1 Transmembrane domains 83
3.3.6.2.2 Leucine zippers 83
3.3.6.2.3 Leucine rich repeats 87
3.3.6.2.4 Signal Peptide 89
3.3.6.2.5 Predicted subcellular localisation 89
3.3.6.2.6 Structural conservation 91
3.3.6.3 Genes homologous to *WM1* genes 91
3.3.7 Northern analysis of *WM1* genes 94
3.3.8 RT-PCR of *WM1* genes 94
3.4 Discussion 97
3.4.1 *WM1* genes cluster 97
3.4.1.1 *WM1* gene structural organisation 98
3.4.1.2 *WM1* gene evolution 99
3.4.2 *WM1* gene analysis 100
3.4.2.1 Leucine rich repeats. 101
3.4.2.2 Leucine zipper 103
3.4.2.3 Sorting signals 103
3.4.3 Relating developmental regulation to disease resistance 104
3.4.4 A tentative model for developmental signal transduction 105
3.5 Conclusions 105

4 CHAPTER 108

***In planta* analysis of the meiotic gene *WM5* 108**
4.1 Introduction 108
4.2 Materials and methods. 111
4.2.1 Plant material and growth conditions 111
4.2.2 Gene constructs 112
4.2.3 Microprojectile bombardment 112
4.2.4 Culture and selection conditions 112
4.2.5 Antibiotic selection in glasshouse 118

4.2.6 Histochemical GUS staining 118
4.2.7 Microscopic detection of GFP fluorescence 120
4.2.7.1 *GFP* reporter gene in monocots 120
4.2.8 PCR analysis 120
4.2.9 Southern analysis 123
4.2.10 Northern analysis 123
4.3 Results 123
4.3.1 Plant regeneration and selection of transformants 123
4.3.2 T_0 and T_1 analysis 125
4.3.2.1 PCR analysis of T_0 and T_1 125
4.3.2.2 Southern analysis of T_0 and T_1 125
4.3.3 Transformation efficiency 128
4.3.4 Northern analysis 128
4.3.5 *WM5* promoter analysis 132
4.3.5.1 Analysis of the *GFP* reporter gene in monocots 132
4.3.5.2 *WM5* GFP reporter analysis 132
4.3.5.3 *WM5* GUS reporter analysis 134
4.3.6 *WM5* phenotypic analysis 134
4.4 Discussion 137
4.4.1 Plant regeneration and selection of transformants 137
4.4.2 T_0 and T_1 analysis 137
4.4.2.1 PCR analysis of T_0 and T_1 137
4.4.2.2 Southern analysis of T_0 and T_1 138
4.4.3 Transformation efficiency 138
4.4.4 Northern analysis 139
4.4.5 *WM5* promoter analysis 140
4.4.5.1 Analysis of the *GFP* reporter gene in monocots 140
4.4.5.2 *WM5* GFP reporter analysis 141
4.4.5.3 *WM5* GUS reporter analysis 142
4.4.6 *WM5* phenotypic analysis 143

5 CHAPTER 144

Structure of the *Ph2* locus 144
5.1 Introduction 144
5.2 Materials and methods. 145
5.2.1 Plant material 145
5.2.1.1 Genetic stocks 145
5.2.1.2 Mapping populations 146
5.2.2 RFLP analysis. 146
5.2.2.1 Southern blot analysis. 146
5.2.2.1.1 *WM1*, *WM3*, *WM5* and *TaMSH7* gene family RFLP mapping 146
5.2.2.1.2 RFLP probes for determining *ph2a* deletion size 147
5.2.2.2 Genetic analysis 147
5.2.3 Comparative mapping between barley, wheat and rice 147
5.2.3.1 YAC, PAC and BAC contig development 148
5.2.3.2 Identification of wheat ESTs from PAC and BAC clones 148
5.2.3.3 Structural analysis of putative meiosis genes 148
5.2.3.4 Analysis of rice and barley meiosis related phenotypic traits 148
5.3 Results 150
5.3.1 RFLP probes for determining *ph2a* deletion size 150

5.3.2 Comparative mapping between barley, wheat and rice 150
5.3.2.1 YAC, PAC and BAC physical map 150
5.3.2.2 Identification of wheat ESTs from PAC and BAC clones 153
5.3.2.3 Structural analysis of putative meiosis genes 153
5.3.2.4 Analysis of rice and barley meiosis related phenotypic traits ... 153
5.4 Discussion 162
5.4.1 Physical size of *ph2a* deletion in wheat 162
5.4.2 Putative gene content of *ph2a* region 162
5.4.2.1 Structural aspects of the *Ph2* region 163
5.4.2.2 Candidate Ph2 genes 163
5.4.2.2.1 Asy1-like protein 163
5.4.2.2.2 TaMSH7 protein 165
5.4.2.2.3 MFP1-like protein 166
5.4.2.2.4 ScII protein 168
5.4.2.2.5 Damage-specific DNA binding protein 170
5.4.2.2.6 WD40 repeat like protein 170
5.4.2.2.7 Other predicted genes. 171
5.4.2.3 Possible role for multiple *Ph2* candidate genes in maintaing strict diploid-like behaviour 171
5.5 Conclusions 172

6 CHAPTER 174

General Discussion 174
6.1 Future direction 174
6.2 Wheat genes have multiple levels of complexity 175
6.2.1 Structural organisation of the *Ph2* region 175
6.2.2 Transcriptional control of meiotic wheat genes 176
6.3 Cereal genome and wheat breeding 177

BIBLIOGRAPHY 178

LIST OF FIGURES

Number *Page*

Figure 1: Mitosis 7
Figure 2: Meiosis 8
Figure 3: Homologous recombination as a mechanism to repair damaged DNA. 12
Figure 4: Resolution of the Holliday junction. 14
Figure 5: The synaptonemal complex (SC). 16
Figure 6: Chromosomal assignment of the *WM1* gene family. 62
Figure 7: Sub-chromosomal arm assignment and localisation of *WM1* gene family members within the region deleted in the *ph2a* mutant. 64
Figure 8: Sub-chromosomal arm assignment of *WM1* gene family members *WM1.1*, *WM1.2* and *WM1.3*. 65
Figure 9: Physical linkage between *WM1* gene family members *WM1.1*, *WM1.2* and *WM1.3*. 66
Figure 10: Determination of overlapping barley and wheat BAC clones through commonality in banding patterns. 68
Figure 11: Co-localisation of *WM1* gene family members on overlapping wheat BAC clones. 69
Figure 12: Co-localisation of homologues to the *WM1* gene family members on overlapping barley BAC clones. 70
Figure 13: Pulsed field gel electrophesis (PFGE) separated wheat BAC DNA of nine overlapping clones known to encode *WM1.1*, *WM1.2* and *WM1.3*. 72
Figure 14: Physical alignment of insert DNA of three overlapping wheat BAC clones (*TtBAC3*, *TtBAC6* and *TtBAC11*) based on Southern hybridisation, pulsed field gel electrophoresis and fingerprinting data. 74
Figure 15: Predicted gene content and structure of a 226 Kb DNA subfragment within 270 Kb of DNA sequence derived from *Triticum taushii* BAC clones *TtBAC3*, *TtBAC6*, *TtBAC11*. 75
Figure 16: Diagram of the predicted nucleotide coding sequence for seven members of the *WM1* gene family. 78
Figure 17: Promoter and terminator analysis of the predicted *WM1* gene family. 79
Figure 18: Relationship between the predicted primary polypeptide sequences for seven members of the *WM1* gene family. 81
Figure 19: Diagram of the structural relationship between the predicted primary polypeptide sequences for seven members of the *WM1* gene family. 85
Figure 20: Structure of leucine rich repeat and leucine zipper-like motifs from within the predicted primary polypeptide sequences for members of the *WM1* gene family. 86
Figure 21: Highly conserved N-terminal and C-terminal blocks of tandemly arrayed leucine rich repeats. 88
Figure 22: Diagram of the structural relationship between the predicted primary polypeptide of *Triticum tauschii* WM1.1 and those polypeptides found to be most homologous. 92
Figure 23: Relationship between the predicted primary polypeptide of *Triticum tauschii* WM1.1 and those polypeptides found to be most homologous 95

Figure 24: Southern blot analysis of electrophoretically separated RT-PCR products of individual members of the *WM1* gene family..... 96
Figure 25: A speculative model for the role of the predicted membrane bound recepetor-like WM1 peptides in perception and signaling leading to the regulation in development of floral tissue of wheat 107
Figure 26: Diagram of the nucleotide and polypeptide sequence of *WM5*. 109
Figure 27: Diagramatic representation of constructs used in biolistic transformation of wheat. 113
Figure 28: *In vitro* culture of wheat explants. 117
Figure 29: *In vitro* culture of wheat calli. 119
Figure 30: Diagramatic representation of the constructs used in the transient bombardment assay.. 121
Figure 31: PCR analysis of T_0 wheat lines.. 126
Figure 32: Example of Southern analysis performed on all T_0 and T_1 wheat lines. 130
Figure 33: Northern analysis showing expression pattern of A- *Npt II*, B- *GFP* C- *GUS* (no expression), D- *WM5* and E- Glyceraldehyde-6-phosphate dehydrogenase (*GAPDH* - loading control), F- Ethidium bromide stained ribosomal RNA in the segregating T_1 transgenic wheat lines. 131
Figure 35: A. Transient expression assay of green fluorescent protein (GFP) in wheat scutella from 1- Cambridge GFP, 2- Monsanto GFP, 3- Cambia GFP and 4- *pBluescript SK(+)* control. B. Whole scutella image of transient expression assay of GFP from Monsanto GFP construct 133
Figure 36: Fluorescent microscopy for GFP fluorescence in the T_1 progeny of transgenic line 16.Y. 135
Figure 37: GUS expression of pre-meiotic spikes from T_1 progeny of 24.Z and 30.Z 136
Figure 38: The region deleted in the *ph2a* mutant of wheat delimited in the three-barley mapping populations Chebec x Harrington, Clipper x Sahara and Galleon x Haruna Nijo). 151
Figure 39: The syntenous *Ph2* regions in rice, wheat and barley 152
Figure 40: Physical alignment of the YAC clones with respect to the rice genetic map, highlighting the estimated physical size of the homologous *Ph2* region in rice with the physical positions of the putative meiotic or meiosis related genes. 154
Figure 41: Putative structure of rice Asy1-like protein 156
Figure 42: Putative structure of rice TaMSH7-like protein 157
Figure 43: Putative structure of rice MFP1-like protein 158
Figure 44: Putative structure of rice ScII-like protein 159
Figure 45: Putative structure of rice damage-specific DNA binding protein 160
Figure 46: Putative structure of rice WD40 repeat like protein 161

LIST OF TABLES

Number *Page*

Table 1: DNA sequences of primers designed to amplify hypervariable and coding sequences for members of the *WM1* gene family. 44
Table 2: Electrophoretic parameters used to resolve restricted HMW DNA 47
Table 3: BAC fingerprinting reaction components and conditions 53
Table 4: Type II restriction enzymes and corresponding incubation temperatures used for BAC fingerprinting 54
Table 5: Gene specific RT-PCR primers for individual members of the *WM1* gene family. RT-PCR control primers for *MSH7* and *Thioredoxin H*. 61
Table 6: RiceGAAS software identified coding sequences from 270 Kb of *Triticum tauschii* genomic sequence. 77
Table 7: Polypeptide sequence similarity and divergence between the predicted primary peptides of seven members of the *WM1* gene family. 82
Table 8: Properties of predicted WM1 primary peptides 84
Table 9: Prediction of subcellular localisation, membrane topology and signal sequence for WM1 predicted polypeptides. 90
Table 10: Polypeptide sequence similarity and divergence between the predicted primary polypeptide of *Triticum tauschii* WM1.1 and those polypeptides found to be most homologous 93
Table 11: Experimental design for co-bombardment of transgene constructs into wheat scutellum. 114
Table 12: Parameters used in microprojectile bombardment of scutellar tissue in wheat. 115
Table 13: PCR primers designed for the amplification of antisense *WM5* (p*WM53.1*), sense *WM5* (p*WM54*), 35S *WM5* (p*WM55.6*), *WM5* GUS (p*WM51*) and *WM5* GFP (p*WM52.1*), and pAct1-F/*Npt II*. 122
Table 14: Restriction endonucleases used for digestion of genomic DNA from T_0 and T_1 plantlets to estimate both copy number and whole transgene integration. 124
Table 15: Kanamycin selected wheat regenerants identified as being transgenic through both PCR and Southern analysis. 127
Table 16: Inheritance of transgenes from T_0 to 10 random T_1 progeny as identified through both PCR and Southern analysis. 129
Table 17: Restriction fragment length polymorphism probes (molecular markers) for determining *ph2a* deletion size 149
Table 18: Putative meiosis related genes identified from rice PAC and BAC sequence along with cereal ESTs. 155
Table 19: Transgenic plants produced from individual microprojectile bombardment experiments 6

ACKNOWLEDGMENTS

I wish to express sincere appreciation to the following people and institutions for their support and assistance during the course of this study:

My supervisor, Professor Peter Langridge for his advice, encouragement, patient supervision, and the critical reading of this manuscript. The unique opportunities provided during my studies are sincerely appreciated and has resulted in extremely fulfilling research.

Research fellows Dr Chongmei Dong and Dr Ute Baumann for their constructive criticism and suggestions throughout the course of my studies.

Friend and colleague Tim Sutton for those many interesting and stimulating discussions about our research that occurred over a few beers both in Adelaide and in the USA.

Dr Jason Able for the critical reading of this manuscript. Dr Petra Wolters, and Dr Dirk Becker for all their technical assistance in DNA sequence analysis and wheat genetic transformation. All the members of the Langridge Lab and those people from the Plant Science Department who helped me during my studies. In particular I would like to thank Patricia Warner, Jodie Kretschmer, Dr Ursula Langridge, Angelo Karakousis, Dr Brendon King and Juan Juttner for their help and friendship.

The Grains Research and Development Corporation, Agriculture Western Australia, Coorporative Research Centre for Molecular Plant Breeding and DuPont Ag Biotech for financial support over the couse of my research.

My friends and family for their support and encouragement over the last few years

Finally I would like to acknowledge the love and support of my soul mate and best friend Nadja.

DEDICATION

This thesis is dedicated to my late grandfather Stanley Harold Whitford (1916 – 1988). I hope that my pursuit for a career in Agricultural Science, rather than working Springhill Farm, would have made him proud.

ABBREVIATIONS

aa	amino acid
BAC	bacterial artificial chromosome
bp	base pairs
cDNA	complementary DNA
cv	cultivar
DNA	deoxyribonucleic acid
DT	ditelosomic
DTT	dithiothreitol
EDTA	ethylenediaminetetracetic acid
EMC	embryo mother cell
EST	expressed sequence tag
g	grams
x g	9.81 m/s^2
GFP	green fluorescent protein
GUS	β-glucuronidase
HMW	high molecular weight
hr(s)	hour(s)
Kb	kilobase
KDa	kilodalton
L	litre
lx	lux
M	molar
mA	milliampere
Mb	megabase
mbar	millibar
min	minutes
mg	milligrams
mJoules	mega Joules
mL	millilitres
mm	millimetres
mM	millimolar

MOPS	morpholino propanesulfonic acid
ng	nanograms
nos	nopaline synthase gene
Npt II	neomycin phosphotransferase
NT	nullisomic-tetrasomic
ORF	open reading frame
PCR	polymerase chain reaction
RT-PCR	reverse transcription polymerase chain reaction
PFGE	pulse field gel electrophoresis
pfu	plaque forming units
PDS	particle delivery system
PMC	pollen mother cell
psi	pounds per square inch
rpm	revolutions per minute
RNA	ribonucleic acid
RT	room temperature
SDS	sodium dodecyl sulphate
sec	seconds
TEMED	tetramethylethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
w/v	weight per volume
X-gluc	5-bromo-4chloro-3-indolyl-β-D-glucuronic acid
UV	ultraviolet
YAC	yeast artificial chromosome
2,4-D	2,4-dichlorophenoxyacetic acid
μg	micrograms
μL	microlitres

1 Chapter

LITERATURE REVIEW

1.1 General introduction

The dynamic cellular processes within higher eukaryotic organisms remain largely unknown. Our growing awareness of cellular development including mitosis and meiosis often results in a tendency to manipulate these processes for further understanding and ultimately for economic and sociological benefit. The pursuit for understanding the mechanisms controlling the developmental process of meiosis is of crucial importance given that it is the basis of sexual reproduction. In bread wheat (*Triticum aestivum*), sexual reproduction followed by fertilisation results in the formation of seed. Derivatives of this seed forms the foundation of the staple diet for the much of the world's population.

Model organisms such as yeast (*Saccharomyces cereviseae*) have provided valuable insight into the genetic control of meiosis, more specifically chromosome pairing and recombination. However, it is still largely unproven whether there is strict commonality in fundamental mechanisms controlling meiosis among all eukaryotic organisms. The polyploid nature of wheat poses an added level of complexity when trying to understand the mechanisms of chromosome pairing. This literature review explores classical views of meiosis through understanding in model organisms and endeavours to shed light on the genetic control of meiosis in polyploid organisms, specifically the control of chromosome pairing and recombination in bread wheat (*T. aestivum*). Additionally, it aims to highlight the importance of manipulating chromosome pairing and recombination with respect to wheat breeding.

1.2 Mitosis and meiosis

Mitosis and meiosis represent two fundamental processes in eukaryote cell division.

Mitosis constitutes a phase in the cell cycle, alternating with interphase that combines two growth periods (G1 and G2), between which the chromosomes replicate (S phase). Following the second growth period, the cell enters the mitotic phase (M) (**Figure 1**).

Meiosis shares many corresponding steps to the events in mitosis. However, meiosis is the essential complement to the genome doubling at fertilisation that occurs in sexually reproducing organisms, as it provides a mechanism to reduce the genome size in half. For the daughter cells and products of fertilisation to be viable, this must be a precise reduction into two functionally equivalent halves - each cell must possess a full haploid genome complement. The way this is achieved by the great majority of organisms is the pairing of homologous chromosomes during prophase of the first meiotic division. When homologous chromosomes are paired, they are able to orient their kinetochores in opposite directions so that at anaphase, half of the chromosomes are pulled to each spindle pole (**Figure 2**). This simple description of pairing belies a cell biological process of immense complexity.

Two important properties distinguish meiosis from mitosis:

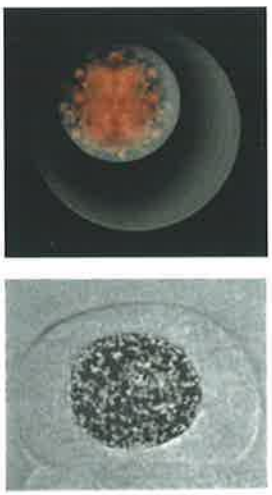
1. In meiosis, the homologous chromosomes pair lengthwise and their chromatids exchange genetic material by crossing-over.

2. The sister chromatids, which are not identical after crossing-over, do not separate from one another in the first nuclear division, and the chromosomes do not replicate between the two nuclear divisions.

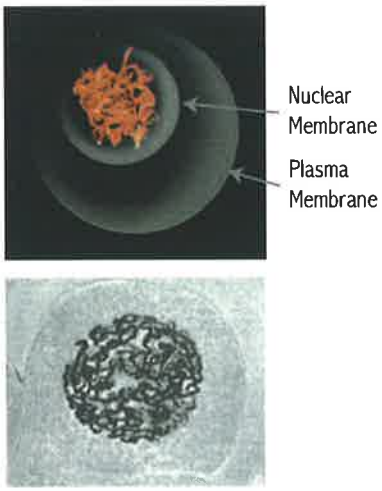
(Raven & Johnson, 1992)

Figure 1: Mitosis. The photographs show nuclei of root tip cells of *Lilium regale*. Modified from (Grimes *et al.*, 2001; McLeish & Snoad, 1958).

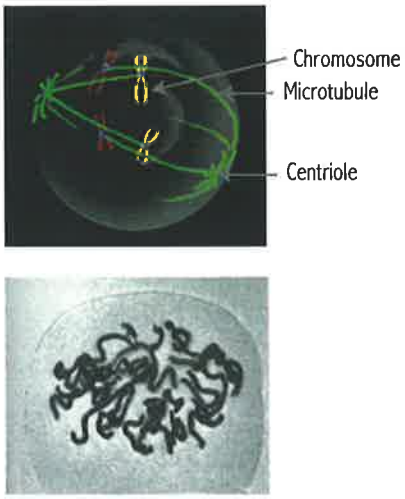
Interphase



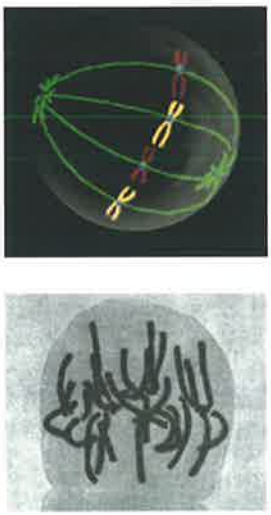
Early Mitotic Prophase



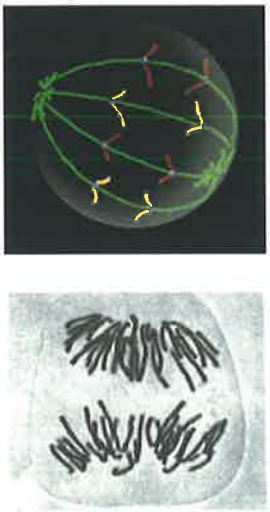
Late Mitotic Prophase



Mitotic Metaphase



Mitotic Anaphase



Mitotic Telophase

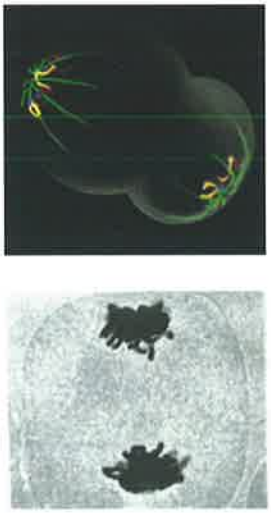
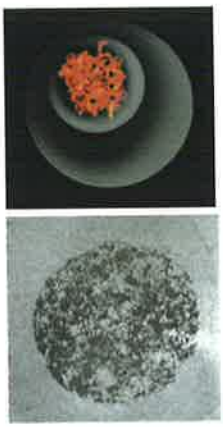
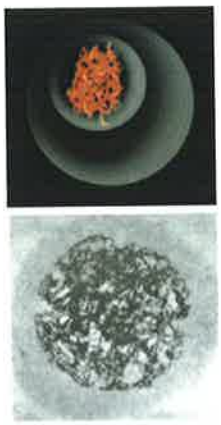


Figure 2: Meiosis. The photographs are of *Lilium regale.* Note: For simplicity, multiple chiasmata are drawn as involving only two chromatids; in reality, all four chromatids can be involved. Modified from (Grimes *et al.*, 2001; McLeish & Snoad, 1958).

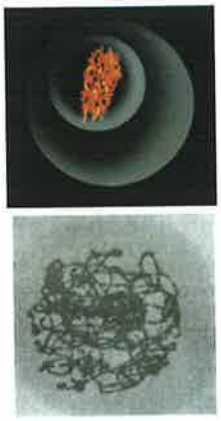
Leptotene



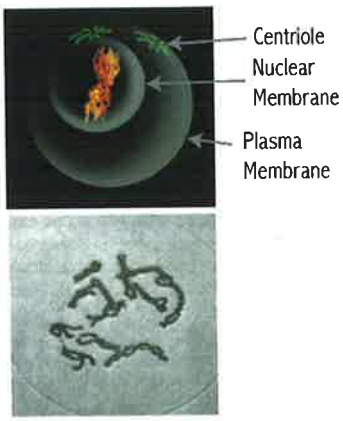
Zygotene



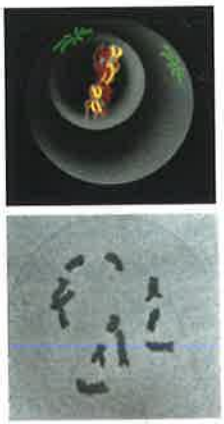
Pachytene



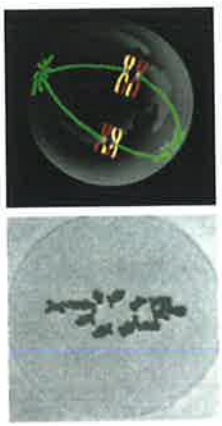
Diplotene



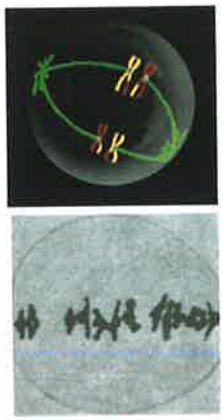
Diakinesis



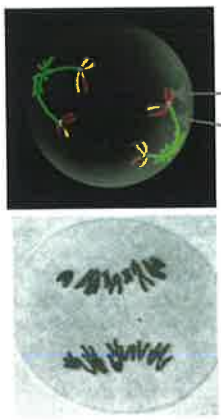
Metaphase I



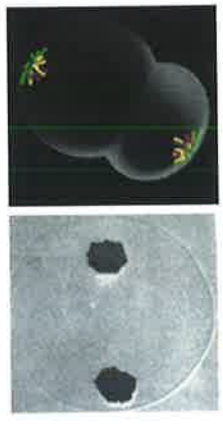
Early Anaphase I



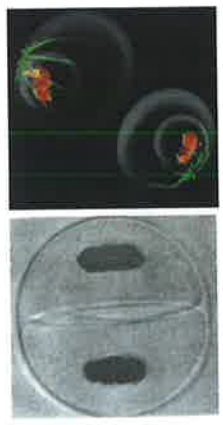
Later Anaphase I



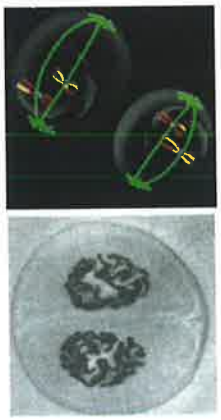
Telophase I



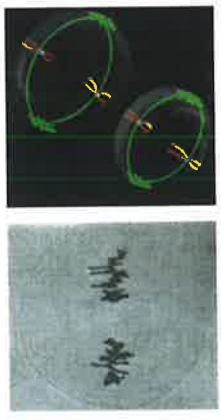
Interphase



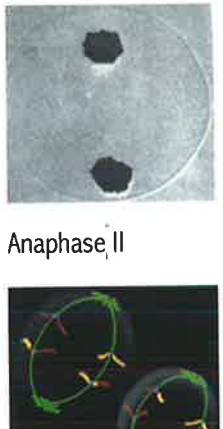
Prophase II



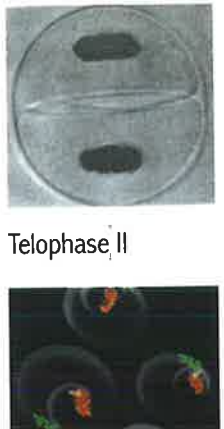
Metaphase II

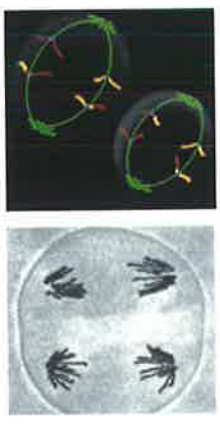


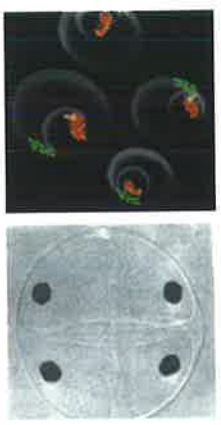
Anaphase II



Telophase II







The meiosis specific mechanism of genetic recombination or crossing over is the exchange of DNA sequences on maternal and paternal chromatids resulting in new combinations of genes. This provides a mechanism for generating genetic diversity beyond that achieved by the independent assortment of chromosomes.

1.3 Model organisms for the study of meiosis

The budding yeast *S. cereviseae* has been extensively studied in relation to meiosis largely due to its ease of manipulation (eg. mutagenesis), availability and production of developmentally defined material, and the availability of whole genome sequence information. Through the use of *S. cereviseae* as a model for the genetic control of meiosis, molecular mechanisms underlying chromosome pairing and recombination are outlined and compared with plant systems.

1.3.1 Chromosome pairing

The literature provides us with an intense debate on the sequence of events surrounding chromosome homology recognition, alignment, synapsis and synaptonemal complex formation (Bennett, 1984; Loidl, 1990; Moens, 1994). Does it follow this particular sequence, another sequence, or does it occur simultaneously? How do homologous chromosomes recognise each other?

Chromosome pairing is generally agreed to be the close association of homologous or homeologous chromosomes at late leptotene in wheat. This chromosome pairing and subsequent chiasmata formation is the mechanism by which genetic information is exchanged between homologous or homeologous chromosomes resulting in new gene combinations.

Meiotic homologues come into homology-directed physical contact with one another no later than leptotene. In yeast, homologues put side by side prior to meiotic S-phase are dislocated during S-phase, and then re-established towards the end of S-phase or directly thereafter (Scherthan *et al.*, 1994; Weiner &

Kleckner, 1994). Whether all monocots conform to this generalisation remains to be determined (Franklin *et al.*, 1999).

In yeast, homologues identify one another in the apparent absence of any chemical interruptions in their DNA, not only during somatic pairing but also during meiotic prophase (Loidl *et al.*, 1994; Nag *et al.*, 1995; Weiner & Kleckner, 1994). This homologue recognition seems to be independent of recombination (Loidl *et al.*, 1994; Weiner & Kleckner, 1994). Chromosome pairing has been associated with heterochromatic regions, along with specific or specialised chromatin structures or compartments in plants, humans and *Drosophila*. Many hypotheses attempting to explain this process occur in the literature. It has been suggested that pairing contacts are unstable and it is expected that there is necessity for the synergistic cooperation of multiple elements to stabilise the contacts. This is likely to be associated with the number and proximity of interacting copies, receptiveness of underlying DNA sequence to DNA/DNA interactions and suitable "stickiness" factors (Zickler & Kleckner, 1999). Certain sequences at the DNA level have been implicated directly in homologous interactions (Zickler & Kleckner, 1998). However, aside from homology at the DNA level, pattern recognition has also been implicated in chromosome pairing (Albini & Jones, 1987; Chandley, 1986).

1.3.2 Homologous recombination

Recombination is any process generating new combinations of pre-existing genetic material. Homologous recombination is specifically homology dependent, but not sequence-dependent, so any two DNA molecules of related sequence can undergo recombination by this process.

The isolation of *S. cerevisiae* mutants defective in DNA damage repair, and blocked during meiosis has been useful for the identification of recombination genes. Recombination repair is any DNA repair or damage tolerance tactic involving homologous recombination.

There are two substrates for homology dependent recombination repair: double strand breaks (DSBs) and single strand gaps. **Figure 3** outlines mechanisms

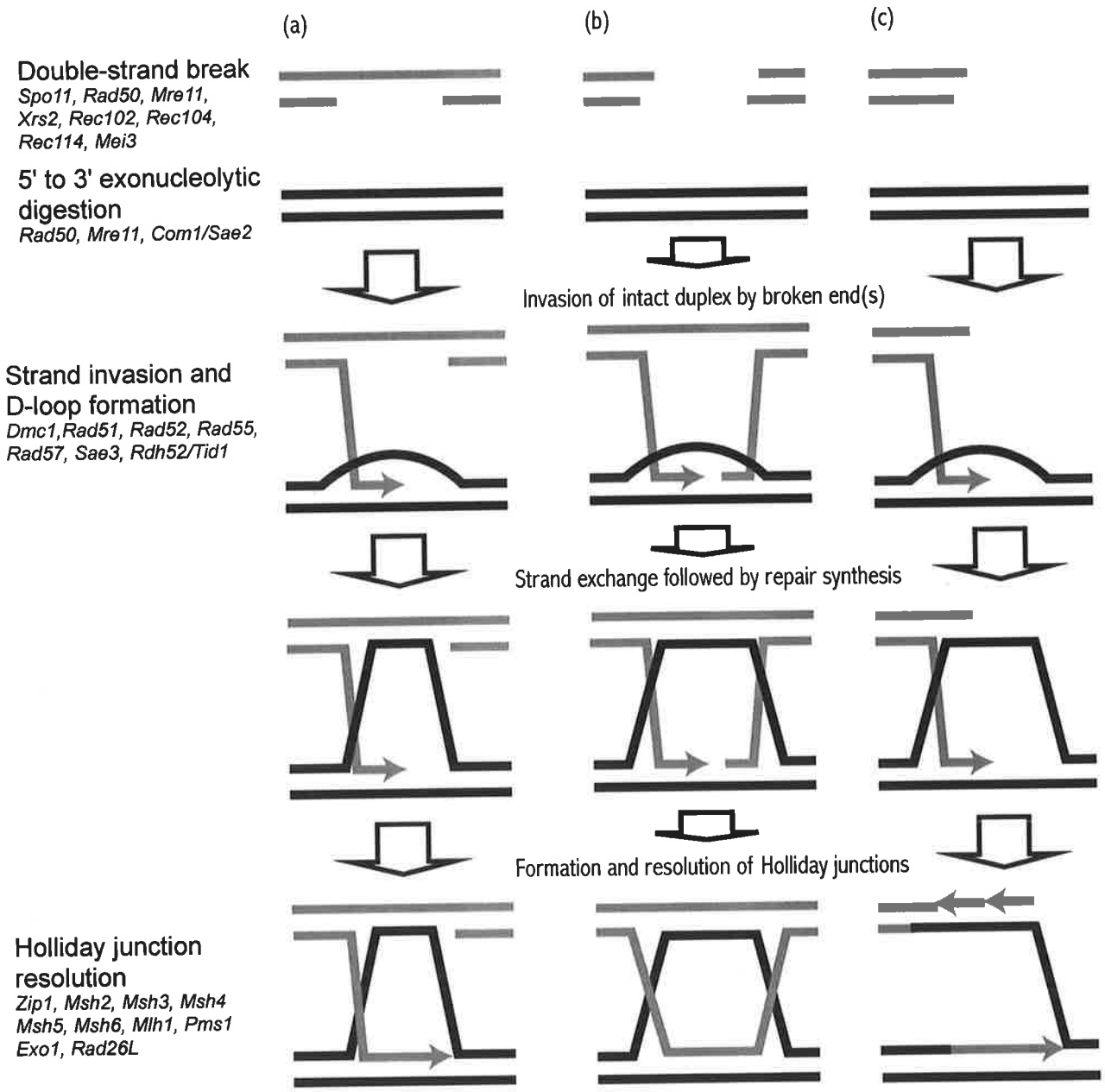
involved in this homology dependent recombination repair. Interestingly, comparable mechanisms form the basis for both double strand breaks and single strand gaps. Described below is the molecular basis for double strand break repair.

1.3.2.1 Double strand break repair

Recombination at the molecular level is a result of a specific pathway involving the processing of DNA double-strand breaks (DSBs) (Game *et al.*, 1989; Sun *et al.*, 1989). These DSBs are site-specific and take place during early to mid prophase (Padmore *et al.*, 1991). They are also accountable for much or all-meiotic recombination in yeast (Sun *et al.*, 1989). DSBs are 5' resected leaving 3' ssDNA of about 600 nucleotides in length (Weiner & Kleckner, 1994). Meiosis-specific DSBs have been shown to arise prior to or at the same time as the synaptonemal complex formation. Ensuing stable joint molecules take shape (Schwacha & Kleckner, 1994) and recombination is achieved by the end of pachytene (Padmore *et al.*, 1991). Such a process would suggest that a RecA-like protein (functions in pairing homologous DNA molecules and the subsequent exchange of DNA strands in *Escherisha coli*) would bind to the ssDNA and assist in the homology searching. However, meiotic homologue pairing transpires in the absence of all four *RecA* homologues (Kleckner, 1996). Instead the *Rad52* epistasis group of genes from yeast have been shown to be involved in recombination repair (Cameriniotero & Hsieh, 1995). *Rad51* (inducible by DNA damage) is the major gene of the group with homologues having been isolated from many eukaryotes. Other genes from the *Rad52* epistasis group have mutant phenotypes, which implicate them in having roles in both mitosis and meiosis.

Other genes have been observed to be involved in recombination repair with six genes accepted to be meiosis-specific and involved specifically in this DSB establishment. These consist of *Spo11*, *Me14*, *Mer2*, *Rec102*, *Rec104* and *Rec114*. There are also three required for the repair of DSBs in non-meiotic cells. These include *Rad50*, *Mre11* and *Xrs2* (Keeney *et al.*, 1997).

Figure 3: Homologous recombination as a mechanism to repair damaged DNA. Repair of single-strand gaps (a) and double strand breaks (b) involves new replication across the lesion using a strand from the undamaged duplex as a template. Completion of the replication is followed by strand ligation, forming Holliday junctions, which can be resolved as shown in **Figure 4**. Recombination involving a partial chromosome (c) generates a Holliday junction intermediate, which is resolved as a replication fork. Arrows represent the direction of new DNA synthesis (Twyman, 1998). Outlined on the left are those genes of yeast found to be involved in the recombination process (Zickler & Kleckner, 1998; Zickler & Kleckner, 1999).


(a)
(b)
(c)
Double-strand break
Spo11, Rad50, Mre11,
Xrs2, Rec102, Rec104,
Rec114, Mei3
5' to 3' exonucleolytic
digestion
Rad50, Mre11, Com1/Sae2
Invasion of intact duplex by broken end(s)
Strand invasion and
D-loop formation
Dmc1,Rad51, Rad52, Rad55,
Rad57, Sae3, Rdh52/Tid1
Strand exchange followed by repair synthesis
Formation and resolution of Holliday junctions
Holliday junction
resolution
Zip1, Msh2, Msh3, Msh4
Msh5, Msh6, Mlh1, Pms1
Exo1, Rad26L

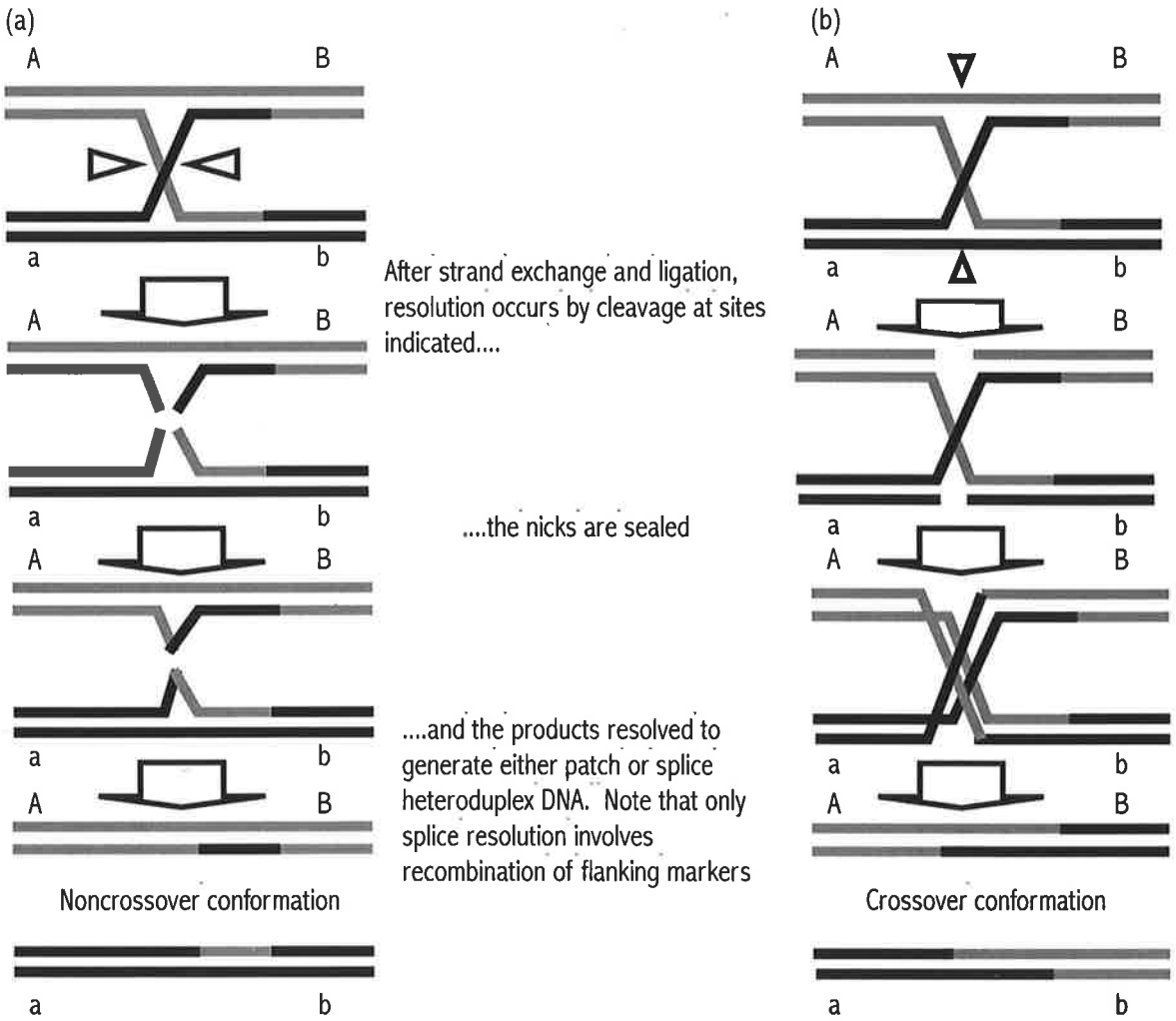
Mechanistically these double strand breaks occur at hotspots of recombination and are made, in conjunction with other proteins, by the Spo11 protein, which becomes covalently attached to the 5' ends at the break. Progressive shortening of the 5' – ended strands, perhaps by the Rad50-Mre11-Xrs2 complex, exposes 3' – ended single strands (**Figure 3**) (Davis & Smith, 2001). Assisted by Rad51, Dmc1 and other strand exchange proteins, these single strands are believed to invade homologue duplex DNA to form a displacement loop (D-loop). D-loop joint molecules are altered into Holliday junctions, which are then transformed into recombinant molecules with either a cross-over or noncrossover configuration (**Figure 4**) (Davis & Smith, 2001). The heteroduplex DNA formed in the Holliday junction is typically corrected by the mismatch repair system.

1.3.2.2 Mismatch repair system

The yeast mismatch repair system is understood to be involved in meiotic recombination through the processing of non-homologous ends during double-strand break-induced recombination (Sugawara *et al.*, 1997). A mismatch repair protein complex, MSH2-MSH6, has also been observed to bind to Holliday junctions during meiotic recombination (Alani *et al.*, 1994; Marsischky *et al.*, 1999).

The affinity of the MSH2-MSH6 complex for Holliday junctions was shown to be at a level comparable to that as for mispaired bases (Marsischky *et al.*, 1999). Apart from mismatch correction, MSH2, MSH3 and MSH6 additionally play a role in inhibiting recombination between divergent DNA sequences (Chambers *et al.*, 1996; Chen & Jinks-Robertson, 1998; Datta *et al.*, 1996; Hunter *et al.*, 1996; Selva *et al.*, 1997; Selva *et al.*, 1995; Sugawara *et al.*, 1997). The suppression of homeologous recombination mediated by the mismatch repair system is believed to be through the detection of mispairs within the recombination intermediates and the prevention of processing of these recombination intermediates (Baker *et al.*, 1995; de Wind *et al.*, 1995; Edelmann *et al.*, 1996).

Figure 4: Resolution of the Holliday junction in either of two planes (a) and (b), generating different products. Only one resolution pathway generates a molecule, which is recombinant for flanking markers A and B, although both pathways generate a region of heteroduplex DNA (Twyman, 1998).


(a)
(b)
A
B
a
b
After strand exchange and ligation, resolution occurs by cleavage at sites indicated....
....the nicks are sealed
....and the products resolved to generate either patch or splice heteroduplex DNA. Note that only splice resolution involves recombination of flanking markers
Noncrossover conformation
Crossover conformation

Consequently the mismatch repair system plays a crucial role in maintaining genetic stability by recognising and processing mismatched nucleotides that may occur during DNA replication, genetic recombination and some types of chemical damage to DNA.

Chromosome associations instigated through DSBs and the mismatch repair system result in the formation of the long "ribbon-like" synaptonemal complex (**Figure 5**). The synaptonemal complex is a meiosis specific proteinaceous structure (Klein, 1994), which initiates at late leptotene and degrades by early diplotene. This synaptonemal complex is termed the chaperone of crossing over and is characterised by a tripartite structure consisting of two lateral elements, a central element and a complex array of transverse filaments (Hasenkampf, 1996). The synaptonemal complex is absolutely crucial in fastening the pairing chromosomes together and mediating the resolution of homologous recombination.

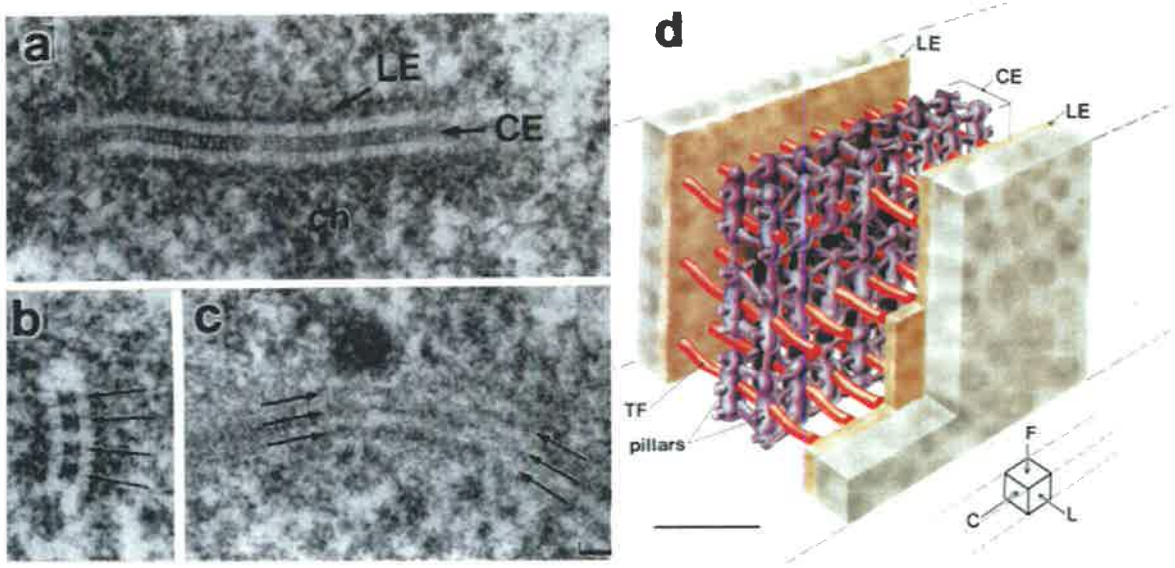
1.4 Meiosis in allohexaploid wheat

1.4.1 Genome organisation

The common bread wheat, *T. aestivum*, is generally considered as an allopolyploid. Bread wheat being an allohexaploid (2n = 6x = 42) consists of three separate genomes, the A, B and D genomes, which seem to have arisen by two successive hybridisation events. The A genome arising from *Triticum uratu* (2x = 14); the B genome from *Aegilops speltoides* (2x = 14) (Blake *et al.*, 1998); and the D genome from *Triticum tauschii* (2x = 14) (Sears, 1976).

Chromosomes from each of the separate genomes are termed homeologues whilst those chromosomes within each genome are termed homologues. Although homeologues share extensive sequence homology, hexaploid wheat behaves in a diploid manner where pairing is strictly between true homologues. Even in haploids very little homeologous chromosome association is observed (Riley, 1960). Consequently homeologous chromosomes do not recombine with each other and this helps maintain genome stability.

Figure 5: The synaptonemal complex (SC). (a) Longitudinal section of *Blaps cribrosa* SC with clear scalariform central element (CE) and two lateral elements (Karpen *et al.*, 1996) surrounded by chromatin (ch). (b) Cross section showing the complex four layered CE. (c) Lateral section with three of the CE layers indicated by arrows (the round dark ball on top of the layers is a recombination nodule). (d) Three-dimensional model of the central region in *B. cribosa*. The multilayered CE are in purple, the flanking LEs in pale yellow (indicated as plates), the transverse filaments (TF) crossing the central region in red, and the surrounding chromatin is in grey. Directions of views indicated *F*, frontal, *L*, lateral, *C*, cross-sectional view. Bar = 50 nm (Schmekel *et al.*, 1993).


a
LE
CE
cm
b
c
d
LE
CE
LE
TF
pillars
F
C
L

At meiotic prophase bread wheat undergoes a two-phase pattern of synaptonemal complex associations. Firstly, a complex network of synaptic interactions establishes at zygotene and early pachytene (Holm, 1986; Holm & Wang, 1988) meaning homology detection and synaptonemal complex formation is taking place between homeologues as well as homologues. Synapsis is progressively limited to homologues by the end of pachytene (Holm, 1986) and by metaphase I, only chiasmate bivalents are present, meaning that crossing over has occurred strictly between homologues and not between homeologues (Feldman & Avivi, 1988; Holm, 1988a; Holm, 1988b; Holm & Wang, 1988; Martinez-Perez *et al.*, 1999; Mikhailova *et al.*, 1998; Sears, 1976). It seems as though chromosome synapsis in bread wheat undergoes a corrective or adjustment step at mid-pachytene to allow only homologous chromosome pairing.

Most allopolyploids, whether natural (*Scilla autumnalis*, *Avena sativa*, *Avena maroccana*, *Aegilops biuncialis*) or hybrids (*Lolium* spp.), seem to behave like bread wheat (Cunado *et al.*, 1996; Jenkins & Jimenez, 1995; Jenkins *et al.*, 1988; Peterson *et al.*, 1996; Thomas, 1990). Synapsis is confined mainly to homologues even as early as zygotene, before the corrective phase of mid-pachytene, as observed in wheat.

Chromosome pairing in autopolyploids is greatly different. Autopolyploids stem from from the multiplication of a single genome. This consequently causes multivalent associations at prophase and metaphase I. Associations between these homologous chromosomes are still mediated by crossovers and synaptonemal complexes (Sherman *et al.*, 1989; Vincent & Jones, 1993). Multivalent formation in autopolyploids often results in reduced fertility.

1.4.2 Chromosome pairing and recombination

Homologues seem to find their counterparts in a very short time; in wheat this is in a matter of hours. This is such a remarkable process considering the nucleus is crowded with chromosomes, which are entangled. In wheat these chromosomes contain large stretches of repeated sequences of which all

homeologous chromosomes share. How each homologous chromosome recognises each other at the molecular level considering these adversities remains a mystery.

It is interesting that plant breeders have been unable to create synthetic allopolyploids that undergo regular bivalent formation as seen in wheat and oats (Benavente & Orellana, 1991). This suggests that the naturally occurring allopolyploids or progenitors must have evolved special mechanisms for determining stringency control in chromosome pairing so as to promote viability.

Genetic determinants have been found that manage the stringency of chromosome pairing in allopolyploids. A complex group of genes have been defined that control the level of pairing between homologous and homeologous chromosomes. These genes have been studied in detail for many decades (Driscoll, 1972; Driscoll, 1973; Feldman, 1966; Riley & Chapman, 1958; Riley *et al.*, 1966; Sears, 1976; Sears, 1977; Sears, 1982; Sears & Okamoto, 1958). These *P*airing *H*omeologous (*Ph*) genes were described based on wheat euploid and mutant lines that had varying abilities to promote or suppress pairing between homeologous chromosomes. Other factors that have been found to regulate chromosome pairing in allopolyploids and amphiploids include heterochromatic B chromosomes. The heterochromatic B chromosomes affect the regularity of chiasmatic associations within complex genomes (Sears, 1976).

1.4.2.1 Suppressors of homeologous chromosome pairing

Suppressors of chromosome pairing act to restrict chiasma formation strictly to homologous chromosomes resulting in strict bivalents forming at metaphase I. The gene expressing the strongest influence on the suppression of homeologous chromosome pairing has been localised to the long arm of chromosome 5 (5BL) and has been termed *Ph1* (Okamoto, 1957; Riley & Chapman, 1958; Sears & Okamoto, 1958). Minor suppressors have been found on 3DS and 3AS. The gene located on 3DS, termed *Ph2*, is a stronger suppressor of pairing than the gene located on 3AS (Cuadrado *et al.*, 1991; Driscoll, 1972; Mello-Sampayo, 1971; Mello-Sampayo & Canas, 1973; Uphadya

& Swaminathan, 1967). An even weaker suppressor of pairing has been located on chromosome 4D (Driscoll, 1973). These suppressors of homeologous chromosome pairing have primarily been identified through euploid wheat lines or mutagenesis.

1.4.2.1.1 The mechanistic action of Ph *genes*

There have been many proposals for the mechanism controlling the actions of the *Ph* genes. One of the earliest models, suggested that pairing involved the ratio of DNA to histones during synapsis (Ansley, 1958). It was found that cells undergoing synapsis had a low ratio of DNA to histones. How this affected the stringency of chromosome pairing was unknown at the time. Ansley's model was superseded by Feldman and colleague's (Avivi & Feldman, 1973a; Avivi & Feldman, 1973b; Feldman & Avivi, 1973) who suggested that *Ph* genes regulated chromosome proximity. Studies were based on the effect of *Ph* gene dosage and colchicine treatment on somatic chromosome associations. It was proposed that pairing suppressors destabilised the interaction between chromosomes and both the microtubules and spindle, whereas the pairing promoters stabilised these interactions. Studies on chromosomes from early meiotic prophase showed that the chromosomes lie closer together in the absence of a certain *Ph* gene, whilst when there are six copies of the same *Ph* gene, chromosomes lie nearly at random. Driscoll *et al.* (1979) on the other hand hypothesised that there is an enzyme system affected which allows pairing events to proceed to chiasma formation. Recently Liu *et al.* (1998) suggested that homology search and initiation of pairing is controlled both on the genic level (ie *Ph* genes) and by homeologous differentiation. Homeologous differentiation is based on low-copy, non-coding, chromosome specific sequences (CSSs) that are clustered in interstitial, subterminal and terminal (subtelomeric) regions of the chromosome.

The debate on the mode of action of *Ph* genes has been raised since the 1960's and to date still continues.

*1.4.2.1.2 Pairing Homeologous gene 1 (*Ph1*)*

Taking *P*airing *H*omeologous gene 1 (*Ph1*) as an example, *ph1*/*ph1* mutant wheat lines (eg *ph1b*, *ph1c*) and those euploid wheat lines missing both copies of chromosome 5BL, chiasma formation is no longer strictly between homologues. Both homologous chromosome synapsis and multivalent formation is observed, which is associated with a decrease in the total number of chiasmata (Holm & Wang, 1988; Vega & Feldman, 1998a). Conversely, six copies of *Ph1* by way of multiplication of 5BL, results in primarily univalent arrangement at metaphase I with a reduction in the levels of chiasmata, even between homologues. It has been shown however that the chromosomes are still interacting, even with high levels of interlocking homeologues (Feldman, 1966; Feldman & Avivi, 1988; Holm, 1988a; Holm, 1988b; Holm & Wang, 1988; Luo *et al.*, 1996; Sears, 1976; Yacobi *et al.*, 1982). These interlockings, which may represent first stages of homology searching, do not result in the formation of chiasmata. This suggests that the dosage effect of *Ph1* results in a loss in the discrimination between homologues and homeologues. Too many (six copies) or too few *Ph1* (zero copies) ultimately results in lower levels of chiasmata both between homologues and between homeologues. Such variations in chiasmata formation are directly reflected in the formation of the synaptonemal complex. Either zero or six copies of *Ph1* results in incomplete synaptonemal complex formation along with a greater number of multiple homologous and homeologous synaptic associations (Holm, 1988a; Holm, 1988b; Holm & Wang, 1988).

Recent studies using fluorescent *in situ* hybridisation on mutant lines of *Ph1* have helped our understanding of the mode of action of the *Ph1* gene. Findings suggest that mutants of *Ph1* have altered chromosome/chromatin organisation and compaction, not only in meiotic cells but also in somatic cells (Aragon-Alcaide *et al.*, 1997; Mikhailova *et al.*, 1998; Vega & Feldman, 1998b). There seems to be a premature separation of sister chromatids and extension of the centromeric chromatin in univalents at anaphase I. Observations also show that there is breakage of centromeres such that the two arms of a chromatid (or chromatid pair) are estranged from one another (Aragon-Alcaide *et al.*, 1997;

Vega & Feldman, 1998b). *Ph1* mutants also seem to have alterations in the relative arrangement of homologous chromosomes both in meiotic and somatic cells (tapetal cells) (Mikhailova *et al.*, 1998). This alteration occurs in meiotic cells from pre-meiotic interphase. It seems that *Ph1* specifies or affects some basic component of chromosome structure.

*1.4.2.1.3 Pairing Homeologous gene 2 (*Ph2*)*

Pairing Homeologous gene 2 (*Ph2*) is a minor suppressor of homeologous chromosome pairing as compared to *Ph1*, which is located on chromosome 3DS. An X-ray induced deletion of the *Ph2* gene, *ph2a* (Sears, 1977) and a point mutation, *ph2b* (Wall *et al.*, 1971), reveal that the removal of the gene induces higher levels of homeologous chromosome pairing in wheat hybrids with alien species, but does not effect chromosome pairing in wheat itself (Sears, 1977).

Ph2 being a minor suppressor of homeologous chromosome pairing as compared to *Ph1* holds true when comparing metaphase I pairing associations within wheat x rye hybrids. The X-ray deletion mutant, *ph1b*, has an average of 10.71 chiasmatic associations per cell, whilst *ph2b* has an average of 3.43 chiasmatic associations per cell. This compares to 1.36 chiasmatic associations in wild-type *Ph1Ph2* wheat (Benavente *et al.*, 1998). In the same study it was suggested that whatever the meiotic mechanism actually affected by a mutation at the *Ph2* locus, the ability to discriminate chromosome similarity among related partners seemed not to be influenced. The greater pairing promotion effect of the *ph1b* mutation appears to be relatively more on distant homeologous partner metaphase I associations, whereas the lower promoting effect of *ph2b* is evenly distributed among all types of homeologous associations. It is also suggested that the resolution of wheat x rye metaphase I associations into wheat x rye recombination events in *ph2b* is lower than that for *ph1b*. This finding reveals that distinct mechanisms are involved in the control on homeologous synapsis and/or chiasma formation by the two *Ph* genes (Benavente *et al.*, 1998).

More detailed ultrastructural analysis comparing *ph1b* and *ph2b* illustrates that only a few nuclei accomplish synapsis (synaptonemal complex formation) in the *ph2b* genotype, whereas most nuclei completed synapsis in the wild type and *ph1b* genotypes (Martinez *et al.*, 2001). Results of this study suggest that neither *Ph1* or *Ph2* affect synaptic restriction to bivalents at early prophase but do have a different effect on later synaptic behaviour. This result further illustrates that *Ph1* and *Ph2* bring about diploidisation of allohexaploid wheat by different mechanisms. *Ph2* seems to affect synaptic progression, probably in a similar way to a diploid species. It has been suggested that *Ph2* itself may not represent a pairing homeologous (*Ph*) locus but a synaptic (*Syn*) locus (Martinez *et al.*, 2001). This raises the question of whether homeologous chromosome pairing as phenotypically determined through analysis of the *Ph2* locus is a consequence of asynapsis?

It is also interesting that less attention has been paid to the effect of *ph2*/*ph2* in wheat x alien hybrids, even though its intermediate level of promotion of homeologous chromosome pairing is thought to be more appropriate for genetic transfers from closely related species (Sears, 1982). If *Ph1* and *Ph2* do indeed represent different mechanisms influencing homeologous chromosome pairing, further understanding of these mechanisms at a molecular level may help in their synergistic integration to produce wheat lines of increased homeologous chromosome pairing and recombination capacity.

1.4.2.2 Promoters of homeologous chromosome pairing

Unlike the suppressors of pairing, *Ph3*, which is located on 5BS is a promoter of pairing (Sears, 1972). Other promoters of pairing have been located on chromosomes 3DL (Driscoll, 1972; Mello-Sampayo & Lorente, 1968), 3BL (Kempanna & Riley, 1962; Sears, 1954), 3AL (Mello-Sampayo & Canas, 1973), 5DL (Feldman, 1966; Feldman, 1968; Mello-Sampayo, 1972; Riley *et al.*, 1966), and 5AL (Feldman, 1966; Feldman, 1968; Riley *et al.*, 1966). These studies revealed that promoters of homeologous chromosome pairing increase levels of multiple chromosome associations. However, there are no proportional

correlations between dosage of suppressors or promoters. More detailed ultrastructural analysis using electron microscopy illustrated that altering the balance of promoters and suppressors causes synapsis arrest in both wheat itself and in interspecific hybrids (Holm & Wang, 1988).

1.4.2.3 Molecular models of chromosome pairing in wheat

Recent studies have increased basic understanding of the mechanistic action of *Ph* genes on chromosome behaviour in allohexapoid wheat. Of particular interest is the recent comparison of *Ph1* to *Ph2*, suggesting two different mechanisms resulting in homeologous chromosome pairing (Martinez *et al.*, 2001). Accordingly several hypotheses have progressed to explain the observed chromosome behaviour.

1.4.2.3.1 Spatial distribution of chromosomes

Feldman and colleague's (Feldman & Avivi, 1988) have suggested that homologous chromosomes pair up before the onset of meiosis (described in Section 1.4.2.1.1). Discrimination between homologues and homeologues occurs prior to meiosis via effects on somatic pairing related to spatial distribution within the nucleus. It is anticipated that the *Ph* genes directly or indirectly influence this spatial distribution and therefore chromosomes would not be completely dispersed within the nucleus. In a haploid cell the chromosomes would be spatially separated, lying side by side or around each other (Schwarzacher, 1997), unlike a diploid cell where the homologues and homeologues would be associated and have relatively fixed positions in relation to each other (Flavell *et al.*, 1987). It was also suggested that there may be "elastic connectors" controlling the position of chromosome homologues or homeologues (Maguire, 1977). This has led to suggestions of the possible presence of intranuclear fibrillar material that has the same influence on the positioning of the chromosomes (Bennett & Smith, 1979). This is certainly the case for dipterans such as *Aedes aegypti* and some higher plants such as maize (Maguire, 1983). However, this is untrue for plants such as *Lilium* (Del Fosse & Church, 1981; Walters, 1970). Weiner and Kleckner (1994) have observed in yeast, pre-meiotic homologous chromosome associations but these

weaken as DNA replication commences. As these cells undergo meiosis the chromosomal associations re-establish. Wheat has been shown to behave slightly differently to other eukaryotes in that centromeres of homologous chromosomes become associated first during floral development. At pre-meiotic interphase over 90% of homologues are already completely associated with their pairing partner (Moore, 1998). Somatic chromosome disposition seems to vary among eukaryotes and no strict rule applies to all.

1.4.2.3.2 Timing of chromosome interactions

Hobolth (1981) and Von Wettstein *et al.*, (1984) suggest that timing of crossovers and chromosome exchanges after diploidisation is the key factor controlled by the *Ph* genes. Hobolth (1981) observed that multivalent synaptonemal complexes regularly form at zygotene but become corrected into strict bivalents at pachytene in euploid wheat. However, with six copies of *Ph1* (trisomic 5BL) the number of multivalents increased. Following these observations, Hobolth believed that with two copies of *Ph1* (wild type), crossing over is delayed until after multivalents are corrected into bivalents. However, with six copies of *Ph1* crossing-over is delayed until diplotene when the synaptonemal complexes begin to degrade and when the condition for recombination has become sub-optimal. Holm and Wang (1988) disagree with this mechanistic model for *Ph1* based on observations in monosomic 5BL and occasionally in euploid wheat, where multivalent associations sometime persist through the crossing-over interval yet chiasmata are not formed between homeologues. Holm and Wang (1988) also argue that it is expected that there would be a reduction of chiasmata in trisomic 5BL, however there is pairing arrest.

1.4.2.3.3 Stringency of both synapsis and homologous exchanges

It has been proposed that the observed chromosome behaviour of mutants at *Ph* loci is a result of alterations in the control of stringency of both synapsis and homologous exchanges (Holm & Wang, 1988). It was also proposed that the control is likely to lie within the DNA mismatch repair system (Holm & Wang, 1988) as crossover suppression in bread wheat can be attributable, at least in

part, to the fact that homeologue DNA affects the basic biochemistry of the recombination process. This is supported by the fact that recombinant chromosomes containing homeologous or alien segments, upon undergoing synapsis, have a reduced number of chiasma between the homologous and homeologous/alien segments (Liharska *et al.*, 1996; Luo *et al.*, 1996). This suggests that homology is processed along the entire length of the chromosomes yet there is disruption at specific sites where there are likely to be recombination interactions.

1.4.3 Improvement of wheat breeding by the control of the *Ph* genes

Plant breeding essentially relies on and makes use of genetic variation, which occurs naturally within and between members of a family, a genus, a species or a subspecies, or through the introduction of genes from other organisms, which may or may not be related to the host plant.

From a plant breeder's perspective, the limits of homeologous recombination define a genetic barrier to the introduction of new and desirable traits into existing plant varieties. This means that land races and related species such as rye are hard to use for wheat improvement. Problems with chromosome pairing and recombination limit the ability of breeders to exploit useful genes through introgression. Homeologous recombination is thus of great importance for plant breeding. Accordingly there is a need for a process of enhancing the frequency of homeologous recombination in plants. In particular, there is a need for a process of increasing homeologous recombination to significantly shorten the length of breeding programs by reducing the number of crosses required to obtain an otherwise rare recombination event.

Understanding the action of the *Ph2* gene for alien gene introgression aims to help increase the frequency of recombination between the genomes. The *Ph* genes of *T. aestivum* maintain a strong barrier to gene flow between crossable species having homeologous chromosomal relationships (Ceoloni & Donini, 1993). Increasing the frequency of recombination will reduce capital costs for wheat breeding programs by reducing the numbers of plants required within

breeding populations. The introduction of alien traits into *T. aestivum* would help increase the genetic base from which breeders can work.

Breeders currently use *ph1b*, a deletion mutant at the *Ph1* locus of Chinese Spring (Sears, 1977) in their alien gene introgression crosses. The effectiveness of the *ph1b* mutant has been demonstrated in wide crosses (Gale & Miller, 1987) but in some cases it's usefulness is limited. With crosses between wheat (lacking *Ph1* locus) and rye, as well as crosses between wheat and barley, desired levels of pairing and recombination are laborious and difficult to obtain (Islam & Shepherd, 1988; Koebner & Shepherd, 1985). To solve the problem of such crosses their needs to be an understanding of the molecular mechanisms involved in chromosome pairing and recombination. Once understood, the system can then be manipulated to increase the efficiency of alien gene introgression.

1.5 Aims of this study

This study primarily aims to outline a systematic method for the analysis of genes likely to be involved in meiosis, in particular, those genes that could possibly represent the *Ph2* gene. Specifically structural analysis of the *W*heat *M*eiosis 1 (*WM1*) gene family in relation to its localition within the region deleted in the *ph2a* wheat mutant was examined. In addition, an attempt at functional analysis of the *W*heat *M*eiosis 5 (*WM5*) gene through the use of biolistic wheat transformation was conducted. Finally comparative mapping among the grasses aims to use the currently available rice genome sequence to identify those genes that may represent the *Ph2* gene. Through the structural, functional and comparative analysis of genes within the grasses, this study aims to increase our awareness of grass genome structure and meiosis.

What are the gene(s) controlling the *P*airing *H*omeologous (*Ph*) phenotype and how do they work? This question has been studied for 40 years and it is anticipated that the approaches taken in this study will shed light on this unanswered question.

2 Chapter

GENERAL MATERIALS AND METHODS

2.1 Materials

Materials used in this study are listed below, together with the suppliers' names. All chemicals for *in vitro* use were at least analytical grade. Where appropriate, solutions were prepared using MilliQ H_2O, and subsequently autoclaved. Descriptions of RFLP clones and genetic material used in this study can be found in the individual chapters.

2.1.1 Chemicals

Bovine serum albumen (BSA) fraction V, spermidine, ampicillin, kanamycin, salmon sperm DNA, N-(2-hydroxyethyl) piperazine-N'-(2-ethane-sulfonic acid (HEPES), Tris (hydroxymethyl) amino-methane (trizma base), ethidium bromide, poly vinyl pyrillidone (PVP, 40,000 molecular weight), *Eschericia coli* t-RNA, salmon sperm DNA, dithiothreitol (DTT): Sigma Chemicals (USA).

Dextran sulphate, ficoll 400: Pharmacia (USA)

Phenol: Wako Industries (Japan).

NaCl, NaOH, Na_2EDTA, $MgCl_2$, potassium acetate (KOAc), sodium acetate (NaOAc), urea, sucrose, glucose, ethanol (EtOH), iso-propyl alcohol, iso-amyl alcohol, chloroform, bromophenol blue, HCl, glacial acetic acid, sodium dodecyl sulphate (SDS): BDH.

Xylene cyanol: Ajax Chemicals.

Oligodeoxyribonucleotides: Synthetic oligodeoxyribonucleotides were either made on an Applied Biosystems (USA) Model 380B DNA synthesiser by Neil Shirley in the Department of Plant Science, University of Adelaide or through

Parallel Array Synthesis (Invitrogen, Australia) comprising of standard cyanoethyl phosphoramidite chemistry. Subsequently oligonucleotides were purified by ion exchange HPLC using a MonoQ column (Pharmacia, USA).

Nucleotides and Radionucleotides: Ultrapure nucleotide triphosphates (NTPs) and deoxynuclotide triphosphates (dNTPs) were obtained from Pharmacia. [α-^{32}P] dATP (10 μCi/μl), and [α-^{32}P] dCTP (10 μCi/μl), were obtained from Bresatec and Amersham (Australia).

2.1.2 Enzymes

Mung bean nuclease: New England Biolabs.

Calf intestinal alkaline phosphatase (CIAP): Boehringer Mannheim (Germany).

Pancreatic RNase A: Sigma (USA).

Taq DNA polymerase, Klenow fragment (large fragment of *Eschericia coli* DNA polymerase I), Thermoscript and Superscript Reverse Transcriptase: Life Technologies (Rockville, MD, USA).

Pfu DNA polymerase: Stratagene (La Jolla, CA)

Restriction enzymes: Bresatec (Australia), Boehringer Mannheim (Germany), New England Biolabs and Promega (USA).

2.1.3 Molecular weight markers and cloning vectors

λ DNA cut with *Hind III* and λ DNA cut with *Bst EII/Sal I*.

pBluescript SK(+) : Stratagene (USA).

Bacterial strains: *Escherichia coli* DH5α: Stratagene (USA).

Bacterial media ingredients: bacto-agar, bacto-tryptone and yeast extract: Difco Laboratories (USA).

2.1.4 Agaroses

Low melting point agarose: BRL (USA).

Other agarose (NuSeive GTG grade): FMC Bioproducts (USA).

2.1.5 Kits

Bresa-Clean: Bresatec (Australia).

Qiagen tip-20: Qiagen (Australia)

Qiagen gel extraction kit: Qiagen (Australia).

Qiagen PCR clean-up kit: Qiagen (Australia)

2.2 Methods

Methods were carried out according to standard procedures (Sambrook *et al.*, 1989) or using manufacturers specifications (except where cited in text). Methods used routinely throughout this study are described below. Specific methods that were used only in particular parts of this study are described in the individual chapters.

2.2.1 Plant growth conditions

Potting soil was prepared by the plant growth facility at the Waite Institute (University of Adelaide). Unless otherwise stated, recycled soil made from discarded soil and plant material was used. UC (University of California) soil mix consisting of four parts washed river sand and three parts (dry volume) moss peat was used when a more consistent soil was required. Both types of soil were steam treated to reduce the viability of contaminating seeds and microorganisms. Unless otherwise stated, plants were grown in 13 to 25 cm pots in the glasshouse, at 18 to 25°C.

2.2.2 Growth of bacteria

Cultures of *E. coli* bacteria were grown overnight at 37°C, using solid or liquid media. Solid media was prepared by dissolving bacterialogical agar (1.5% w/v)

in boiling LB broth (1.0% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl, pH 7.0). Liquid cultures were grown in 10 mL tubes containing 3.0 mL LB broth or 2YT broth (1.6% (w/v) bacto-tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl, pH 7.0). Each liquid culture was inoculated using a single bacterial colony or a scraping from a frozen glycerol culture, and grown on a rotator. The antibiotic ampicillin or kanamycin was added to bacterial growth media at concentrations of 100 μg mL^{-1} or 50 μg mL^{-1} respectively.

2.2.3 Transformation of *E. coli* with plasmids by electroporation

Procedures used to prepare electrocompetent cells were those supplied with the Gene-Pulser (Bio-Rad, USA). One litre of LB culture was inoculated using 10 mL of an overnight culture of *E. coli* strain DH5α (BRL, USA), grown to an optical density (λ600) of 0.9) with the flask chilled on ice for 15 to 30 min. The culture was transferred to 200 mL tubes and the cells pelleted in a GSA rotor at 3,000 rpm for 15 min at 4°C. The supernatant was discarded and the cells gently resuspended in 0.5 L of ice-cold 10% glycerol solution. The cells were then pelleted as above, the supernatant discarded, and resuspended in 20 mL of ice-cold 10% glycerol solution. Cells were transferred to 30 mL tubes, pelleted in a HB4 rotor at 4,000 rpm for 15 min at 4°C, and resuspended in 2.0 mL of ice-cold, 10% glycerol solution. The electrocompetent cells were transferred to 1.5 mL Eppendorf tubes in aliquots of 140 μl, snap frozen in liquid nitrogen, and stored at –80°C until use.

Transformation of electrocompetent cells with plasmids was performed according to the recommendations supplied with the Gene-Pulser. Electrocompetent cells (40 μl) were combined with 1.0 μL MilliQ H_2O containing 5.0 ng of plasmid DNA or 60 ng of DNA from a ligation reaction. The mixture was transferred to an ice-cold, disposable electroporation cell (0.1 cm electrode gap, supplied with the Gene-Pulser), and subject to electroporation using a Gene-Pulser (Bio-Rad), set at 1.8 kV, 125 μFD and 200 Ω. Immediately following electroporation, the cells were mixed with 1.0 mL LB broth without antibiotic, and grown at 37°C in a 1.5 mL Eppendorf tube for 1 hr on a shaker.

Two-hundred μL aliquots of culture were then plated onto solid media containing antibiotic, and grown at 37°C overnight.

2.2.4 Mini-preparation of plasmid DNA

The protocol used for plasmid DNA isolation was essentially the procedure described by Sambrook *et al.* (1989) for the small-scale isolation of plasmid DNA by alkaline lysis. All steps were performed at room temperature unless otherwise stated, and centrifugations were performed in an Eppendorf 5415C bench centrifuge. Liquid culture of plasmid-containing bacteria was used to fill a 1.5 mL Eppendorf tube and the cells pelleted by centrifuging at 14,000 rpm for 30 sec. The supernatant was discarded and the cells resuspended in 100 μL ice-cold Plasmid I buffer (50 mM glucose, 25 mM Tris-HCl, 10 mM Na_2EDTA, pH 8.0) by vortexing for 1 min. Two-hundred μL of freshly made, ice-cold 0.2 M NaOH, 1.0% SDS was added and mixed in by gentle inversion of the tube. One-hundred and fifty μL of 3.0 M KOAc, 11.5% glacial acetic acid was added to the lysate and the tube vortexed gently before placing on ice for 3 to 5 min. The tube was then centrifuged for 5 min at 14,000 rpm and 350 μL of the supernatant transferred to a new tube. Nucleic acid was precipitated by combining the supernatant with 35 μL of 3.0 M NaOAc, pH 5.2, and 400 μL ice-cold ethanol, and pelleted by centrifugation at 14,000 rpm at 4°C. Pellets were washed in 1.0 mL ice-cold 70% ethanol, dried completely, and resuspended in 20 μL of 10 mM Tris-HCl, 0.1 mM Na_2EDTA, pH 8.0, containing 40 mg mL^{-1} pancreatic RNase A.

2.2.4.1 PCR amplification of cloned insert DNA

The cloned insert DNA was amplified using the polymerase chain reaction with M13 forward and reverse primers. The oligonucleotide primers M13 -40P (5'- CAG GGT TTT CCC AGT CAC GAC –3') and M13 RSP (5'- ACA GGA AAC AGC TAT GAC CAT G –3') for clones in the plasmid vectors *pBluescript SK(+)*, *pUC18*, *pUC19* and pUC119, or the primers SP6 (5'- GAT TTA GGT GAC ACT ATA G –3') and T7 (5'- TAA TAC GAC TCA CTA TAG GG –3') for clones in *pGEM-4*. These primers were synthesised using an Applied Biosystems 392

oligonucleotide synthesiser according to the manufacturers instructions. The reaction mix contained 50 mM Tris HCl, 6.5 mM $MgCl_2$ (25 mM), 48 μM each dNTP, 0.3 μg each primer, 1 μg plasmid DNA and 1 unit *Taq* polymerase in a total volume of 50 μl. Temperature conditions for PCR were 95°C for 3 min, followed by 35 cycles of 96°C for 1 minute, 55°C or 58°C for 2 min, 72°C for 2 min with a final step of 72°C for 5 min and 25°C for 5 min. This reaction was performed in a PTC-150 Mini Cycler (MJ Research, USA). A $MgCl_2$ concentration of 1.3 mM was effective for the amplification of most inserts, although inserts larger than 1.5 Kb generally required the $MgCl_2$ concentration to be optimised.

2.2.4.2 Recovery of insert DNA from agarose gels using glass-milk and Qiagen gel extraction kit

Following gel electrophoresis of the PCR product, the required band was excised from the gel after staining with ethidium bromide and visualising on a long wave ultraviolet transilluminator (340 nm). The excised agarose block was transferred to a 1.5 mL eppendorf tube and weighed. The DNA was recovered according to the standard protocol supplied with the Geneclean kit (Bio 101) [Bresatec, Australia] or with the standard protocol that was supplied with Qiagen gel extraction kit (Qiagen, Australia).

2.2.5 Large scale preparations of plasmid DNA

The following method was used to purify large amounts (150 μg) of plasmid DNA. A plasmid-containing bacterial culture was grown overnight to stationary phase (~16 hr) in 400 mL of LB containing appropriate antibiotic(s), in a baffled 2 litre flask at 37°C. Cells were sedimented by centrifugation (5,000 rpm, 15 min, 4°C, Sorvall GSA rotor) and washed in STE buffer (50 mM NaCl, 1 mM EDTA, 25 mM Tris-HCl, pH 8.0). The cells were pelleted by centrifugation as before then resuspended in 4 mL of Plasmid I buffer containing 1 mg mL^{-1} lysozyme. The resuspended cells were incubated on ice for 10 min, before the addition of 8 mL of freshly prepared 0.2 M NaOH, 1% SDS, followed by gentle mixing. After incubation on ice for 10 min the mixture was centrifuged as

previously except that centrifugation was at 10,000 rpm for 10 min. The supernatant was removed and strained through four layers of cheesecloth, before precipitation of nucleic acids by the addition of 12 mL of ice-cold 2-propanol. The pellet was washed with ice-cold 70% ethanol, after which it was dried in air at RT before resuspension in 3 mL TE. LiCl (10 M) was added to a final concentration of 2.5 M to precipitate RNA, and the solution placed on ice for 10 min. RNA was pelleted by centrifugation in a Sorvall HB4 rotor at 10,000 rpm for 15 min at 4°C. The supernatant was removed and DNA precipitated by addition of an equal volume of ice-cold 2-propanol, followed by centrifugation as previously described to pellet DNA. The pellet was washed in ice-cold 70% ethanol and air-dried at RT. DNA was resuspended in 400 μL TE and transferred to a micro-centrifuge tube. RNase A was added to 20 μg mL^{-1} and the mixture incubated at 37°C for 1 hr. The solution was extracted twice with phenol:chloroform (2.2.6), and once with chloroform to remove proteins. Plasmid DNA was precipitated from solution by the addition of an equal volume of 13% PEG 8000, 1.6 M NaCl, followed by incubation at RT for 5 min. DNA was recovered by centrifugation at full speed in a bench micro-centrifuge for 5 min at RT, and the pellet resuspended in 400 μL TE. DNA was again precipitated from solution by addition of 3 M sodium acetate pH 4.6 to a concentration of 0.3 M, and 2.5 volumes of ice-cold ethanol. The DNA was pelleted by centrifugation as previously described, and the pellet washed in 400 μL ice-cold 70% ethanol before re-centrifugation. The pellet was dried and resuspended in 400 μL TE.

2.2.6 Phenol:chloroform extraction and ethanol precipitation of DNA

DNA solutions were mixed with one volume of phenol:chloroform (containing one volume of redistilled phenol (BDH, Australia) equilibrated in 50 mM Tris-HCl, pH 8.0, and one volume of chloroform) and centrifuged for 10 min at RT (full speed in an Eppendorf micro-centrifuge for small quantities, or 10,000 rpm in a Sorvall HB4 rotor for larger solutions). The aqueous phase was recovered and the extraction repeated as necessary.

DNA was routinely precipitated from solutions with ethanol. Briefly, 1/10th volume of 3 M sodium acetate (pH 4.6) was added followed by 2.5 volumes of ice-cold ethanol. The solutions were incubated on ice for 15 min, followed by centrifugation at high speed at RT for 15 min in an Eppendorf micro-centrifuge for small volumes, or at 10,000 rpm at 4°C for 15 min in a Sorvall HB4 rotor for larger volumes. Pellets were washed in 70% ethanol prior to drying in a speedyvac or on the bench at RT.

2.2.7 Agarose gel electrophoresis

Large-scale agarose gels were cast from 100 mL of 0.9 to 1.3% (w/v) molten agarose solution containing 1 x TAE buffer (0.04 M Tris-acetate, 1.0 mM Na_2EDTA, pH 8.0), using a 15 x 20 cm mould and a comb for making wells of 15 μL volume (for Southern analysis) or 30 to 50 μL volume (preparative electrophoresis). DNA samples were mixed with 0.2 volume 6 x FLB loading buffer (15% (w/v) ficoll 400, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) and electrophoresed overnight at 40 to 50 V and 18 to 50 mA in 1 x TAE buffer. DNA size markers made from *Hind III* cut λ DNA or *Bst EII*/*Sal I* cut λ DNA (0.2 to 0.5 μg) were run alongside the sample DNAs when needed. Gels were soaked in 0.5 mg/l solution of ethidium bromide for 10 min, de-stained by rinsing in water for 10 min and photographed for future reference using UV light at 302 nm (preparative gels) or 260 nm (gels for Southern analysis).

Mini agarose gels were cast by pouring 15 mL of 0.9 to 1.5% (w/v) molten agarose solution containing 1 x TAE buffer (0.04 M Tris-acetate, 1.0 mM Na_2EDTA, pH 8.0) onto a 6.0 x 7.0 cm glass plate with the appropriate comb set above it. DNA samples were mixed with 0.2 volumes of 6 x FLB loading buffer (15% (w/v) ficoll 400, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) prior to electrophoresis at 80 to 180 mA in 1 x TAE buffer. Mini agarose gels were stained and photographed as described for large-scale agarose gels.

2.2.8 Polyacrylamide gel electrophoresis

Denaturing polyacrylamide gels were prepared from 50 mL solutions containing 6% (w/v) acrylamide, 0.3% (w/v) bisacrylamide, 7 M urea, and 1xTBE.

Polymerisation was initiated by the addition of 400 μL freshly prepared 10% (w/v) ammonium persulfate and 40 μL of TEMED. The polymerising solution was poured into gels of 20 x 40 x 0.04 cm, with well formation by a shark toothcomb. Gels were allowed to set for at least 60 min, and then pre-electrophoresed at 50 W until gel temperature was approximately 50°C. Gels were electrophoresed at 50°C at constant power after loading and denaturing of samples in formamide loading solution (95% (v/v) formamide, 0.1% bromophenol blue, 0.1% xylene cyanol, 10 mM EDTA).

2.2.9 Cereal genomic DNA preparation

The method for genomic plant DNA preparation is based on Guidet *et al.*, (1991). All centrifugations were performed at 14,000 rpm in an Eppendorf bench centifuge at RT.

2.2.9.1 Small-scale genomic DNA preparation

The method used for the small-scale extraction of DNA from leaves was modified from Guidet *et al.*, (1991). A 10 cm long piece of healthy leaf was placed in a 2 mL Eppendorf tube and frozen in liquid nitrogen. The sample was then crushed with a small pestle to a fine powder after which 600 μl DNA extraction buffer (1% sarkosyl, 100 mM Tris-HCl, 100 mM NaCl, 10 mM EDTA, pH 8.5) was added and homogenised with the leaf powder to form a slurry. Extraction was performed by adding 600 μL of cold phenol solution (phenol:chloroform:iso-amylalcohol (25:24:1) equilibrated with 1 M Tris-HCl, pH 8.0) and mixing on an orbital rotor for 10 min. The sample was centrifuged for 10 min at 20,160 x g and the supernatant transferred to a fresh tube to repeat the phenol extraction step. After the supernatant was collected, 60 μL of 3 M NaAc (pH 4.8) and 600 μL isopropanol was added and mixed gently at room temperature to allow the DNA to precipitate. The DNA was then pelleted by centrifugation for 5 min at 20,160 x g and the supernatant discarded. After washing the pellet with 1 mL 70% ethanol, the DNA was air-dried and resuspended overnight at 4°C in 50 μL R40 (40 μg mL^{-1} RNAse A in TE buffer).

2.2.9.2 Medium scale genomic DNA preparation

For medium scale genomic DNA preparation approximately 2 g of the harvested leaf material was crushed to a fine powder in liquid nitrogen using a mortar and pestle. When the powder had partially thawed, 4 mL of extraction buffer was added and mixed to form a slurry. The slurry was transferred to a 10 mL plastic tube together with 4 mL of cold phenol solution and mixed on an orbital rotor for 10 min.

After centrifugation for 10 min at 12,900 g the supernatant was dispensed into a silica matrix tube and re-extracted with 4 mL phenol/chloroform/iso-amylalcohol (25:24:1). The tubes were centrifuged again for 10 min at 2060 g and the extraction step repeated. After the third extraction the supernatant was dispensed into fresh 10 mL plastic tubes to which 400 μL 3 M NaAc (pH 4.8) and 4 mL isopropanol were added. The tubes were mixed gently to precipitate the DNA and centrifuged for 10 min at 12,900 g to pellet the DNA. The supernatant was poured off and the pellet washed with 4 mL 70% ethanol, air-dried and resuspended overnight at 4°C in 350 μL R40 (Guidet *et al.*, 1991).

To calculate the concentration of the DNA, 10 μL of the sample was added to 990 μL TE buffer and vortexed before reading the absorption in a spectrophotometer at A^{260} (concentration μg μL^{-1} = A^{260} x 100 x 50/ 1000).

2.2.10 DNA restriction, electrophoresis and Southern transfer

For RFLP analysis, 7.0 μL of genomic DNA preparation (approximately 3.0 μg DNA) was digested at 37°C for 3 to 5 hrs in 11 μL reactions containing 1.0 mg mL^{-1} BSA, 1.3 mM spermidine and 1 x concentration of the buffer supplied with the enzyme. Genomic DNA digests were electrophoresed on 1.3% agarose gels (Section 2.2.7), until the bromophenol blue had run 3/4 of the gel length. DNA transfer procedures used were essentially those supplied by Amersham with the Hybond N+ membrane. Each gel was shaken gently in 300 mL denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 30 min, rinsed briefly in water, and shaken gently in 300 mL neutralising solution (1.5 M NaCl. 0.5 M Tris-HCl, 1.0 mM Na_2EDTA, pH 7.2) for a further 30 min. DNA was transferred from the gels

to Hybond N+ membrane (Amersham) by Southern transfer (Southern, 1975) overnight, using 20 x SSC (3.0 M NaCl, 0.3 M trisodium citrate, pH 7.0) as the transfer buffer. DNA was fixed to the membrane by placing the membrane DNA side up on a pad made from three sheets of Whatmann 3MM paper, soaked in 0.5 M NaOH. After 20 min, membranes were shaken in a solution of 2 x SSC for at least 5 min.

2.2.11 Preparation of ^{32}P-labelled DNA based probes

Radioactively labelled probes were synthesised by random priming (Feinberg & Vogelstein, 1983). Purified cloned insert DNA (20 ng) was combined with 6.0 μL random sequence 9-mer oligonucleotide (0.1 mg mL^{-1}), and the mixture incubated at 95°C for 3 min to denature the DNA. The mixture was cooled on ice for 5 min and combined with 10 μL probe labelling buffer (0.5 M HEPES, 0.125 M Tris-HCl, 12.5 mM DTT, 12.5 mM $MgCl_2$, 1.0 mg mL^{-1} BSA)[1], 2.5 μL dNTP mixture (0.2 mM dATP, 0.2 mM dTTP, 0.2 mM dGTP), 1.0 U Klenow fragment, 3.0 μL [α-^{32}P]dCTP (10 μCi/μl), and enough MilliQ H_2O to make the total volume 25 μL. This mix was incubated for 1 hr at 37°C and then passed through a Sephadex G-100 mini-column saturated with TE buffer to separate the labelled probe from the unincorporated [α-^{32}P]dCTP. A total volume of 200 μL salmon sperm DNA (5 mg mL^{-1}) was then added to the sample.

2.2.12 Hybridisation and autoradiography

The pre-hybridisation and hybridisation procedures used were based on protocols supplied with the Hybond N+ membrane (Amersham). Each 10 mL of pre-hybridisation/hybridisation solution was made by combining 3.0 mL of 5 x HSB solution (3.0 M NaCl, 0.1 M PIPES, 25 mM Na_2EDTA, pH 6.8), 2.0 mL of 50 x Denhardt's solution (2.0% (w/v) BSA, 2.0% (w/v) ficoll 400, 2.0% (w/v) PVP), 3.0 mL 25% (w/v) dextran sulphate, 2.0 mL MilliQ H_2O, and 200 μL of 10 mg mL^{-1} salmon sperm DNA which had been denatured at 95°C for 10 min prior to its addition. Membranes were placed in hybridisation bottles containing 5.0 to

[1]The Tris-HCl and HEPES stock solutions used to make the probe labelling buffer were 1 M, pH 8.0 and 0.8 M, pH 6.6, respectively.

20 mL of pre-hybridisation/hybridisation solution (up to 10 membranes per bottle) and pre-hybridised for 3 to 5 hr at 65°C. Following pre-hybridisation, the probe(s) was denatured by heating for 10 min at 95°C, cooled on ice for 5 min, added to the hybridisation mixture, and hybridisation performed at 65°C for 12 to 20 hrs. Membranes were then washed (i) three times for 5 min in 2.0 x SSC, 0.1% SDS at 65°C, and then (ii) two times for 10 min in 0.2 x SSC, 0.1% SDS at 65°C. Two final washes of 10 min in 1.0 x SSC, 0.1% SDS at 55°C were used instead of step (ii) when probes derived from non-wheat species were used. Autoradiography was performed for 5 to 25 days at –80°C with Konica-HR X-ray film and an intensifying screen.

Following autoradiography, the probes were stripped from the membranes by placing the membranes in a container with between 0.5 - 1.0 L of boiling 0.1% SDS solution. The container was left on a shaker at room temperature for at least 10 min before replacing the solution with fresh 0.1% SDS at room temperature. Membranes were stored in this solution at 4°C until reused.

2.2.13 Total plant RNA isolation

Total plant RNA was isolated using Trizol reageant (Gibco-BRL, Australia) from approximately 1 g of tissue according to manufacturers specifications. The isolated RNA pellet was washed in 70% ethanol and then sedimented, air-dried and dissolved in 1 x TE buffer. Quality and concentration was determined using a spectrophotomer and gel electrophoresis.

2.2.14 Northern blot hybridisation

Gel preparation: 2.16 g of agarose was boiled in 156.6 mL of water and cooled to 60°C. A volume of 5.4 mL 37% formaldehyde (BDH, Australia) and 18 mL of 10 x MOPS buffer [0.23 M MOPS (3-N-Morpholino) propanesulfonic acid; Sigma, Australia) (pH7.0), 0.01 M EDTA and 0.05 M NaAC] were then added before the gel was poured into a gel box (14 x 2 cm). The running buffer was 1 x MOPS.

Sample preparation: 5 μg of RNA, 2.5 μL of 10 x MOPS, 4.4 μL formaldehyde and 12.5 μL formamide (BHD, Australia) were incubated at 65°C for 15 min. A 5 μL volume of formamide loading buffer (FLB: 95% formamide, 20 mM EDTA, 0.5% bromophenol blue, 0.05% xylene cyanol) was then added before loading. The gel was run at 6 V/cm until the first bromophenol blue was approximately 3/4 of the gel length.

Northern blotting and hybridisation: Before transfer, the gel was soaked in 10 x SSC for 15 min. The RNA transfer sandwich consisted of the following layers (from bottom to top): 4 sheets of Whatmann 3MM pre-soaked with 10 x SSC, the gel (upside down), the nylon membrane (Hybond N+, Amersham), 2 sheets of Whatmann 3MM pre-soaked with 10 x SSC and dry paper towels. After overnight transfer, the sandwich was disassembled. RNA on the membrane was fixed by UV light (GS Genelinker, BIORAD) and photographed under short wavelength UV light. Pre-hybridisation was performed at 45°C for a minimum of 8 hrs in a bottle containing 20 mL of solution consisting of 3 mL 50 x Denhardts reagent, 5 mL 20 x SSC, 1 mL 10% SDS, 2 mL carrier DNA (10 mg mL^{-1}) and 9 mL of deionised formamide (BDH, Australia). The hybridisation was performed at 45°C in a bottle containing 20 mL of hybridisation solution consisting of 2 mL of 50 x Denhardts reagent, 5 mL of 20 x SSC, 1 mL of 10% SDS, 1 mL of carrier DNA (10 mg mL^{-1}) and 9 mL of deionised formamide (BDH, Australia) and 2 mL of dextran sulphate (25%). Membranes where then washed, subjected to autoradiography and stripped and stored as described for Southern analysis in Section 2.2.12.

2.2.15 Purification of DNA clone inserts

To excise insert DNA from DNA clones, 10 μL of plasmid mini-preparation was digested for 3 hr at 37°C, in 20 μL reactions containing 20 U of the appropriate restriction enzyme, 1.0 mg mL^{-1} BSA, 1.3 mM spermidine and 1 x concentration of the buffer supplied with the enzyme. Alternatively the cloned insert was amplified via PCR as outlined in Section 2.2.4.1.

Plasmid digests and PCR reactions were electrophoresed in large scale agarose gels (Section 2.2.7), the bands corresponding to the insert DNA excised, and the DNA purified from the gel slices using either Bresa-Clean kit (Bresatec, Australia) or Qiagen gel extraction kit, according to the manufacturer's instructions. Samples of the purified clone insert DNA preparations were electrophoresed on mini agarose gels (Section 2.2.7), alongside known quantities of DNA markers in order to verify the sizes of the inserts and to estimate concentration.

3 Chapter

STRUCTURAL ANALYSIS OF THE *WM1* GENE FAMILY

3.1 Introduction

A partial cDNA sequence from wheat meiocytes encoding a leucine rich repeat (LRR)-containing protein has previously been described (Ji & Langridge, 1994). This cDNA was termed *AWJL3* (Ji & Langridge, 1994) and is now described as *W*heat *M*eiosis 1.0 (*WM1.0*). The encoded protein has three 24 amino acid long leucine-rich repeats at the N-terminus and a leucine-zipper-like domain at the C-terminus. The *WM1.0* cDNA is interesting, as it represents a family of 21 genes of which certain members have meiosis specific expression (Ji, 1992). The genes identified by the *WM1.0* cDNA clone have been named *WM1.1* to *WM1.21* (*W*heat *M*eiosis genes 1.1 to 1.21). A few members of the family have been mapped to the chromosome groups three and five in bread wheat (*T. aestivum* L.). These chromosomes are known to contain several genes involved in the control of chromosome pairing in wheat. The strongest effects on pairing are shown by *Ph1* on the long arm of chromosome 5B (5BL) and *Ph2* on 3DS. Both genes are suppressors of chromosome pairing while *Ph3*, on 5BS, is a promoter of pairing (Sears, 1972). The *WM1.0* cDNA clone previously identified three or four genes on the short arm of chromosome 3DS all of which were found to be missing in wheat lines with a deletion at the *Ph2* locus (Ji, 1992). The deletion mutant of *Ph2*, *ph2a*, shows altered meiotic chromosome pairing in interspecific hybrids (Sears, 1982). On this basis it was proposed that one or more of these genes may represent the *Ph2* locus (Ji & Langridge, 1994).

This study looks specifically at those *WM1* genes that seem to be localised on chromosome 3DS in the vicinity of the *Ph2* locus. Two of these genes are termed *WM1.1* and *WM1.2*. Earlier RT-PCR studies suggested that *WM1.1* and *WM1.2* have peak expression at leptotene and diplotene/diakinesis respectively (Ji, 1992). The chromosome localisation and expression patterns of putative

meiosis-specific genes suggest that they are important in the developmental process of meiosis. This chapter attempts to describe the structure and confirm expression of these *WM1* family members.

3.2 Materials and methods

3.2.1 Genetic stocks

Seeds of euploid wheat (*Triticum aestivum* cv. Chinese Spring) and aneuploids were obtained from Dr Ken Shepherd, Waite Agricultural Research Institute. The *ph2a* and *ph2b* mutants were obtained from Prof. Moshe Feldman, Plant Genetics Institute, Israel.

Plants were grown in a glasshouse and leaves of young seedlings were harvested for DNA extraction. Plants used for anther collection were kept under controlled growth conditions with a 16 hr photoperiod at 18°C/14°C.

Wheat nullisomic-tetrasomic lines were obtained from Terry Miller (John Innes Centre, Norwich). Each nullisomic-tetrasomic line is deficient for one of the 21 wheat chromosomes, with the missing chromosome compensated by the addition of an extra copy of one of its homeologous chromosomes from a different genome. Four seeds of each line were grown in small pots in the glasshouse at 25 °C for several weeks before leaf material was collected for DNA extraction. The DNA extracted from these lines were used to assign the *WM1* gene family members to specific chromosomes.

3.2.2 DNA sequence and partial genomic clones of *WM1* genes

The DNA sequence and partial genomic clones for *WM1* genes were provided for this study (Ji, 1992). The genomic clones were isolated from two wheat genomic libraries using the cDNA clone *AWJL3* now described as *WM1.0* (Ji & Langridge, 1994), with each library consisting of about $1x10^6$ pfu. Hybridising clones were subcloned and sequenced with clones being designated *WM1.1* to *WM1.21*.

3.2.3 Amplification of *WM1.1*, *WM1.2* and *WM1.3* specific sequences

PCR primers were designed based on partial genomic sequence for *WM1.1*, *WM1.2*, *WM1.3* and *WM1.5* (Ji, 1992). A hypervariable region within the coding sequence for *WM1.1*, *WM1.2*, *WM1.3* and *WM1.5* was amplified from each of the respective genomic clones using primers as described in **Table 1**. These amplicons were termed a "linker" based on the deduced structural characteristics of the peptide sequence in this region. Additionally the whole putative coding sequence for *WM1.1* was amplified from Chinese Spring DNA using primers of the *WM1.1* ORF (**Table 1**).

3.2.4 Nullisomic tetrasomic determined chromosome localisation

The *WM1* family was mapped using Southern membranes of Chinese spring and the full set of nullisomic tetrasomic wheat lines (refer to Section 3.2.1). The DNA on these membranes had been digested with *Eco RV* and *Dra I* and were screened with the *WM1.1* ORF probe (refer to Section 3.2.3) according to the procedures outlined in Sections 2.2.9 to 2.2.12. After chromosome localisation, selected nullisomic tetrasomic, ditelosomic and meiotic mutant wheat lines were analysed by Southern analysis, using the same restriction endonuclease digestions, for sub-chromosomal localisation. These membranes were firstly screened with the *WM1.1* ORF probe and subsequently with the *WM1.1* linker, *WM1.2* linker, and *WM1.3* linker probes (refer to Section 3.2.3).

3.2.5 Pulse field gel electrophoresis

Southern membranes of restriction endonuclease digested Chinese Spring DNA that had been separated by pulse field gel electrophoresis (PFGE) was kindly provided by Dr Ursula Langridge. All protocols used in DNA isolation, digestion and PFGE separation are outlined in Sections 3.2.5.1 to 3.2.5.3

Table 1: DNA sequences of primers designed to amplify hypervariable and coding sequences for members of the *WM1* gene family.

Amplicon	Forward Primer 5'→ 3'	Reverse Primer 5'→ 3'
WM1.1 linker	AGAATTCAATACCTCGGCATCTGT	AGAATTCATATAAGTTGCTGCCCTTT
WM1.2 linker	AGAATTCAATACCTTGTCATCTGT	AGAATTCATATAAGATGTTGTCTTTT
WM1.3 linker	GAGCCACAGTGCACTTTCTGG	TGCCTCGATTATCAGACCTA
WM1.1 ORF	TCATTCAATAGGTTGCAGCAC	TGGGCCTGTAGCATACTCTAGTGT

NB. Conditions for the polymerase chain reaction are outlined in section 2.2.4.1

3.2.5.1 Protoplasts

HMW DNA was isolated from protoplasts obtained from leaf material of 14-day-old Chinese Spring wheat seedlings (Cheung & Gale, 1990). The leaves were removed and surface sterilised for 15 min in 10% Domestos, followed by 4 washes in sterile distilled water. Leaves were then transferred to a petri dish containing 4 mL of protoplast buffer (10 mM $CaCl_2$, 1 μM $CuSO_4$, 0.2 mM KH_2PO_4, 10 μM KI, 1 mM KNO_3, 0.7 M Mannitol, 2 mM MES, 1 mM $MgSO_4$, pH 5.8 with KOH), and 2 mL of enzyme solution. The enzyme solution contained 0.8% Cellulase Onokzuka R-10 (Yakult, Tokyo, Japan), 0.4% Macerocyme Onokzuka R-10 (Yakult, Tokyo, Japan), and 0.04% Pectoylase Y23 (Seishin, Tokyo, Japan) made to volume in protoplast buffer. Leaf tissue was sliced longitudinally into very fine pieces with a scalpel blade and placed in a vacuum for 20 min at 700 mbar. Petri dishes containing leaf material were then removed and incubated at normal pressure at 25°C in the dark for 4 hrs for cell wall digestion. Petri dishes were then transferred to an orbital shaker with gentle shaking (25 rpm) for protoplast release, which was monitored by microscopic observation. Tissue was then flushed with 4 mL of protoplast buffer for further protoplast release, before being sequentially filtered through 40 μm and 20 μm nylon sieves and then pelleted at 26 x g for 10 min. Protoplasts were then washed in protoplast buffer before being pelleted again at 26 x g for 10 min. Protoplasts were resuspended at a final concentration of 4.5×10^6 protoplasts mL^{-1}, mixed with an equal volume of low melt agarose solution containing 2% agarose, 0.17 M EDTA, in protoplast buffer, at 45°C. The molten mixture was aliquoted into plastic moulds (10 mm height × 5 mm width × 1.5 mm thickness) and allowed to solidify at 4°C for 30 min. The blocks were incubated for 48 hrs at 50°C in two changes of lysis buffer (0.5 M EDTA pH 8.0, 10 mM Tris pH 8.0, 1% sarkosyl, 2 mg mL^{-1} proteinase K) followed by 4 washes of 0.5 M EDTA at 50°C. The resulting DNA plugs were then stored at 4°C in 0.5 M EDTA.

3.2.5.2 Restriction enzyme digestion of HMW DNA

Before restriction enzyme digestion, agarose plugs containing HMW DNA were washed 6 times in 10-20 volumes of ice-cold TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0). The first wash for 2 hrs, the second wash overnight, the following 4 washes each for duration of 15 min. The plugs were then ready for restriction enzyme digestion. The agarose plugs were incubated in 500 μL of restriction digestion buffer at 4°C, in the presence of 8 mM spermidine and 2 mg mL^{-1} of BSA for 2 hrs. Restriction enzyme digestions were carried out in 500 μL of the same buffer with fresh reaction mix and 20 units of enzyme at temperatures recommended by the manufacturers, between 4 to 12 hrs. Restriction enzymes used for restriction of HMW DNA were *Not I*, *Sma I*, and *Xho I*. After digestion, the activity of the restriction endonuclease in each sample was stopped by incubation in 0.5 M EDTA. The agarose blocks were then ready for electrophoresis and were stored at 4°C.

3.2.5.3 PFGE separation, transfer and hybridisation

The digested HMW DNA was loaded onto a 0.5 × TBE 1% agarose gel and sealed in position using the same agarose as the gel. Pulsed-field gel electrophoresis was performed using a CHEF DRII (Bio-Rad, USA) run at 14°C and 200 volts. Pulse and run times used to separate HMW DNA are presented in **Table 2**. The DNA was stained with ethidium bromide, photographed, and nicked using 60 mJoules of UV light in the GS gene linker (Bio-Rad, USA). The DNA was blotted onto Hybond-N^+ membrane (Amersham, USA) in 0.4 M NaOH instead of 20 × SSC as described in 2.2.10. ^{32}P labelled probe preparation and hybridisation conditions were as described in Sections 2.2.11 and 2.2.12 respectively. Kodak double emulsion BIOMAX MS film and a BIOMAX MS intensifying screen were used for PFGE autoradiography. Filters were exposed from 24-144 hrs at -80°C.

Table 2: Electrophoretic parameters used to resolve restricted HMW DNA

Switch Time	Run Time	Gel	Voltage	Resolution Window
30 min	72 hrs	0.6%, 0.5 x TBE	50 V	2.2 Mb – 4.6 Mb
60-90 sec	22-24 hrs	1.0%, 0.5 x TBE	200 V	220 Kb – 2.2 Mb
15 – 50 sec	24 hrs	1.0%, 0.5 x TBE	200 V	24 Kb – 610 Kb
1-12 sec	14 hrs	1.0%, 0.5 x TBE	200 V	4.36 Kb – 145 Kb

3.2.6 BAC library screening

3.2.6.1 Screening of a barley and wheat bacterial artificial chromosome (BAC) library

Filters of a 6.5 fold haploid genome equivalent barley BAC library constructed from barley (*Hordeum vulgare*) cultivar Morex were imported from the Clemson University Genomics Institute (USA). Tim Sutton performed initial screening of BAC filters as described in Sections 3.2.6.1.1 → 3.2.6.1.5.

A 3.7 fold haploid genome equivalent wheat BAC library, from the diploid D-genome progenitor of wheat (*Triticum tauschii*) was kindly screened by Dr E. Lagudah (CSIRO Plant Industry, Canberra) (Moullet *et al.*, 1999).

3.2.6.1.1 Purification of DNA probe template

Wheat and barley BAC clones were isolated by hybridisation with a 3-Kb *WM1.1* ORF probe generated by the amplification of a genomic clone from wheat (*T. aestivum* cv. Chinese Spring) with primers *WM1.1* ORF (refer to **Table 1**). No vector sequence was present in the probe since this could result in cross-hybridisation with vector sequence of the BAC vector.

3.2.6.1.2 BAC filter preparation

Barley BAC filters were pre-hybridised in 30 mL of hybridisation buffer (160 mL of 1 M sodium phosphate pH 7.2 buffer, 112 mL of 20% SDS, 0.6 mL of 0.5 M EDTA, 330 μL of 10 mg mL^{-1} denatured salmon sperm and 27 mL MQH_2O) at 65°C overnight. The pre-hybridisation solution was then discarded and fresh hybridisation solution containing denatured probe was added and incubated for 24 hrs at 65 °C. Filters were then washed in the bottle twice with 2 × SSC, 0.1% SDS at RT. Radioactive counts were measured to check signal strength before subsequent washing down to 0.5 × SSC at 50 °C. Subsequent washes were undertaken until membrane signal strength was no greater than counts 15 counts/second but not less than 10 counts/second. After washing, filters were wrapped in plastic film and exposed to Konica-HR X-ray film with an intensifying

screen at –80 °C for a minimum of 10 days to ensure that background was sufficiently strong enough to help correctly address positive BAC signals. This overexposure was used to help with the orientation and identification of autoradiographs of barley BAC membranes. Barley BAC addresses (positions on the filter) were identified according to instructions at the Clemson University Genomics Institute homepage (http://www.genome.clemson.edu).

Wheat BACs identified, in the primary library screen, to hybridise strongly to *WM1.1* ORF were isolated by Dr E. Lagudah (CSIRO, Plant Industry, Canberra).

3.2.6.1.3 Growth of BAC clones

Clones received as stab cultures from CSIRO Plant Industry (Canberra) and Clemson University Genomics Institute were immediately grown in a liquid LB broth as described in Section 2.2.2. The antibiotic chloramphenicol was added to bacterial growth media at 25 μg mL^{-1}. Single colonies were picked and glycerol cultures made for each clone.

3.2.6.1.4 Mini-preparation of BAC plasmid DNA

A single colony was picked and grown in 1.5 mL of L-broth with 25 μg mL^{-1} chloramphenicol in a 10 mL cap tube overnight at 37°C on a rotating wheel. The culture was then poured into an eppendorf tube and the cells pelleted by centrifugation for 3 min at 14,000 rpm. The supernatant was discarded and the cell pellet resuspended in 90 μL of TES (25 mM Tris-HCl (pH 8.0), 10 mM EDTA, 15% w/s sucrose) buffer by vortexing. Lysis solution (180 μL)(0.2 M NaOH, 1% SDS) was added and gently mixed by inversion without vortexing. One volume of 135 μL 3 M NaOAc, pH 4.6 was added and gently mixed by inversion. The bacterial chromosome DNA was then pelleted by centrifugation at 14,000 rpm for 15 min. The supernatant was removed to a fresh tube. A 2 μL volume of 10 mg mL^{-1} RNase A was added at 37 °C for 1 hr. TE-saturated phenol (400 μL) and chloroform (400 μL) were added to samples, and then gently mixed by inversion. Samples were centrifuged at 14,000 rpm for 5 min and the supernatant extracted. BAC DNA was then precipitated using 2-3 volumes of -20 °C absolute ethanol for 10 min at RT. DNA was then pelleted at

14,000 rpm for 15 min before being washed with 500 μL 70% ethanol. DNA was resuspended in 20-50 μL of TE (10 mM Tris-HCl, 1.0 mM EDTA) buffer. DNA was stored at 4°C prior to subsequent analysis.

3.2.6.1.5 Confirmation of BAC clones

To confirm positive identification of barley and wheat BACs imported from Clemson University and CSIRO Plant Industry, Tim Sutton performed dot blot analysis on all BAC clones received. Samples of purified BAC DNA (10 μL) were incubated at 94 °C for 5 min before being put on ice. Samples (5 μL) were then spotted on Hybond N^+ membrane and left to air dry for 10 min. DNA was then cross-linked to the nylon membrane using 125 mJoules of UV light using a BIORAD GS gene linker. Membranes were washed three times in 2 × SSC before being ready for hybridisation. General ^{32}P labelled probe preparation, hybridisation, and autoradiography were as described in Sections 2.2.12 and 2.2.11.

Hybridising wheat and barley BAC clones were verified with the same 3-Kb *WM1.1* ORF probe used for the initial library screen (Section 3.2.6.1.1).

Individual members of the *WM1* gene family were identified by hybridisation with PCR amplified probes from linker regions of *WM1.1*, *WM1.2* and *WM1.3* (**Table 1**).

3.2.6.1.6 Southern fingerprint analysis of BAC clones

Purified BAC DNA (10 μg) was digested with *Hind III* to release the BAC vector and digest wheat and barley DNA inserts so BAC clones could be compared using Southern analysis for common restriction fragment bands of the same molecular weight that may infer an overlap between two or more different BAC clones. *Dra I* and *Eco RV* restriction endonuclease digests were also performed in order to determine overlapping BAC clones. Southern transfers were performed on gels as described in 2.2.10. Membranes were then hybridised with radiolabelled BAC vector DNA (Section 3.2.6.1.7). PCR amplified *WM1.1* ORF, *WM1.1* linker, *WM1.2* linker and *WM1.3* linker (refer to Section 3.2.3 and

Table 1) were also radiolabelled and used as probes in Southern analysis to identify the whole *WM1* gene family and it's individual members.

3.2.6.1.7 Preparation of ^{32}P-labelled total BAC DNA probes

Total BAC vector DNA (200 ng) nicked with 60 mJoules of UV was radiolabelled with ^{32}P (Feinberg & Vogelstein, 1983) according to Section 2.2.11.

3.2.6.1.8 Pulsed field gel electrophoresis

Wheat and barley BAC clones containing *WM1* genes were sized using PFGE as described in 3.2.5. Purified BAC (15 μg) was first digested with *Not I* to release the insert and linearise the DNA so it could be sized. All pulse field gels were run for 22 hrs at 200 volts with a linear ramp of 5-20 sec.

3.2.7 BAC fingerprinting

BAC DNA of clones was isolated by using alkaline lysis procedure (Sambrook *et al.*, 1989). BAC fingerprinting was performed according to procedures developed by DuPont (Morgante, *pers. commun.*) and is outlined below.

BAC DNA (500 ng) for clones *TtBAC3*, *TtBAC6*, *TtBAC11* was used as template for a type II restriction endonuclease digest according to **Table 3**. These individual type II restriction endonuclease (**Table 4**) digests were each performed in combination with a *Taq I* restriction endonuclease digest. The addition of *Taq FS* polymerase using big dye terminators ddATP, ddTTP and ddGTP and unlabelled ddCTP allowed the specific fluorescent labelling of the ambiguous restriction sites generated by the type II restriction enzymes. The materials for the restriction digests and dye incorporation are outlined in **Table 3** and **Table 4**. Unincorporated dyes were removed from the fingerprinting reaction through Centrisep clean-up spin columns (pre-warmed before use). The sample was rehydrated with loading dye (2.5 μl) and Rox500 size standard (0.5 μl) and subsequently electrophoretically separated on either an ABI 377 or 3700 (PE Applied Biosystems).

3.2.8 BAC contig assembly

WM1.1 ORF probed Southerns of *Hind III*, *Dra I*, and *Eco RV* digested barley and wheat BACs were analysed manually for common bands inferring overlap. This was compared to the *Not I* digested, PFGE data of the wheat BACs.

This data allowed an approximate physical size to be assigned to component fragments and overall BAC insert size.

The fingerprinting data from each of *TtBAC3*, *TtBAC6* and *TtBAC11* generated from the ABI 377 or 3700 (PE Applied Biosystems) was analysed for commonality in banding patterns. The analysis performed on the computing software at DuPont (Morgante, *pers. commun.*) resulted in a probable contig assignment of the three wheat BAC clones. The contig generated manually through Southern analysis was confirmed by the BAC fingerprint analysis.

3.2.9 BAC DNA sequencing

Overlapping clones consisting of *TtBAC3*, *TtBAC6* and *TtBAC11* were chosen for complete sequencing (refer to **Figure 14**). Each of these BAC clones was sequenced using a shotgun approach (Qiagen, 1998). BAC subcloning and removal of subcloned vector were performed by Luke Gumaelius (DuPont, USA) with details outlined in 3.2.9.1 and 3.2.9.2

3.2.9.1 BAC subcloning

BAC DNA was purified according to manufacturers specifications for the Qiagen Large-Construct Kit and was sheared by nebulisation (Qiagen, 1998). End repair was performed by using *Pfu* DNA polymerase (Stratagene, La Jolla, CA) treatment according to the manufacturer's directions. DNA fragments were then size fractionated by gel electrophoresis and cloned into the *Sma I* site of *pUC18* (Amershan Pharmacia Biotech, Piscataway, NJ). Three separate subcloned libraries were constructed for each individual BAC clone.

Table 3: BAC fingerprinting reaction components and conditions

Reactant	Volume
BAC DNA template (500 ng)	18 μL
ddATP (big dye terminator)	0.5 μL
ddTTP (big dye terminator)	0.5 μL
ddGTP (big dye terminator)	0.5 μL
ddCTP	0.5 μL
Taq FS (8U/μL)	0.67 μL
Type II restriction enzyme (eg. *Ear1* – 8U/μL)	1.0 μL
Taq I restriction enzyme	0.5 μL
10xBuffer	2.5 μL
ddH_2O	0.33 μL
Final volume	25.0 μL
Incubation at optimal temperature for Type II restriction enzyme (refer to Table 4) for 1 hr. For an additional 1 hr the reaction is heated to 72°C for the *Taq I* restriction and for *Taq FS* polymerase activity.	

Table 4: Type II restriction enzymes and corresponding incubation temperatures used for BAC fingerprinting

Enzyme	Recognition Site	Optimal Incubation Temperature
Bsp MI	5'…ACCTGC$(N)_4$∇…3' 3'…TGGACG$(N)_8$∇…5'	37°C
Bbs I	5'…GAAGAC$(N)_2$∇…3' 3'…CTTCTG$(N)_6$∇…5'	37°C
Bsa I	5'…GGTCTC$(N)_1$∇…3' 3'…CCAGAG$(N)_5$∇…5'	50°C
Bsb MI	5'…CGTCTC$(N)_1$∇…3' 3'…GCAGAG$(N)_5$∇…5'	55°C
Eco NI	5'…CCTNN∇NNNAGG…3' 3'…GGANNN∇NNTCC…5'	37°C
Bsm FI	5'…GGGAC$(N)_{10}$∇…3' 3'…CCCTG$(N)_{14}$∇…5'	65°C
Fnu4 HI	5'…GC∇NGC…3' 3'…CGN∇CG…5'	37°C
Ear I	5'…CTCTTC$(N)_1$∇…3' 3'…GAGAAG$(N)_4$∇…5'	37°C

Two 1-2 Kb and one 5-6 Kb subclone library was constructed for each BAC clone through transformation into electrocompetent cells of *E. coli* DH10B (Life Technologies, Rockville, MD). For each BAC clone 786 random recombinant clones were picked from 1-2 Kb subclone library and 384 from the 5-6 Kb subclone library using a Q-bot (http://www.genetix.co.uk/qbot.html).

3.2.9.2 Removal of subcloned BAC vector

Binary BAC vector *pCLD0454* subclones were eliminated from DNA sequencing through vector hybridisation to clones arrayed on Hybond N+ membranes (Amersham, UK). Duplicates of each of the 3456 bacterial clones were arrayed on Hybond N+ membranes utilising a 384 prong high density-replicating tool (http://www.genetix.co.uk/qbot.html). Bacterial colonies were grown on membranes at 37°C until the colonies were 1 to 2 mm in diameter. Membranes were then removed and placed colony side up on a single pad of absorbent Whatmann filter paper soaked firstly in 0.5 N NaOH, 1.5 M NaCl for 4 min. Secondly the membrane was placed on soaked Whatmann filter paper containing 1.5 M NaCl, 1.0 M Tris-base, pH 7.4 for 4 min. Finally the membrane was immersed with shaking at 55°C for 1 hr in 1% sarkosyl, 100 mM NaCl, 50 mM Tris-HCl (pH7.4), 50 mM EDTA and 250 μg mL^{-1} proteinase K. Membranes were then UV crosslinked and dried at RT. *pCLD0454* vector was digoxigenin-labeled (Boehringer Mannheim, Germany) and mixed with unlabelled *pUC18* then hybridised to the Hybond N+ membranes according to manufacturers specifications. Autoradiography was performed for 10 –15 min as per Section 2.2.12.

The autoradiographs were manually analysed for strong hybridisation signals of the *pCLD0454* vector to individual colonies. Those clones that showed a strong signal were highlighted and subsequently removed from sequencing by the re-arraying of the 384-well microtitre plates with only those clones showing no or weak hybridisation (http://www.genetix.co.uk/qbot.html).

3.2.9.3 DNA sequencing

DNA templates for sequencing were prepared using a 96-well alkaline lysis miniprep kit (Advanced Genetic Technologies Corp, Gaithersberg, MD). Sequencing reactions were performed by using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction kit with *FS AmpliTaq* DNA Polymerase (PE Applied Biosystems, Foster City, CA) and analysed on both ABI 377 sequencing gels and by capillary electrophoresis on an ABI 3700 (PE Applied Biosystems).

3.2.9.4 DNA sequence assembly

The sequence data was assembled into contigs by PHRED/PHRAP, CROSSMATCH software in combination with CONSED (Gordon *et al.*, 1998). Utilising CROSSMATCH software, vector sequence of the subcloning (*pUC18*) and BAC vector (cosmid binary vector *pCLD0454*) was removed from the alignment. Sequence alignment was cross-checked via CONSED for misalignment of repetitive sequences (Wolters, *pers. commun.*).

3.2.9.5 Sequence analysis

3.2.9.5.1 BLAST on non-redundant and EST databases

Homology searches against public (ITEC) and private (DuPont Ag Biotechnology) databases were used to identify candidate genes in the region. The final sequences of the region were divided into 3-Kb overlapping fragments and searched for nucleic or protein homologies by using the BLAST 2.0 blastn program (Altschul *et al.*, 1997).

3.2.9.5.2 Gene prediction

The assembled BAC sequence data was analysed using RiceGAAS software (http://rgp.dna.affrc.go.jp/) for identification of possible coding sequences. RiceGAAS software bases gene prediction on *Arabidopsis thaliana* and *Zea mays* GENSCAN (Burge & Karlin, 1997; Burset & Guigo, 1996), and Rice HMM (Sakata *et al.*, 1999) incorporating results from BLAST 2.0 blastn and blastp (Altschul *et al.*, 1997) searches against the rice EST database (Rice Genome Project, Japan) and the non-redundant database (NCBI, USA). Rice GAAS

predicted genes were annotated through the use of Vector NTI Suite Version 6.0 software (Bethesda, MD, USA).

3.2.9.5.2.1 Signal prediction

TATA-box and poly-adenylation signals were predicted with HcTATA and HcPOLYA respectively (Milanesi *et al.*, 1999; Milanesi *et al.*, 1996). The TATA-box prediction was performed on the consensus sequence generated from an AlignX (Vector NTI Suite Version 6.0 software) alignment of 1 Kb of sequence 5' to the predicted translation start site for each of the predicted *WM1* genes. Likewise the poly-adenylation signal prediction was performed on the consensus sequence generated from the AlignX alignment of 500 bp of sequence 3' to the predicted translation stop site for each of the predicted *WM1* genes.

3.2.9.5.2.2 Promoter prediction

One kilobase 5' from the precited tanslation start site for each of the predicted *WM1* genes was analysed for promoter-like sequences. Promoters were predicted using the Neural Network Promoter Prediction program (http://www.fruitfly.org/seq_tools/promoter.html).

3.2.9.5.2.3 cis-acting regulatory element binding site prediction

One kilobase of sequence 5' to the predicted translation start site for each of the predicted *WM1* genes, was aligned using AlignX program in Vector NTI. The consensus sequence was analysed for plant *cis*-acting regulatory elements using the PLACE program (Higo *et al.*, 1999; Prestridge, 1991). Transcription factor binding sites were predicted by the TfSitescan program (http://www.ifti.org/cgi-bin/ifti/Tfsitescan.pl). Both *cis*-acting regulatory elements and transcription factor binding sites were annotated through the use of Vector NTI.

3.2.9.5.3 Identification of repetitive elements and sequence charactersistics

For all 270 Kb of genomic sequence, structural (CpG islands) and repetitive elements such as retrotransposons, and microsatellites were determined by both REPEATMASKER (http://ftp.genome.washington.edu/.html) and GrailEXP

(Hyatt *et al.*, 2000a; Hyatt *et al.*, 2000b). GrailEXP- and REPEATMASKER-identified repeat sequences were annotated through the use of Vector NTI.

3.2.9.5.4 Matrix attachment region prediction

Matrix attachment regions (MARs) or scaffold attachment regions (SARs) were predicted by the program MAR-FINDER (http://www.futuresoft.org/). MARs were annotated in Vector NTI.

3.2.10 Structural analysis of the WM1 predicted peptides

All predicted structural features of the WM1 primary peptides were annotated on the sequence in Vector NTI. Peptide alignments and phylogenetic analysis were performed by the use of AlignX within Vector NTI and Megalign program (DNASTAR, 1997).

3.2.10.1 Transmembrane domains

For each of the predicted WM1 primary peptides the TMpred program (Hofmann & Stoffel, 1993) was used to make predictions of membrane-spanning regions and their orientation for each of the predicted *WM1* primary peptides.

3.2.10.2 Leucine zipper and leucine rich repeats

Every predicted WM1 primary peptide was screened against both the PROSITE peptide motif database (Hofmann *et al.*, 1999) and INTERPRO V3.2 (Apweiler *et al.*, 2001) to identify peptide structural motifs. Leucine zipper motifs from each of the predicted WM1 primary peptides were compared using AlignX program in Vector NTI. The leucine zipper consensus of the alignment was then analysed using HelicalWheel (GCG program, Accelrys Inc.).

3.2.10.3 Signal peptides

SignalP V2.0 (Nielsen *et al.*, 1997) was used to predict and distinguish between signal peptides and non-signal peptides as well as the recognition of cleavage sites and the prediction of their cleavability of all predicted WM1 primary peptides. Signal peptides and their cleavage were also confirmed with PSORT (Nakai & Kanehisa, 1992; von Heijne, 1986).

3.2.10.4 Predicted subcellular localisation

Every predicted WM1 primary peptide sequence was analysed with the PSORT program (Nakai & Kanehisa, 1992). PSORT was used to predict protein localisation sites in cells. Specific motifs including the tyrosine YXXϕ endosomal/lysosomal sorting signal sequences, were determined by manual analysis of the sequence (Letourneur & Klausner, 1992; Marks *et al.*, 1996).

3.2.11 Genes homologous to *WM1* genes

Each predicted WM1 primary peptide sequence was screened using the BLAST 2.0 blastp program (Altschul *et al.*, 1997) across the non-redundant Genbank database. Peptide sequences were aligned in AlignX, and annotated within the Vector NTI.

3.2.12 Northerns of *WM1* genes

Genomic DNA amplified *WM1.1* ORF (**Table 1**) was purified and probed onto a Northern blot containing total RNA isolated from various wheat tissue as per Sections 2.2.13 and 2.2.14. The same *WM1.1* ORF PCR fragment was used to probe a Northern blot of 10 μg of polyA isolated mRNA from a young wheat spike with meiocytes at metaphase I. The mRNA was isolated from wheat meiotic spike total RNA using the Message Maker kit (Gibco-BRL) according to manufacturers specifications.

3.2.13 RT-PCR of *WM1* genes

First strand cDNA was synthesized from 5 μg of total RNA extracted from; root tissue, young leaf tissue (3 weeks old), and spikes at early meiosis (Chinese Spring, *ph2a* and *ph2b*). Thermoscript reverse transcriptase (Gibco-BRL, Australia) and *WM1F*amRT, *WM1*ZFRT, TOLP, WtMSHLo1 as the initiation primers (**Table 5**) were used for first strand cDNA synthesis. The composition of each reaction was as follows 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM DTT, 30 ng for each reverse transcrition primer (**Table 5**), 500 μM of each dNTP and 200 units Superscript II reverse transcriptase. Reverse transcription was carried out at 45°C for 1 hr and terminated by heating to 75°C for 15 min. RNA was degraded by the addition of 2 units of RNase H and

RNase T1 (Gibco-BRL) and incubation at 37°C for 10 min. cDNA was recovered by ethanol precipitation and the concentration determined by spectrometric measurement. Approximately 0.1 μg cDNA was used per PCR reaction.

3.2.13.1 Gene specific primers

Specific *WM1* gene family RT-PCR primers were designed from *TtBAC* DNA sequence using Vector NTI. Primers were targeted to the non-conserved regions (refer to **Figure 16**) of the predicted coding sequences for each member of the *WM1* gene family. **Table 5** describes RT-PCR primer sequences for the *WM1* genes and for three control sets; *Thioredoxin H*, *MSH7* (positive controls) and Glutenin promoter (genomic contamination control). The specific PCR was performed with the primer pairs as outlined in **Table 5** using a Programmable Thermal Controller (MRJ, USA).

Amplifications were performed under the following conditions: 1 min 94°C, 1 min annealing temperature[2], 1 min 72°C for 36 cycles and final extension of 10 min at 72°C.

RT-PCR products were electrophoretically separated on 2% agarose gel and blotted onto Hybond-N+ (Amersham). Southern hybridisation was performed using *WM1.1* ORF PCR fragment as a probe (refer to Sections 2.2.10 to 2.2.12).

3.3 Results

3.3.1 Nullisomic tetrasomic determined chromosome localisation

Southern analysis of *WM1.1* ORF probe on CS and nulli-tetra lines (Section 3.2.4) shows a high level of cross-hybridisation between different members of the *WM1* gene family given that probe hybridisation was at 65°C and the washing stringency was 0.2 x SSC. Both the *Eco RV* and *Dra I* Southerns (**Figure 6**) suggest that there are 14 members in this gene family. Eleven of these 14 potential genes were located to chromosome 3D.

[2] Optimum annealing temperature for each primer pair was determined using VectorNTI Suite (V6.0) software (Bethesda, MD, USA).

Table 6: RiceGAAS software identified coding sequences from 270 Kb of *Triticum tauschii* genomic sequence.

Predicted Proteins	Number of encoding genes
WM1 family proteins	7
Putative ripening-related proteins from grapevine	3
Bowman-Birk type trypsin inhibitor proteins	4
Transposable element component proteins with homology to hypothetical proteins from *Arabidopsis thaliana*, *Oryza sativa*, *Triticum aestivum*.	17
Predicted proteins	25
Total	**56**

Figure 6: Chromosomal assignment of the *WM1* gene family. Southern hybridisation *of WM1.1 ORF* to wheat genomic DNA restriction endonuclease digested with A- *Eco RV* and B- *Dra I*. All lines prefixed with NT are nullisomic-tetrasomics (ie N1A-T1D, nullisomic chromosome 1A, tetrasomic 1D) whilst CS represents Chinese Spring euploid wheat.

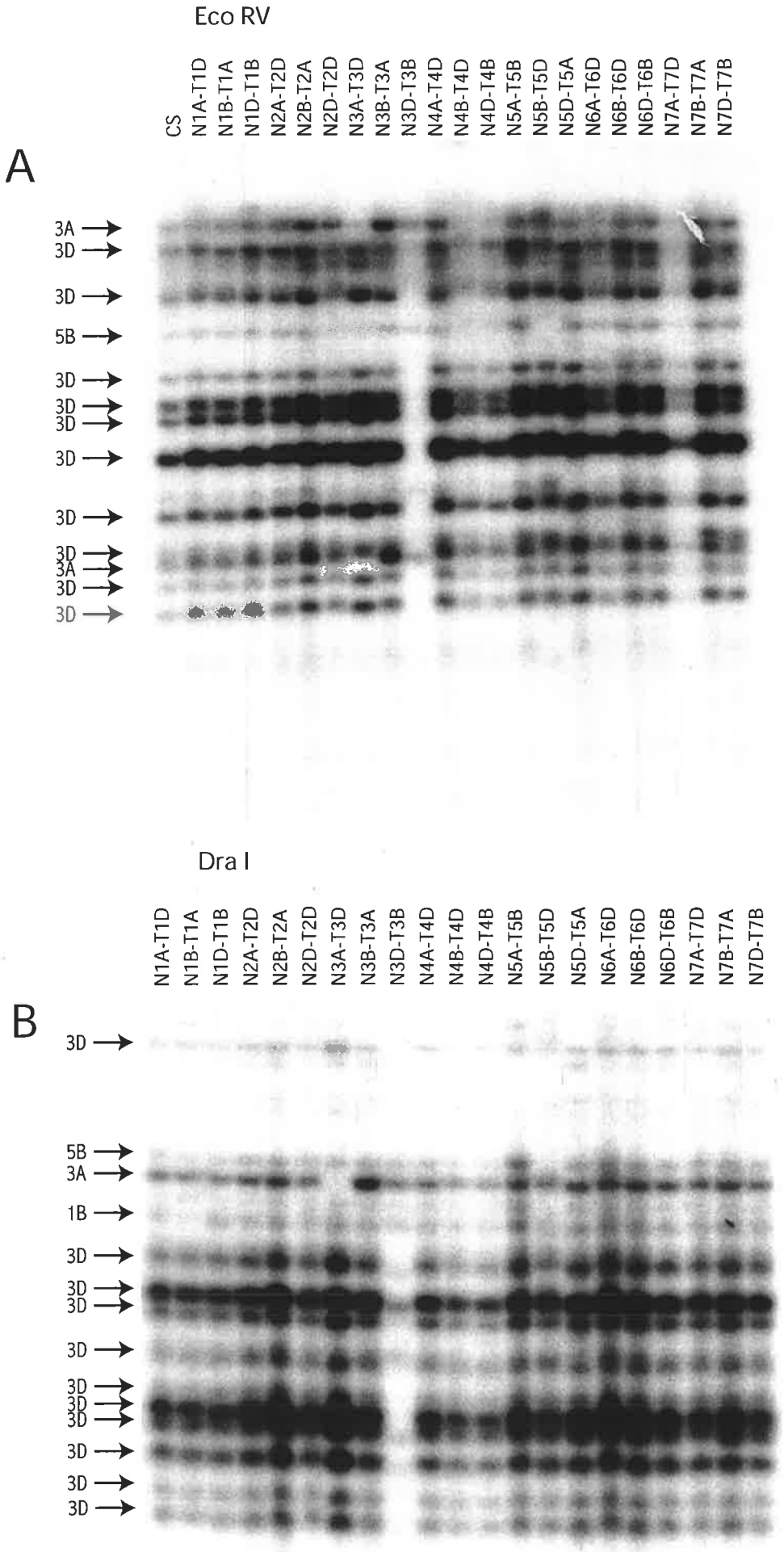

Eco RV
CS
N1A-T1D
N1B-T1A
N1D-T1B
N2A-T2D
N2B-T2A
N2D-T2D
N3A-T3D
N3B-T3A
N3D-T3B
N4A-T4D
N4B-T4D
N4D-T4B
N5A-T5B
N5B-T5D
N5D-T5A
N6A-T6D
N6B-T6D
N6D-T6B
N7A-T7D
N7B-T7A
N7D-T7B
A
3A
3D
3D
5B
3D
3D
3D
3D
3D
3D
3A
3D
3D
Dra I
N1A-T1D
N1B-T1A
N1D-T1B
N2A-T2D
N2B-T2A
N2D-T2D
N3A-T3D
N3B-T3A
N3D-T3B
N4A-T4D
N4B-T4D
N4D-T4B
N5A-T5B
N5B-T5D
N5D-T5A
N6A-T6D
N6B-T6D
N6D-T6B
N7A-T7D
N7B-T7A
N7D-T7B
B
3D
5B
3A
1B
3D
3D
3D
3D
3D
3D
3D
3D
3D

Figure 6 also showed that a further 3 potential genes were scattered across the genome with one being located on chromosome 3A, one on chromosome 1B and the last on chromosome 5B.

Given that the majority of the gene family was located on chromosome 3D, Southern analysis on chromosome 3 selected nullisomic tetrasomic, ditelosomic and meiotic mutants was performed using the *WM1.1* ORF probe. Results (**Figure 8** – A and **Figure 8** - B) confirmed that 11 of the 14 potential genes in the *WM1* gene family were located on chromosome 3D. Furthermore ditelosomic lines allow localisation to the short arm of chromosome 3D and with the meiotic mutants also showing that all 11 potential *WM1* genes on chromosome 3DS were located within the region defined by the *ph2a* deletion mutant. However no restriction fragment length polymorphisms were observable between the Chinese Spring and the *ph2b* mutant. Southern analysis of specific members of the *WM1* gene family for which genomic sequence was available (Ji, 1992), showed that *WM1.1*, *WM1.2* and *WM1.3* were all located within the region defined by the *ph2a* deletion mutant (**Figure 8**).

3.3.2 Pulse field gel electrophoresis

In order to determine if the 11 potential *WM1* genes were clustered within the region defined by the deletion in the *ph2a* meiotic mutant, restriction analysis on high molecular weight (HMW) DNA was performed.

Restriction endonucleases (*Sma I*, *Not I* and *Xho I*) known to cut at low frequency were used to yield large fragments of DNA (between 40 and 600 Kb). Southern analysis of Chinese Spring HMW DNA restricted with *Sma I*, *Not I* and *Xho I* and probed with *WM1.1* linker, *WM1.2* linker and *WM1.3* linker probes (refer to Section 3.2.3) showed that *WM1.1* and *WM1.3* members of the *WM1* gene family were not co-localised on restriction fragments greater than 100 Kb (**Figure 9**). Because of the reduced resolution in size of bands less than 100 Kb (*Xho I* digest) co-localisation of *WM1.1*, *WM1.2* and *WM1.3* genes could not be deciphered (**Figure 9**).

Figure 7: Sub-chromosomal arm assignment and localisation of *WM1* gene family members within the region deleted in the *ph2a* mutant. Southern hybridisation *of WM1.1 ORF* to wheat genomic DNA restriction endonuclease digested with A- *Eco RV*, B- *Dra I* and C- *Eco RI*, *Bam HI*, *Hind III*, *Dra I* and *Eco RV*. All lines prefixed with NT are nullisomic-tetrasomics (ie N1A-T1D, nullisomic chromosome 1A, tetrasomic 1D), DT are ditelosomic (ie DT3AS, ditelosomic short arm of chromosome 3A), whilst CS represents Chinese Spring euploid wheat and *ph2a* and *ph2b* are the deletion and point mutant lines for the *Ph2* locus in euploid wheat.

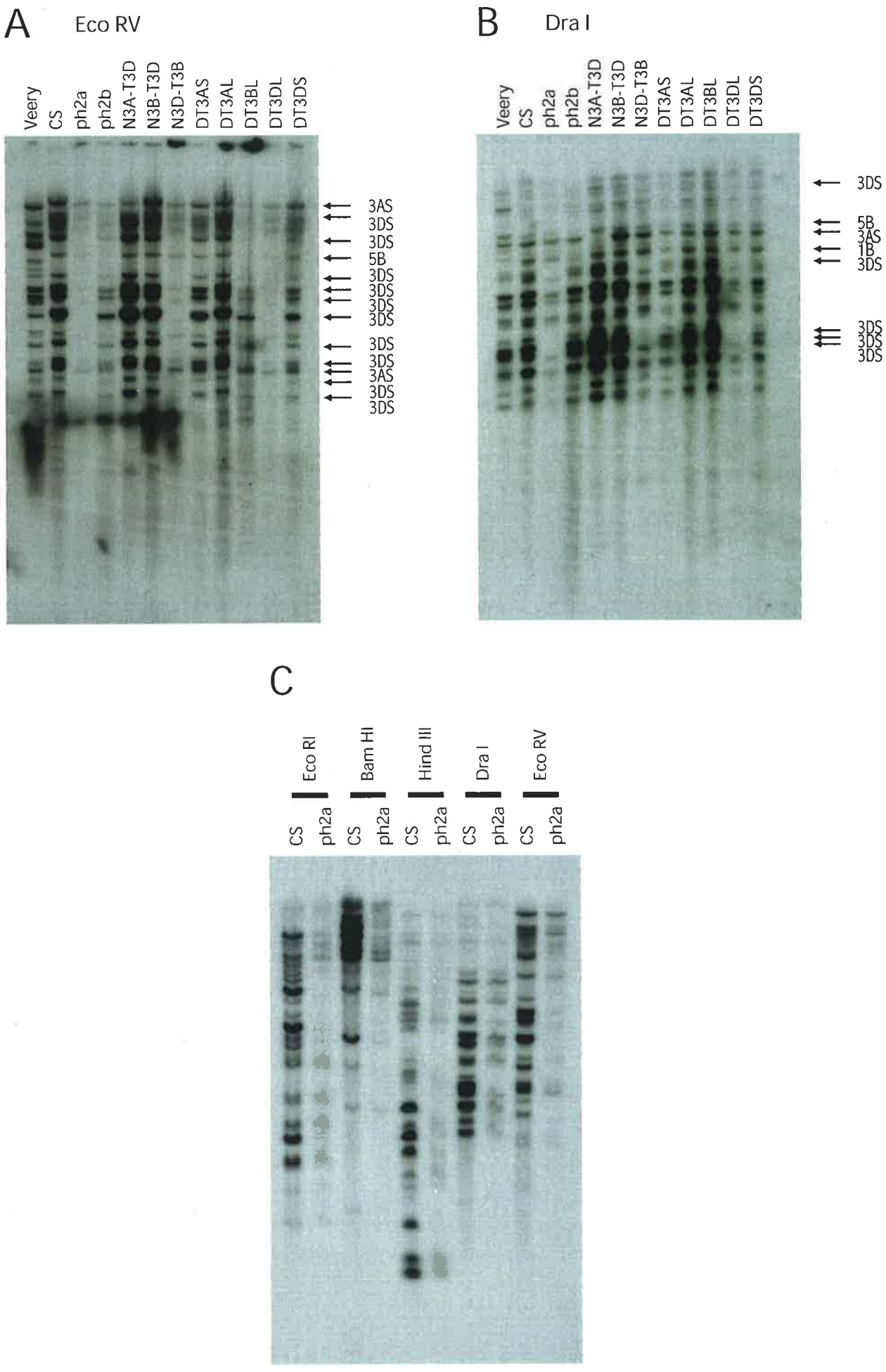

A
Eco RV
Veery
CS
ph2a
ph2b
N3A-T3D
N3B-T3D
N3D-T3B
DT3AS
DT3AL
DT3BL
DT3DL
DT3DS
3AS
3DS
3DS
5B
3DS
3DS
3DS
3DS
3DS
3DS
3AS
3DS
3DS
B
Dra I
Veery
CS
ph2a
ph2b
N3A-T3D
N3B-T3D
N3D-T3B
DT3AS
DT3AL
DT3BL
DT3DL
DT3DS
3DS
5B
3AS
1B
3DS
3DS
3DS
3DS
C
Eco RI
Bam HI
Hind III
Dra I
Eco RV
CS
ph2a
CS
ph2a
CS
ph2a
CS
ph2a
CS
ph2a

Figure 8: Sub-chromosomal arm assignment of *WM1* gene family members *WM1.1*, *WM1.2* and *WM1.3*. Southern hybridisation *of WM1.1 linker*, *WM1.2 linker* and *WM1.3 linker* to wheat genomic DNA restriction endonuclease digested with A- *Eco RV* and B-*Dra I*. All lines prefixed with NT are nullisomic-tetrasomics (ie N1A-T1D, nullisomic chromosome 1A, tetrasomic 1D), DT are ditelosomic (ie DT3AS, ditelosomic short arm of chromosome 3A), whilst CS represents Chinese Spring euploid wheat and *ph2a* and *ph2b* are the deletion and point mutant lines for the *Ph2* locus in euploid wheat.

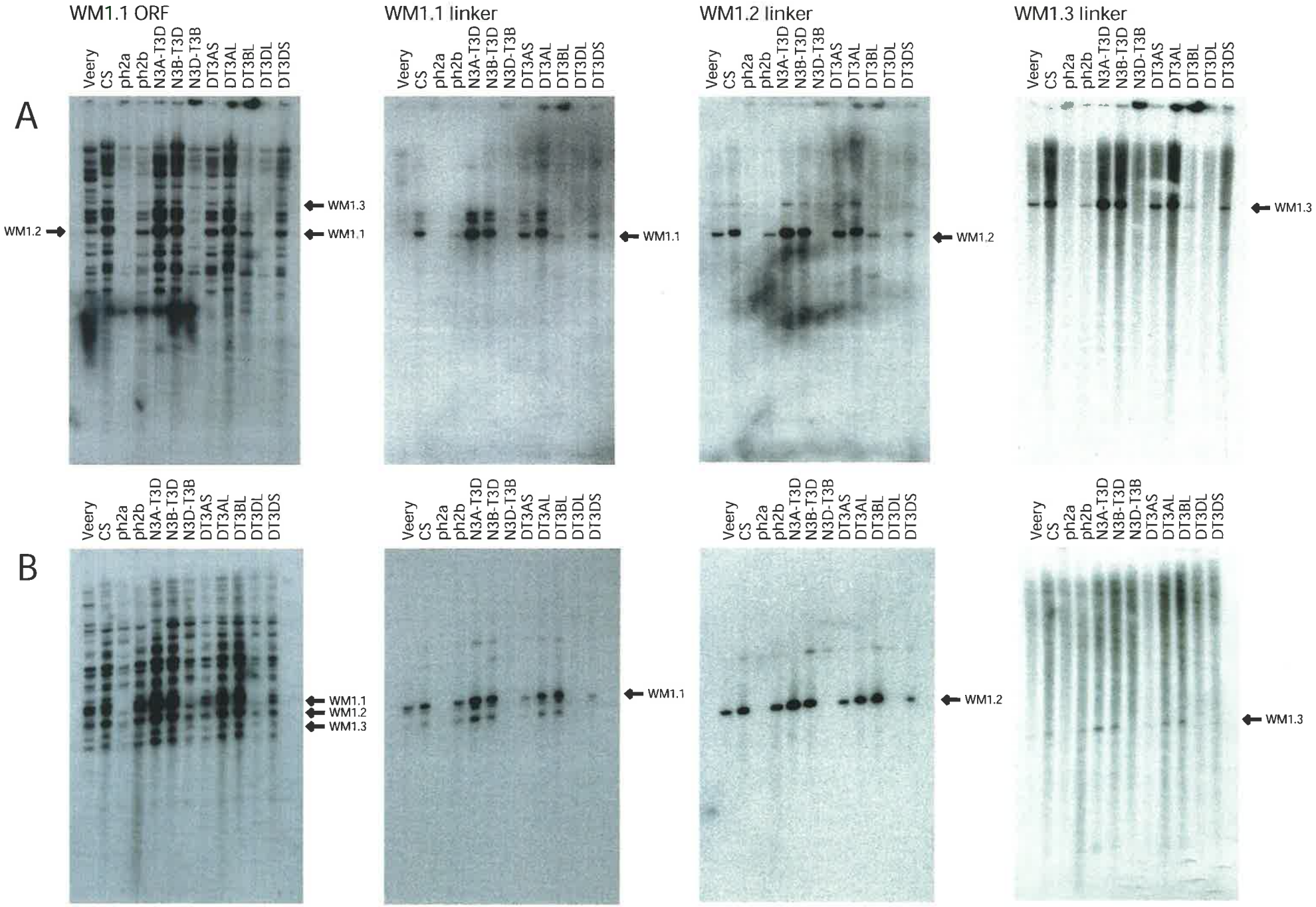
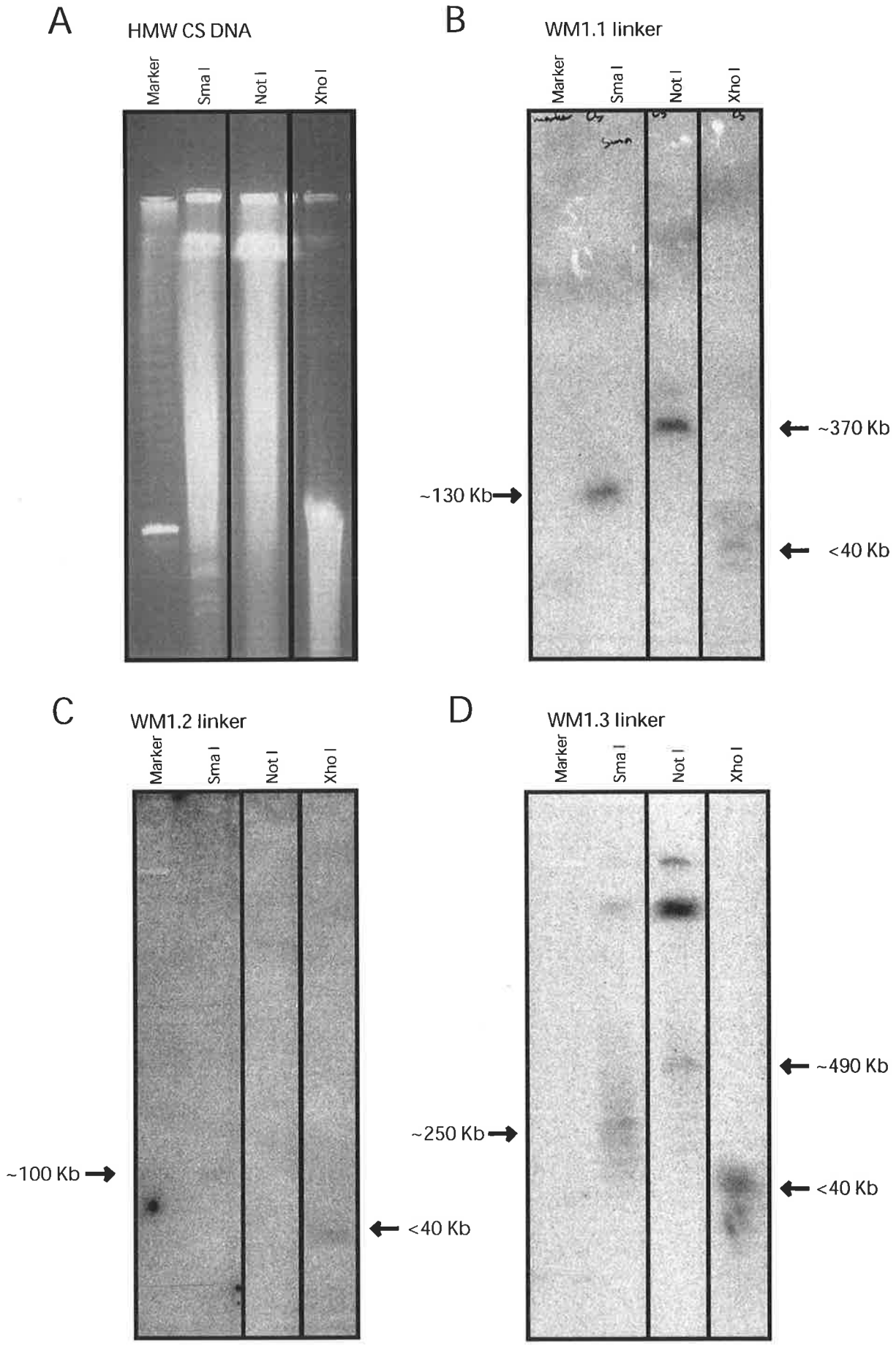

WM1.1 ORF
WM1.1 linker
WM1.2 linker
WM1.3 linker
Veery
CS
ph2a
ph2b
N3A-T3D
N3B-T3D
N3D-T3B
DT3AS
DT3AL
DT3BL
DT3DL
DT3DS
A
B
WM1.1
WM1.2
WM1.3

Figure 9: Physical linkage between *WM1* gene family members *WM1.1*, *WM1.2* and *WM1.3*. Southern hybridisation *of WM1.1 linker*, *WM1.2 linker* and *WM1.3 linker* to wheat high molecular weight (HMW) genomic DNA restriction endonuclease digested with *Sma I*, *Not I* and *Xho I*. All genomic DNA is derived from Chinese Spring euploid wheat. Each hybridising band size is estimated based on long-range λ concatemer DNA ladder (Marker – New England Biolabs).


A
HMW CS DNA
Marker
Sma I
Not I
Xho I
B
WM1.1 linker
Marker
Sma I
Not I
Xho I
~370 Kb
~130 Kb
<40 Kb
C
WM1.2 linker
Marker
Sma I
Not I
Xho I
~100 Kb
<40 Kb
D
WM1.3 linker
Marker
Sma I
Not I
Xho I
~490 Kb
~250 Kb
<40 Kb

3.3.3 BAC library screening

Fourteen individual barley BAC clones were isolated from the preliminary *H. vulgare* library screen whilst 20 individual wheat BAC clones were isolated from the preliminary *T. tauschii* library screen.

The barley clones have been described as *Hordeum vulgare* BAC clone numbers 1 to 14 (*HvBAC*1→14)

The wheat clones have been described as Triticum tauschii BAC clone numbers 1 to 20 (TtBAC1→20).

3.3.3.1 Southern fingerprint analysis of BAC clones

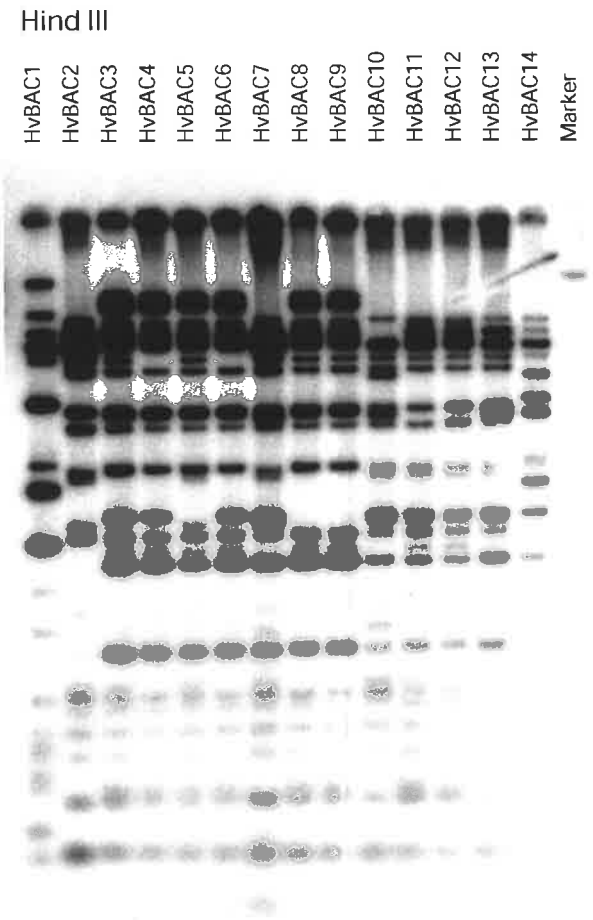
Hind III, *Dra I*, and *Eco RV* restriction endonuclease digests and subsequent Southern analysis was performed in order to determine overlapping barley and wheat BAC clones. Both barley and wheat membranes that were hybridised with radiolabelled total BAC DNA (barley - **Figure 10** - A, wheat – **Figure 10** - B) showed numerous common bands suggesting overlap between barley clones and overlap between wheat clones.

Further Southern analysis of the wheat BAC clones with *WM1.1* ORF (**Figure 11**), *WM1.1* linker (**Figure 11**), *WM1.2* linker (**Figure 11**) and *WM1.3* linker (**Figure 11**) (refer to Section 3.2.3 and **Table 1**) probes showed the *WM1* gene family cluster in wheat. It is anticipated that there is approximately 11 genes located within the region. Only an estimate can be made as it is unknown whether or not multiple *Hind III*, *Dra I* or *Eco RV* restriction sites within the coding sequences for individual *WM1* genes are causing duplication of bands. Such an analysis in barley could not be performed due to the sequence specificity for wheat for each of the linker probes from *WM1.1*, *WM1.2* and *WM1.3* (data not shown). However, probing barley BAC clones that had been digested with *Hind III*, with *WM1.1* ORF suggests that a similar gene cluster is present in barley (**Figure 12**).

Figure 10: Determination of overlapping barley and wheat BAC clones through commonality in banding patterns. Commonality in banding patterns as determined by Southern hybridisation of *total BAC DNA* to restriction endonuclease digested barley (A – *Hind III*) and wheat (B – *Eco RV*) BAC DNA. For both Panels A and B each separate lane corresponds to an independent BAC clone labelled either *HvBAC#* (*Hordeum vulgare* BAC clone number #) or *TtBAC#* (*Triticum tauschii* BAC clone number #). Lanes labelled marker contain *Bst EII/Sal I* restriction endonuclease digested λ DNA.

A

Hind III



B

Eco RV

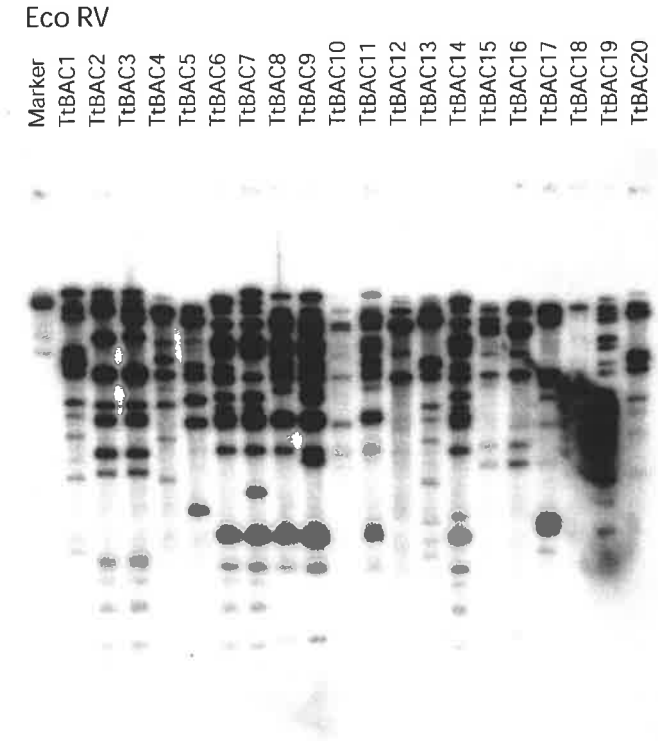


Figure 11: Co-localisation of *WM1* gene family members on overlapping wheat BAC clones. Physical co-localisation of *WM1* genes as determined by Southern hybridisation of *WM1.1 linker*, *WM1.2 linker* and *WM1.3 linker* to restriction endonuclease digested wheat (A – *EcoRV*, B – *Dra I* and C - *Hind III*) BAC DNA. For panels A, B and C each separate lane corresponds to an independent BAC clone labelled *TtBAC#* (*Triticum tauschii* BAC clone number #). Lanes labelled Marker contain *Bst EII/Sal I* restriction endonuclease digested λ DNA.

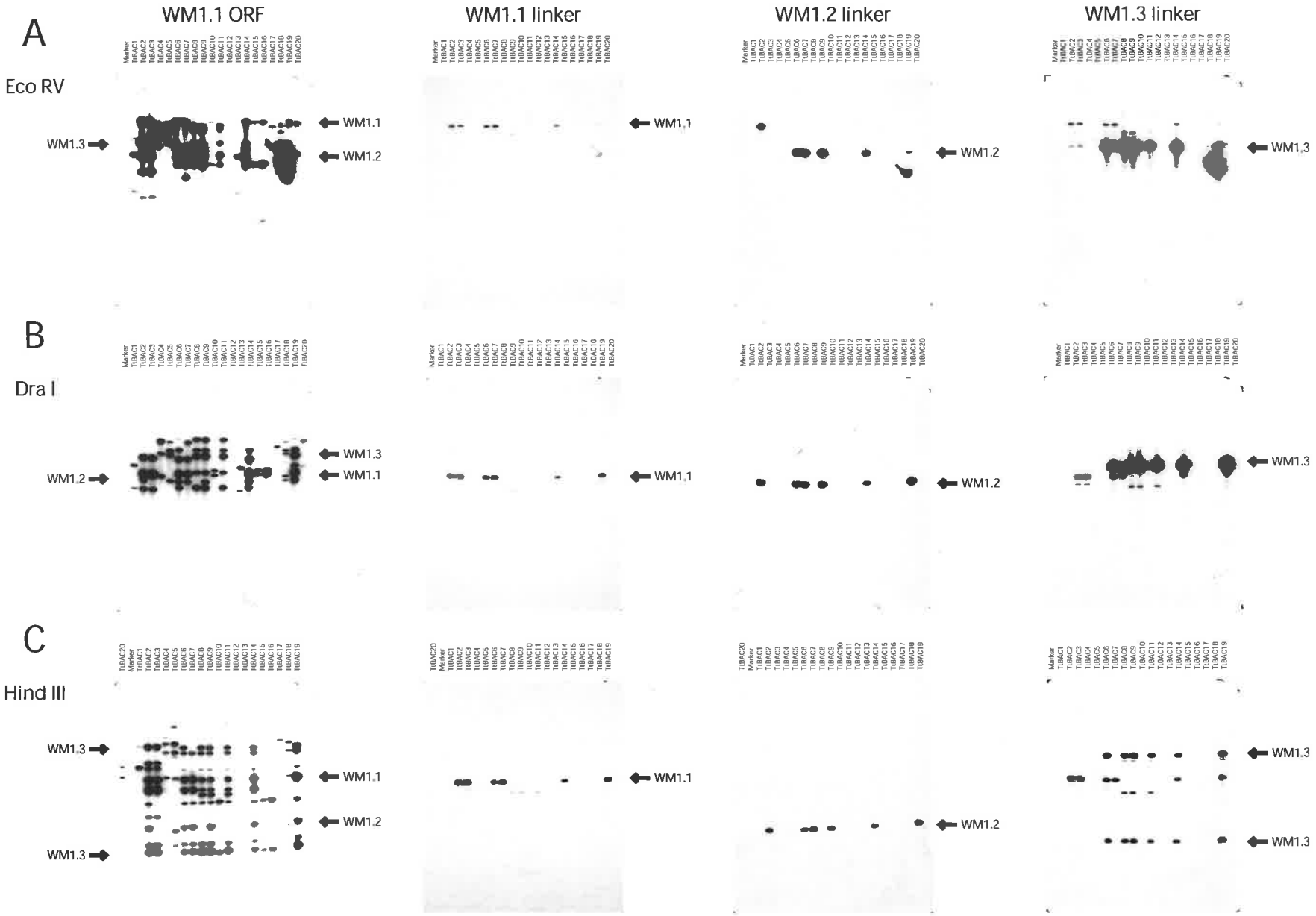
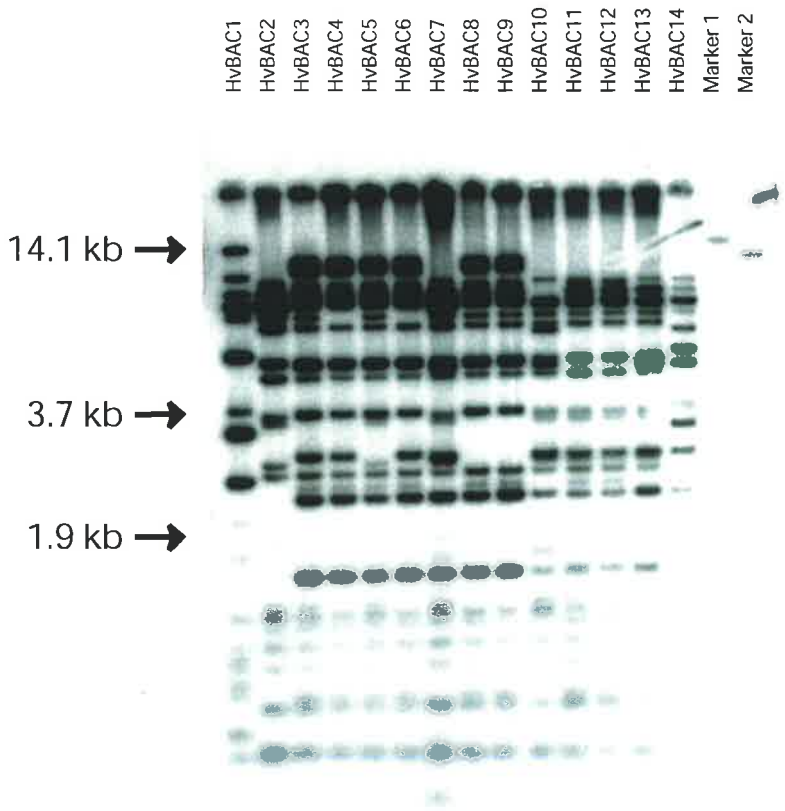

A
WM1.1 ORF
WM1.1 linker
WM1.2 linker
WM1.3 linker
Eco RV
WM1.3
WM1.1
WM1.2
B
Dra I
C
Hind III

Figure 12: Co-localisation of homologues to the *WM1* gene family members on overlapping barley BAC clones. Physical co-localisation of barley homologues to the *WM1* gene family as determined by Southern analysis using A- total BAC DNA and B - *WM1.1* ORFprobes on *Hind III* restriction endonuclease digested barley BAC DNA. For Panels A and B each separate lane corresponds to an independent BAC clone labelled *HvBAC#* (*Hordeum vulgare* BAC clone number #). Lanes labelled Marker 1 and Marker 2 contain *Hind III* and *Bst EII/Sal I* restriction endonuclease digested λ DNA respectively.

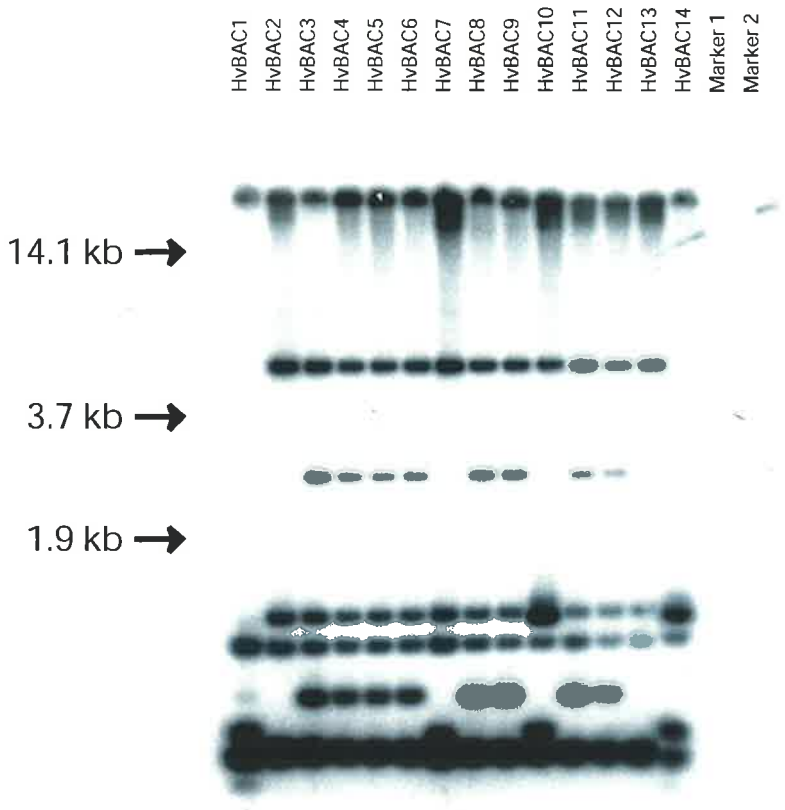
A

Hind III



B

Hind III



Southern analysis on barley BACs suggests that there are 4 or 5 genes comprising the *WM1* gene cluster in barley as opposed to the anticipated 11 genes in *T. tauschii*. Observations on commonality in Southern banding patterns for each of the wheat BAC and each of the barley BAC clones suggests that the insert DNA sequence was derived from two different regions of the genome in wheat whilst in barley it was from one region.

3.3.3.2 *Pulsed field gel electrophoresis*

Figure 13 illustrates the PFGE separated BAC DNA that has been probed with *WM1.1* ORF and illustrates that the insert sizes for the BAC clones range from 36 Kb to 92 Kb. Given that *Not I* cuts DNA in GC rich sequences, the multiple banding patterns in **Figure 13** show the region harbouring the cluster of *WM1* genes is likely to contain GC rich DNA sequences.

3.3.4 BAC fingerprinting

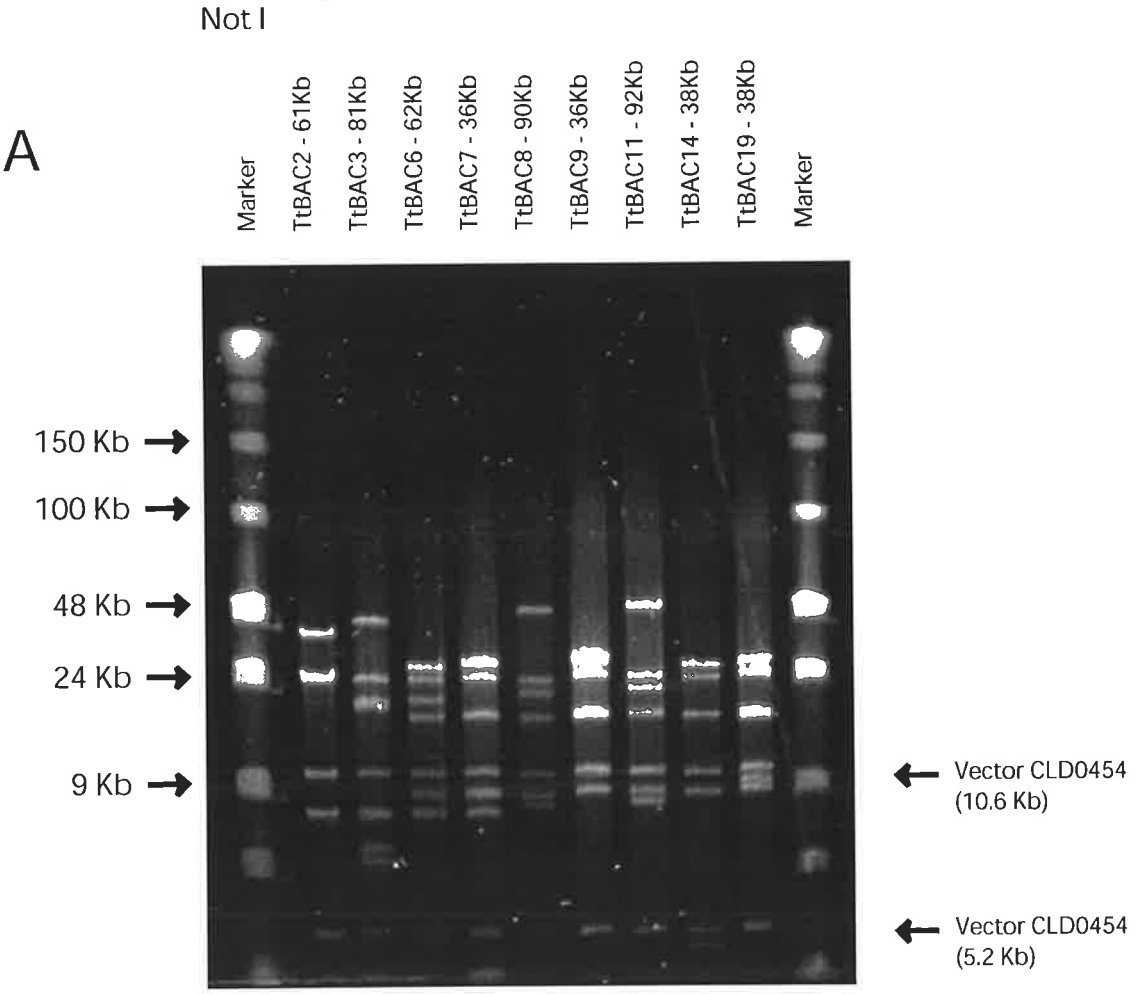
The fingerprinting data from each of *TtBAC3*, *TtBAC6* and *TtBAC11* generated from the ABI 377 or 3700 was analysed using the computing software at DuPont (Morgante, *pers. commun.*). The raw fingerprinting data for each of the eight type II restriction endonuclease digests (**Table 4**) was firstly converted into the three classes based on the type of ddNTP big dye terminator incorporated into the ambiguous type II restriction endonuclease site. Each restriction fragment for each BAC clone was then assigned a size (bp) based on the Rox500 size standard. Band sizes and ddNTP classes were analysed for each BAC clone and commonality in banding patterns was deciphered (data not shown).

3.3.5 BAC contig assembly

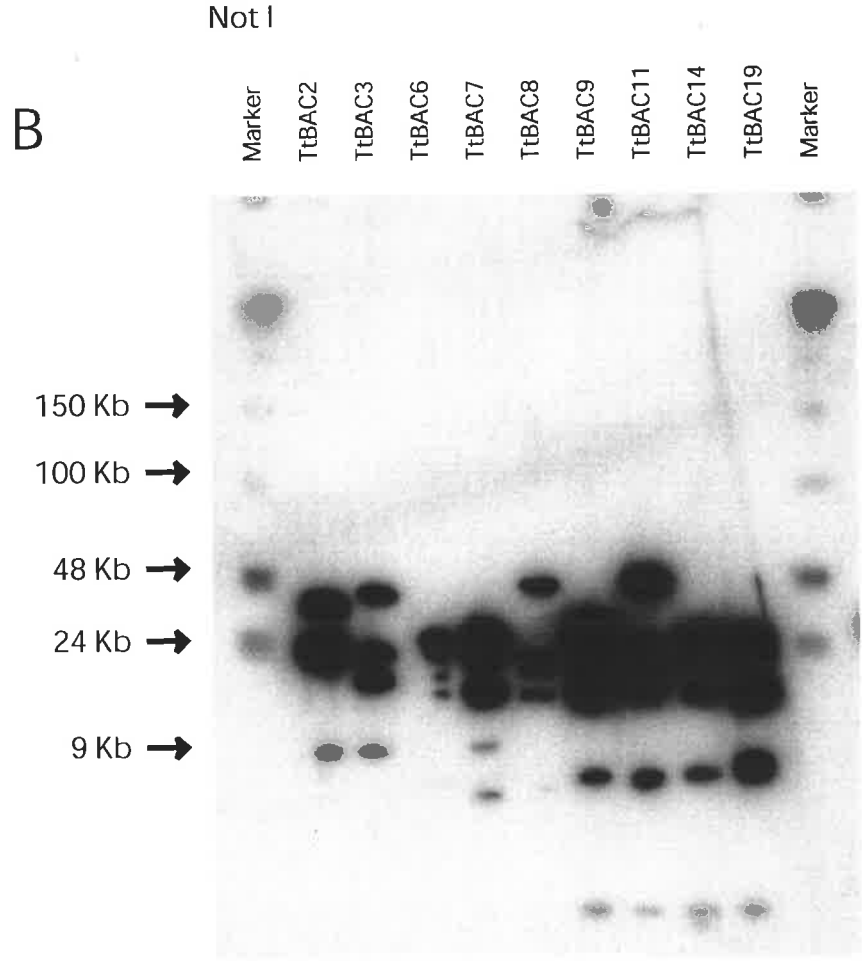
WM1.1 ORF probed Southerns of *Hind III*, *Dra I*, and *Eco RV* digested barley and wheat BACs were analysed manually for commonality in banding patterns inferring overlap. The results were compared to the *Not I* digested, PFGE data of the wheat BACs. This allowed an approximation of physical size to be assigned to component fragments and overall BAC insert size.

Figure 13: Pulsed field gel electrophesis (PFGE) separated wheat BAC DNA of nine overlapping clones known to encode *WM1.1*, *WM1.2* and *WM1.3*. Each of the nine overlapping BAC clones is restriction endonuclease digested with *Not I* to release insert DNA and size separated by PFGE (Panel A). Panel B shows Southern hybridisation of *WM1.1 ORF* to the size seperated BAC DNA for each of the nine BAC clones. Lanes labelled marker contain low range λ concatomer DNA ladder (New England Biolabs).

A


Not I
Marker
TtBAC2 - 61Kb
TtBAC3 - 81Kb
TtBAC6 - 62Kb
TtBAC7 - 36Kb
TtBAC8 - 90Kb
TtBAC9 - 36Kb
TtBAC11 - 92Kb
TtBAC14 - 38Kb
TtBAC19 - 38Kb
Marker
150 Kb
100 Kb
48 Kb
24 Kb
9 Kb
Vector CLD0454 (10.6 Kb)
Vector CLD0454 (5.2 Kb)

B


Not I
Marker
TtBAC2
TtBAC3
TtBAC6
TtBAC7
TtBAC8
TtBAC9
TtBAC11
TtBAC14
TtBAC19
Marker
150 Kb
100 Kb
48 Kb
24 Kb
9 Kb

Based on commonality in banding patterns between BAC clones for individual fingerprints (Section 3.3.4) a probable contig assignment of the three wheat BAC clones is illustrated (**Figure 14**). The contig generated manually through Southern analysis was confirmed by the BAC fingerprint analysis.

3.3.6 BAC sequence analysis

Sequence data generated from BAC subclones minus the BAC vector subclones (refer to Sections 3.2.9.1 to 3.2.9.3) for each of *TtBAC3*, *TtBAC6* and *TtBAC11* were compiled through the use of the bioinformatic software available at DuPont. Computer programs PHRED/PHRAP, CROSSMATCH and CONSED generated 46 independent stretches of DNA from 272617 bp of DNA sequence. Orientations of the DNA stretches were determined by comparison of the DNA sequence to Southern analysis data (3.3.3.1 and 3.3.3.2) and a database of barley and wheat transposable and repetitive elements (Wicker *et al.*, 2001). Gaps in the sequence were predominantly caused by $(G)_n$ repeats (n = ~20 bp). Other gaps were caused by misalignment of DNA sequences encoding the long terminal repeats (LTR) of transposable elements. **Figure 15** displays the predicted gene content within a 226 Kb subfragment from the 270 Kb of DNA sequence derived from *TtBAC3*, *TtBAC6*, and *TtBAC11*.

Of the total 56 coding sequences predicted from the 270 Kb of sequence, 17 were component genes of transposable elements. This demonstrates a gene density of one gene every 19 Kb excluding transposable element coding sequences and predicted proteins with no homology to Genbank.

Figure 14: Physical alignment of insert DNA of three overlapping wheat BAC clones (*TtBAC3*, *TtBAC6* and *TtBAC11*) based on Southern hybridisation, pulsed field gel electrophoresis and fingerprinting data. Physical location of *WM1.1*, *WM1.2* and *WM1.3* to individual *Not I* restriction fragments within the BAC contig is illustrated.

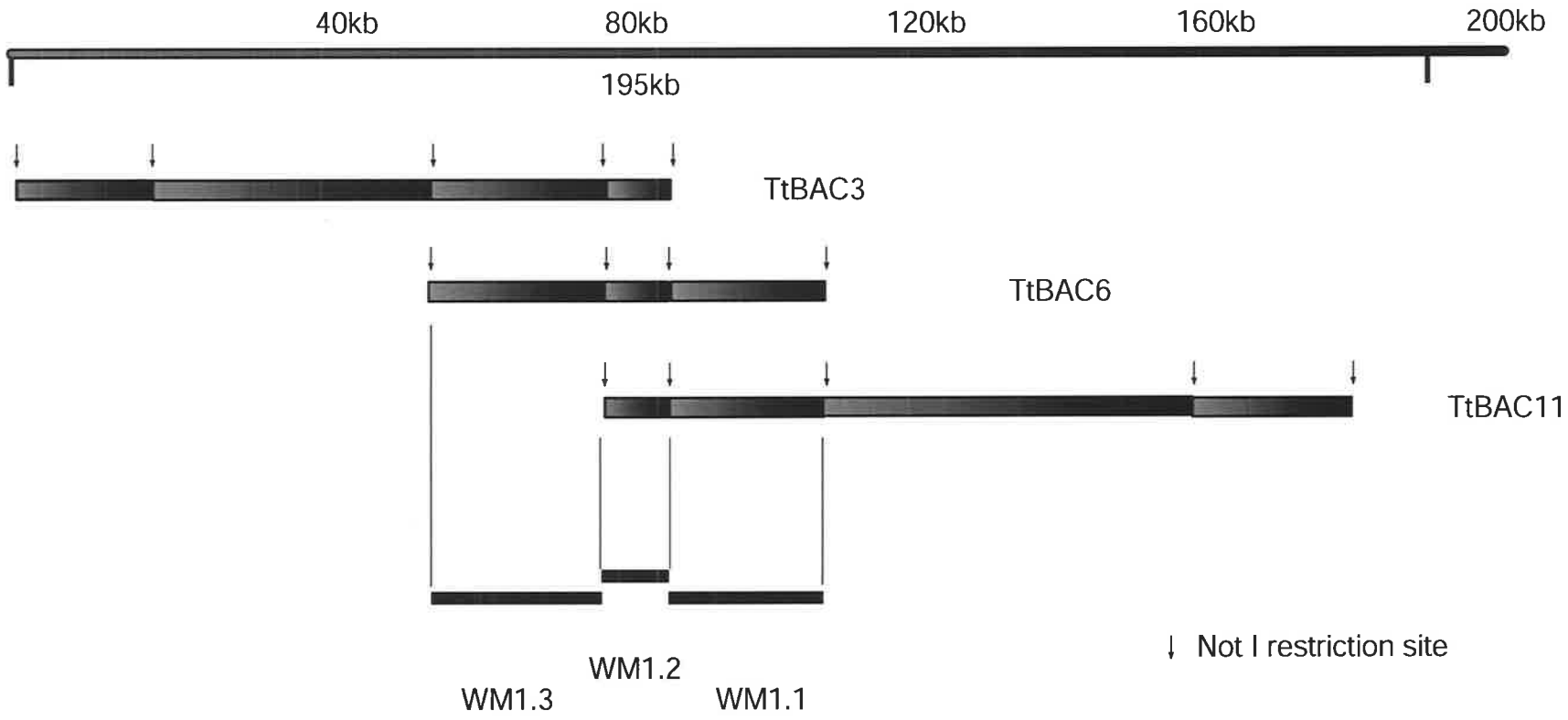

40kb
80kb
120kb
160kb
200kb
195kb
TtBAC3
TtBAC6
TtBAC11
Not I restriction site
WM1.2
WM1.3
WM1.1

Figure 15: Predicted gene content and structure of a 226 Kb DNA subfragment within 270 Kb of DNA sequence derived from *Triticum taushii* BAC clones *TtBAC3*, *TtBAC6*, *TtBAC11.*

Panel A – Diagramatic representation of the transposable and unclassified repetitive element distribution as determined by homology to a characterised transposable element database (Wicker, *pers. commun.*).

Panel B – Diagramatic representation of the RiceGAAS (http://rgp.dna.affrc.go.jp/) predicted genes, minus those genes with homology to transposable elements and predicted matrix attachment regions. Gaps in the DNA sequence are labelled as GAP.

Panel C – Diagramatic representation of the microsatellite distribution as determined by both REPEATMASKER (http://ftp.genome.washington.edu/) and GrailEXP (Hyatt *et al.*, 2000a; Hyatt *et al.*, 2000b).

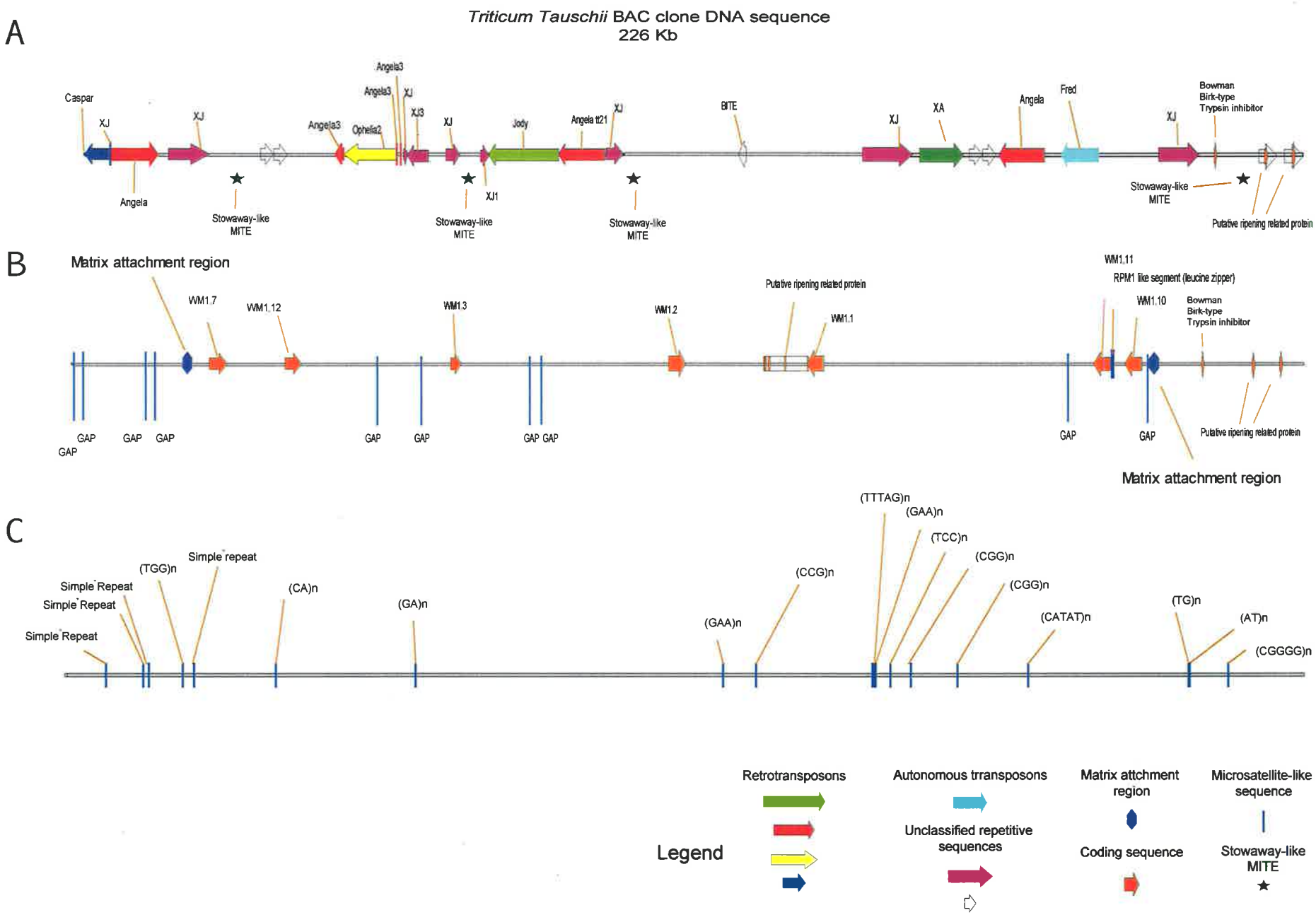

A
Triticum Tauschii BAC clone DNA sequence
226 Kb
Caspar
XJ
Angela
XJ
Stowaway-like MITE
Angela3
Angela3
Angela3
Ophelia2
XJ
XJ3
XJ
XJ1
Stowaway-like MITE
Jody
Angela tt21
XJ
Stowaway-like MITE
BITE
XJ
XA
Angela
Fred
XJ
Bowman Birk-type Trypsin inhibitor
Stowaway-like MITE
Putative ripening related protein
B
Matrix attachment region
WM1.7
WM1.12
WM1.3
WM1.2
Putative ripening related protein
WM1.1
WM1.11
RPM1 like segment (leucine zipper)
WM1.10
Bowman Birk-type Trypsin inhibitor
GAP
GAP
GAP
GAP
GAP
GAP
GAP
GAP
GAP
GAP
Putative ripening related protein
Matrix attachment region
C
Simple Repeat
Simple Repeat
Simple Repeat
(TGG)n
Simple repeat
(CA)n
(GA)n
(GAA)n
(CCG)n
(TTTAG)n
(GAA)n
(TCC)n
(CGG)n
(CGG)n
(CATAT)n
(TG)n
(AT)n
(CGGGG)n
Legend
Retrotransposons
Autonomous transposons
Unclassified repetitive sequences
Matrix attchment region
Coding sequence
Microsatellite-like sequence
Stowaway-like MITE

Table 6 lists the predicted gene content from the three BAC clones.

3.3.6.1 WM1 *genes*

Seven *WM1* genes (*WM1.1-1.3*, *WM1.7* and *WM1.10-12*) were predicted in the 270 Kb of *T. tauschii* DNA (**Figure 16**). Four (*WM1.1*, *WM1.2*, *WM1.7* and *WM1.10*) of the seven *WM1* genes have a single open reading frame suggesting that they do not contain introns.

The three remaining *WM1* genes (*WM1.3*, *WM1.11*, and *WM1.12*) have multiple stop codons within what seemed to be an original single open reading frame.

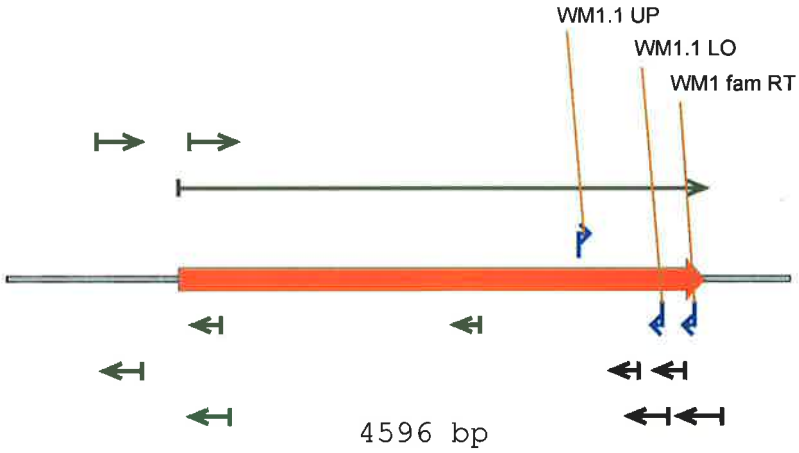
3.3.6.1.1 Prediction of transcription regulatory elements

The promoters of class II genes, such as the *WM1* genes, can be subdivided into the basal promoter located at the transcriptional start site and the upstream promoter elements, comprising of consititutive and regulatory motifs. Alignment of approximately 1 Kb of 5' DNA sequence to the predicted coding regions of the *WM1* genes and analysis of the resulting consensus sequence highlights the common location of the predicted basal promoter (**Figure 17** - A and B). Within the basal promoter there is a TATA-box, which is expected to be at a position – 25 from the transcription start site (**Figure 17** – A). Neural Network Promoter Prediction of this region including a CDF1-RS binding signature (**Figure 17** – A and B), found to bind to basal promoters of chloroplastic genes (Lam *et al.*, 1988) helps support the hypothesis that this region is responsible for RNA polymerase II binding. **Figure 17** - B shows that this basal-like promoter region also posesses a binding signature for plant specific Dof transcription factors (Yanagisawa, 1997; Yanagisawa, 1998; Yanagisawa, 2000; Yanagisawa & Schmidt, 1999; Yanagisawa & Sheen, 1998). It is anticipated that CDF1-RS proteins represent basal transcription initiation factors and will form a part of the pre-initation complex at the basal promoter. This is a likely requirement for RNA polymerase binding and subsequent activity.

Table 6: RiceGAAS software identified coding sequences from 270 Kb of *Triticum tauschii* genomic sequence.

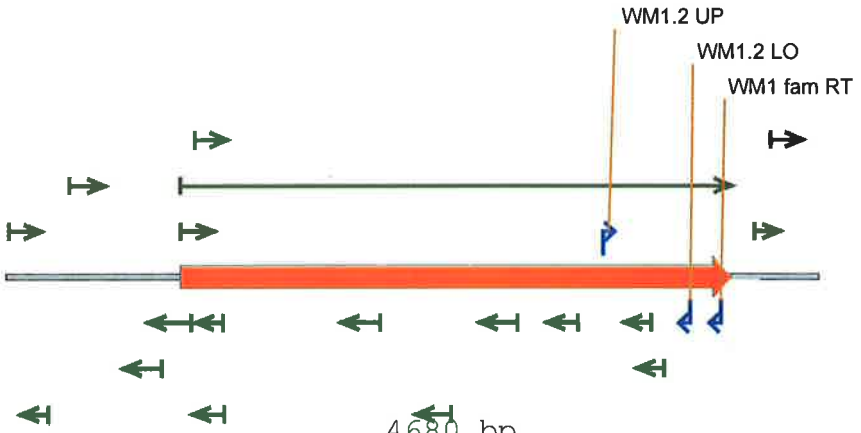
Predicted Proteins	Number of encoding genes
WM1 family proteins	7
Putative ripening-related proteins from grapevine	3
Bowman-Birk type trypsin inhibitor proteins	4
Transposable element component proteins with homology to hypothetical proteins from *Arabidopsis thaliana*, *Oryza sativa*, *Triticum aestivum*.	17
Predicted proteins	25
Total	**56**

Figure 16: Diagram of the predicted nucleotide coding sequence for seven members of the *WM1* gene family. For each predicted coding sequence (orange arrows), green arrows outline open reading frame maps and blue arrows designate the relative position of RT-PCR primers used to distinguish the different members of the gene family.



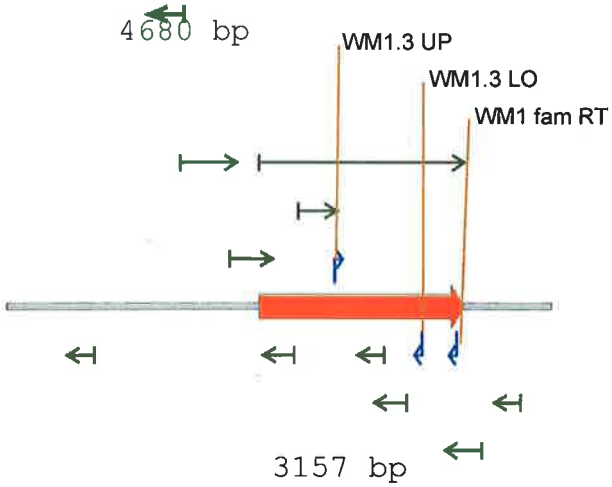
WM1.1

4596 bp



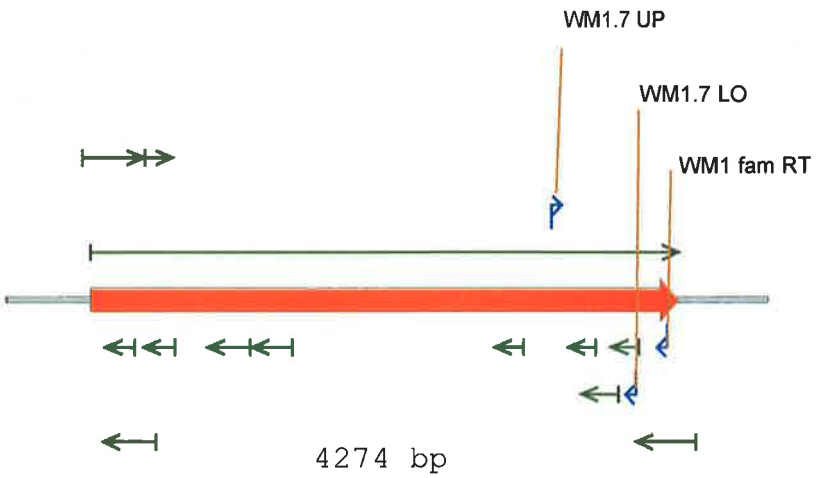
WM1.2

4680 bp



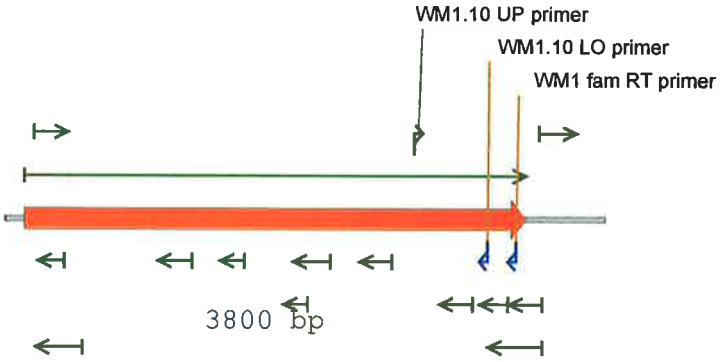
WM1.3

3157 bp



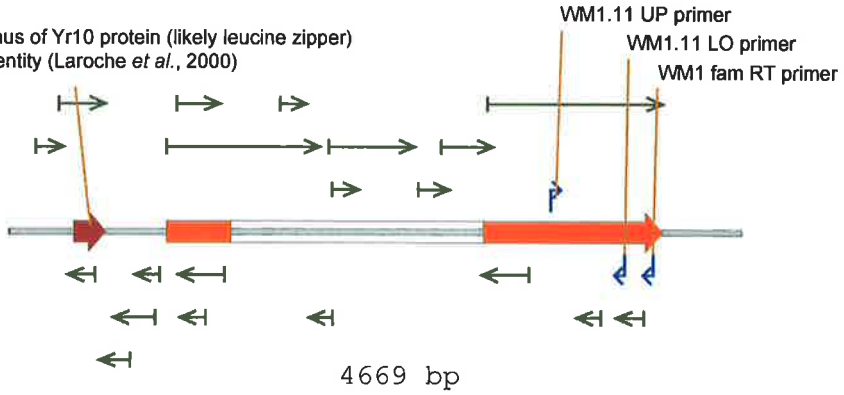
WM1.7

4274 bp

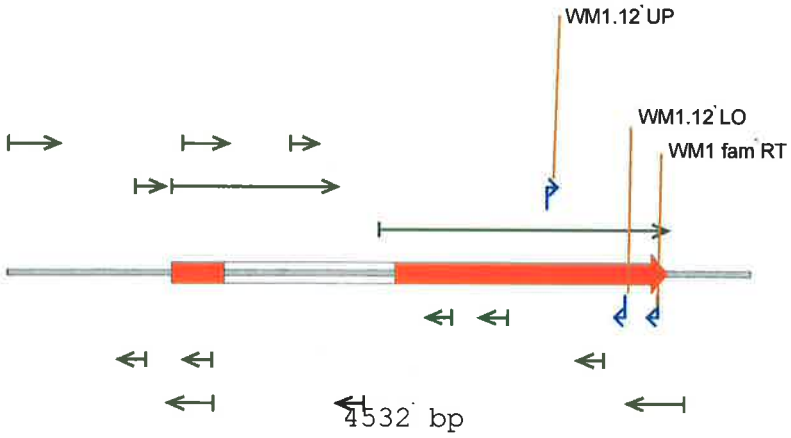


WM1.10

Homology N-terminus of Yr10 protein (likely leucine zipper)
62% amino acid identity (Laroche *et al.*, 2000)



WM1.11



WM1.12

Figure 17: Promoter and terminator analysis of the predicted *WM1* gene family.

A – Prettybox alignment of DNA sequence 5' to the predicted coding regions for each of *WM1.1*, *WM1.2*, *WM1.7, WM1.10-12.* Highlighted by a red line is the predicted basal promoter; a blue line, the predicted TATA-box; in green letters is the predicted binding site for the CDF1-RS *cis*-regulatory element and red letters, the predicted transcription start site.

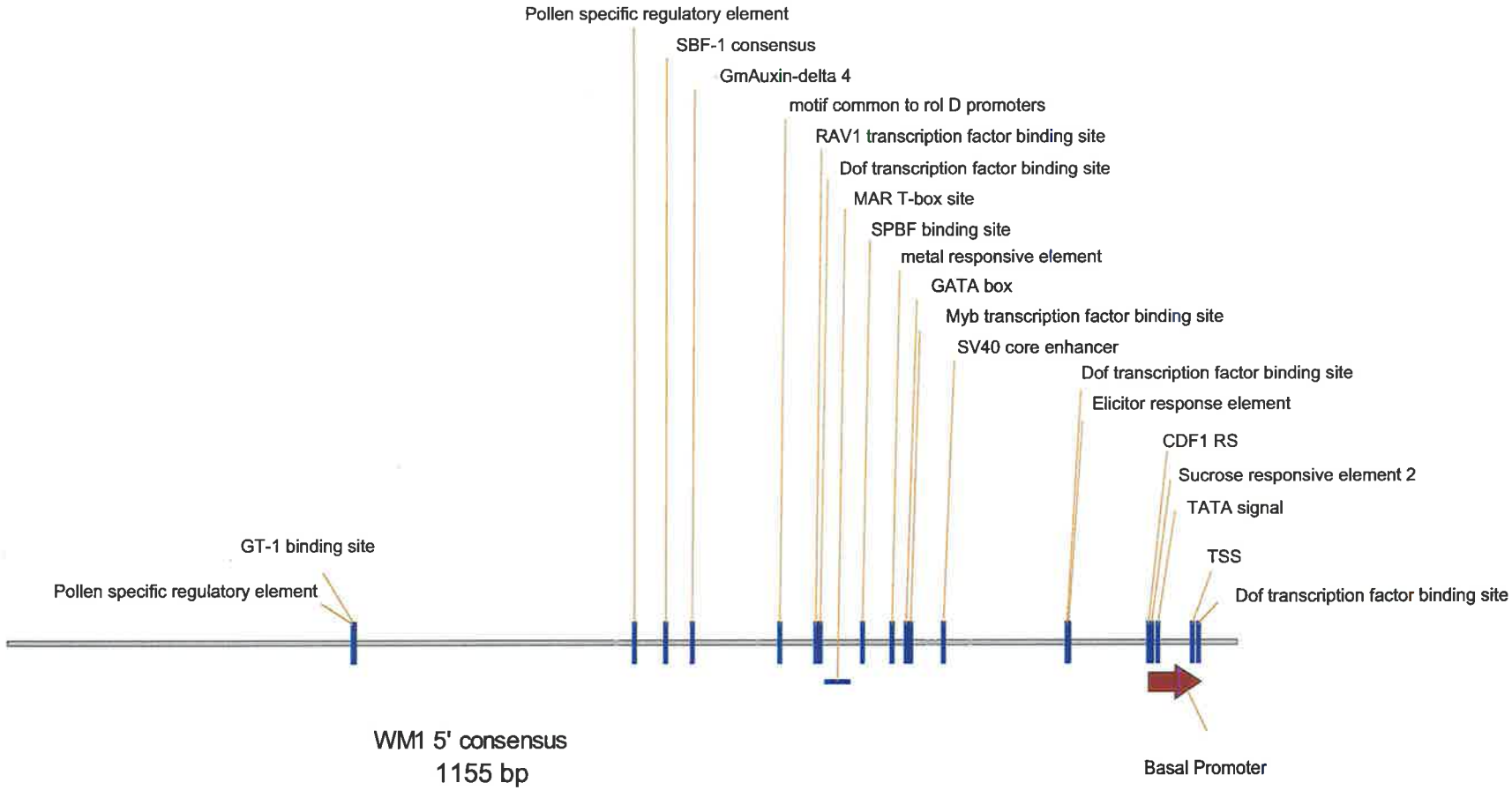
B – 1. Diagram of the predicted *cis*-acting regulatory element binding sites on *WM1* gene family promoter consensus sequence. The consensus sequence produced by an AlignX alignment of 1 Kb of DNA sequence 5' to the predicted coding regions for each of *WM1.1*, *WM1.2*, *WM1.7, WM1.10-12.*

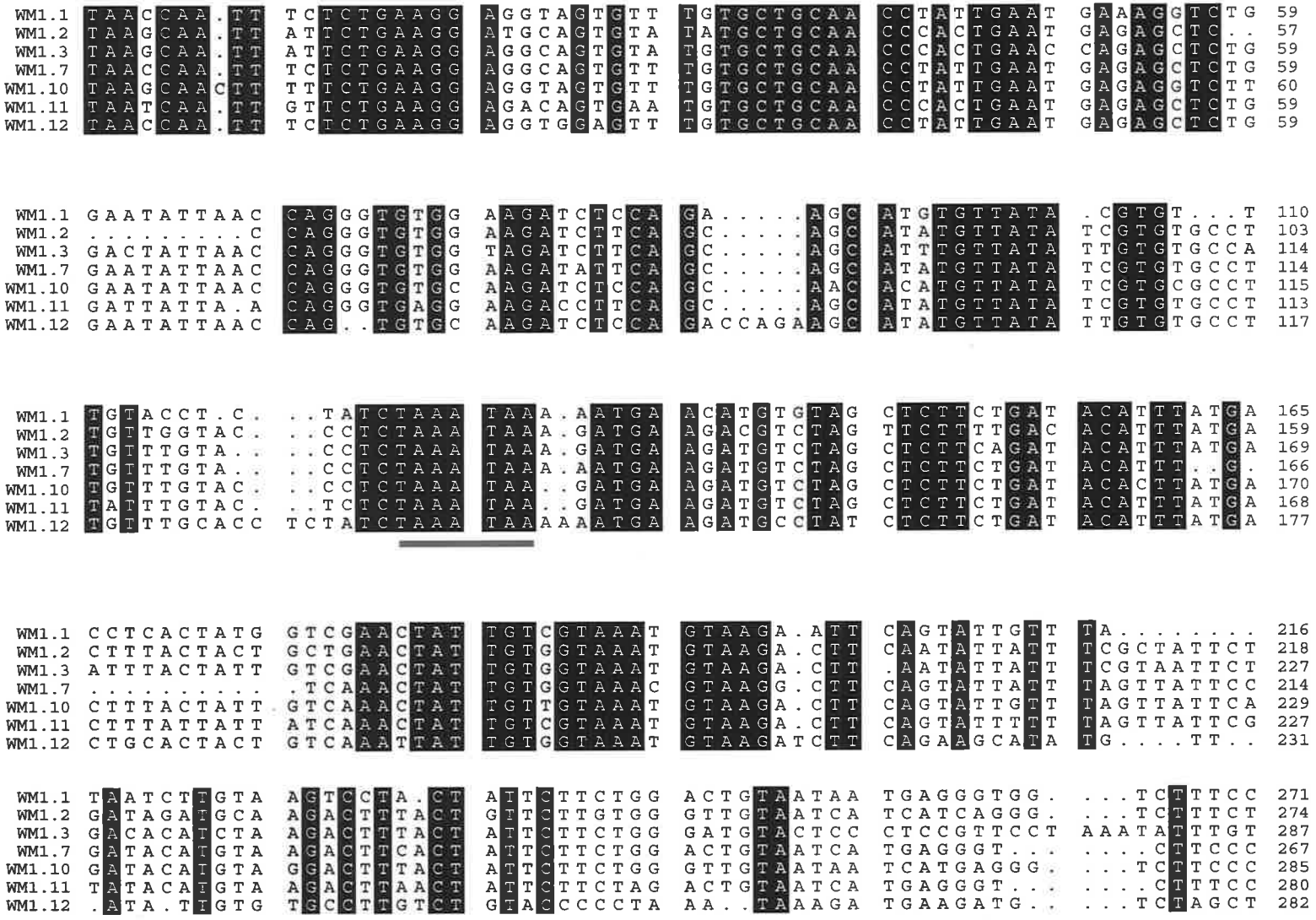
C – Prettybox alignment of 500 bp of DNA sequence 3' to the predicted coding regions for each of *WM1.1*-3, *WM1.7, WM1.10-12.* Predicted A and T rich poly-adenylation signals are highlighted by red lines.

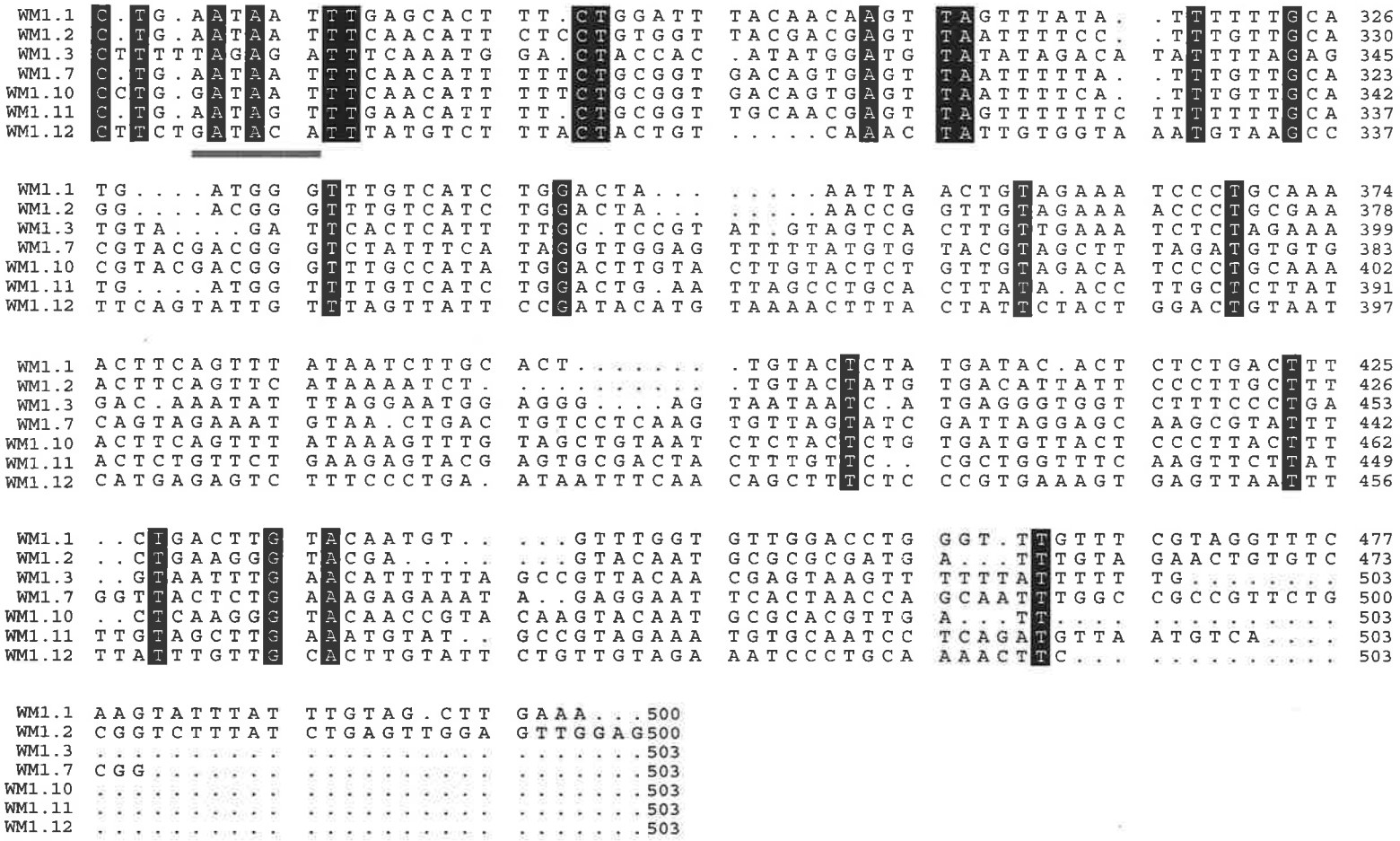
A


WM1.1 AGTCGAGTGG CTGGTGACTC CTAACATAGT CTGGCCACAA ATTTGTC... ...TGAACCA 815
WM1.11 AGTTGAGTAG CCGC..... ....CAA AGTTGAT... ...TGAATGG 832
WM1.12 AGTCAAGTGG CTGGTGACTC CCATCACACT CTGGCCACAA ATTTGTA... ...TGAACCA 871
WM1.2 AGTC.AGTGG CCGC..... ....CAG AGTTGAT... ...TGAATGG 848
WM1.7 AAATAGCTTG CTGA..... ..AAAAAAAT TACTGGTTAA AGATCTG... ...GGAATCA 288
WM1.10 0
WM1.1 AA.CT..T.. GATTCAGTTC CCTGAAGGAC GGCTACTTTG ACCGTAGACT CCATGAGTCT 870
WM1.11 GATCT..C.. CATGGAGCAC A..GAAGGAC GAAGACTTTG ACCGCAGACT CCGTCGGTCT 886
WM1.12 GA..... ....CTTTG ACCGTAGACT CCGTGAGTCT 898
WM1.2 GATCT..C.. CATGGAGCAC C..CAAGGAC GACAACTTTG ACCGTAGACT CCATCGGTCA 902
WM1.7 TACCA..AAG GACTCAGTTC CGTGAAGGAC GGGTACTTGG ACCGTAGA.. .....GTCT 338
WM1.10 ....TCT 3
WM1.1 AGGTATATTC TGGCCAAACA TCTCCGGTAG TA...TGGCT CTGGTAC.GT ACCTAAATAC 926
WM1.11 AGTTTTG... TGGCCACACA TCCCTGGCAG CA...TAGCT CTGCCACAGT AGCTAAATAC 940
WM1.12 AGGAAT..TC TGGCCAAACA TCTCTGCTAG TA...TGGCT CTGGTAC.GT ACCTAAATAC 952
WM1.2 A......... .... .CCATGGCAG CA...TAGCT CTGGCACAAT AGCTAAATAC 939
WM1.7 AGGTAT..TC TGGCCAAACA TCTCTGGTAG TA...TGGCT CTTCTAC.AT AC.TAAATAC 391
WM1.10 AGTTTTG... TGGCCCACAT CCCTGGGCAG CA...TAGCT CTGCCACAGT AGCTAAATAC 57
WM1.1 TAATAAAACT GGGCCTGTAG CATACTCTAG TGTGCATGCT TTGCTTCCCC AGCTAGACAG 986
WM1.11 TAA.....CC GAGCCTGTAA CATACTCT.. ....GCT TTGATTCCCC AGCTAGACAG 986
WM1.12 AAAT..... .... AAAACTGG.. ....GCT CTGCTTCC.C AGCTAGACGG 986
WM1.2 TAC.....CC GAGCCTGTAA CATACTCT.. ....GCT TTGCTTCCCC AGCTAGACAG 985
WM1.7 TATAAAACT GGGCCTGTAG CATACTCTAG TGTGTATGCA TTGCTTCCTC AGCTAGACGG 451
WM1.10 TAA.....CC GAGCCTGTAA CATACTCT.. ....GCT TTGATTCCCC AGCTAGACAG 103
WM1.1 .CTCGTTCTT CCAGC1000
WM1.11 .CTCGTTCTT CCAGC1000
WM1.12 .CTCGTTCTT CCAGC1000
WM1.2 GCTCGTTCTT CCAGC1000
WM1.7 .CTCGTTCTT CCAGC465
WM1.10 .CTCGTTCTT CCAGC117

B


Pollen specific regulatory element
SBF-1 consensus
GmAuxin-delta 4
motif common to rol D promoters
RAV1 transcription factor binding site
Dof transcription factor binding site
MAR T-box site
SPBF binding site
metal responsive element
GATA box
Myb transcription factor binding site
SV40 core enhancer
Dof transcription factor binding site
Elicitor response element
CDF1 RS
Sucrose responsive element 2
TATA signal
TSS
Dof transcription factor binding site
GT-1 binding site
Pollen specific regulatory element
WM1 5' consensus
1155 bp
Basal Promoter


WM1.1 TAACCAA.TT TCTCTGAAGG AGGTAGTGTT TGTGCTGCAA CCTATTGAAT GAAAGGTCTG 59
WM1.2 TAAGCAA.TT ATTCTGAAGG ATGCAGTGTA TTATGCTGCAA CCCACTGAAT GAGAGCTC.. 57
WM1.3 TAAGCAA.TT ATTCTGAAGG AGGCACTGTA TGTGCTGCAA CCCACTGAAC CAGAGCTCTG 59
WM1.7 TAACCAA.TT TCTCTGAAGG AGGCAGTGTT TGTGCTGCAA CCCTATTGAA TGAGAGCTCT 59
WM1.10 TAAGCAACTT TTTCTGAAGG AGGTAGTGTT TGTGCTGCAA CCCTATTGAA TGAGAGGTCT 60
WM1.11 TAATCAA.TT GTTCTGAAGG AGACAGTGAA TGTGCTGCAA CCCACTGAAT GAGAGCTCTG 59
WM1.12 TAACCAA.TT TCTCTGAAGG AGGTGGAGTT TGTGCTGCAA CCCTATTGAA TGAGAGCTCTG 59
WM1.1 GAATATTAAC CAGGGTGTGG AAGATCTTCA GA.....AGC ATGTGTTATA .CGTGT...T 110
WM1.2 ....C CAGGGTGTGG AAGATCTTCA GC.....AGC ATATGTTATA TCGTGTGCCT 103
WM1.3 GACTATTAAC CAGGGTGTGG TAGATCTTCA GC.....AGC ATTTGTTATA TTGTGTGCCA 114
WM1.7 GAATATTAAC CAGGGTGTGG AAGATATTCA GC.....AGC ATATGTTATA TCGTGTGCCT 114
WM1.10 GAATATTAAC CAGGGTGTGC AAGATCTTCA GC.....AAC ACATGTTATA TCGTGCGCCT 115
WM1.11 GATTATTA.A CAGGGTGAGG AAGACCTTCA GC.....AGC ATATGTTATA TCGTGTGCCT 113
WM1.12 GAATATTAAC CAG..TGTGC AAGATCTTCA GACCAGAAGC ATATGTTATA TTGTGTGCCT 117
WM1.1 TGTACCT.C. ..TATCTAAA TAAA.AATGA ACATGTGTAG CTCTTCTGAT ACATTTATGA 165
WM1.2 TGTTGGTAC. ..CCTCTAAA TAAA.GATGA AGACGTCTAG TTCTTTTGAC ACATTTATGA 159
WM1.3 TGTTTGTA.. ..CCTCTAAA TAAA.GATGA AGATGTCTAG CTCTTCAGAT ACATTTATGA 169
WM1.7 TGTTTGTA.. ..CCTCTAAA TAAA.AATGA AGATGTCTAG CTCTTCTGAT ACATTTT..G. 166
WM1.10 TGTTTGTAC. ..CCTCTAAA TAA..GATGA AGATGTCTAG CTCTTCTGAT ACACTTATGA 170
WM1.11 TATTTGTAC. ..TCTCTAAA TAA..GATGA AGATGTCTAG CTCTTCTGAT ACATTTATGA 168
WM1.12 TGTTTGCACC TCTATCTAAA TAAAAAATGA AGATGCCTAT CTCTTCTGAT ACATTTATGA 177
WM1.1 CCTCACTATG GTCGAACTAT TGTCGTAAAT GTAAGA.ATT CAGTATTGTT TA........ 216
WM1.2 CTTTACTACT GCTGAACTAT TGTGGTAAAT GTAAGA.CTT CAATATTATT TCGCTATTCT 218
WM1.3 ATTTACTATT GTCGAACTAT TGTGGTAAAT GTAAGA.CTT .AATATTATT TCGTAATTCT 227
WM1.7 .... .TCAAACTAT TGTGGTAAAC GTAAGG.CTT CAGTATTATT TAGTTATTCC 214
WM1.10 CTTTACTATT GTCAAACTAT TGTTGTAAAT GTAAGA.CTT CAGTATTGTT TAGTTATTCA 229
WM1.11 CTTTATTATT ATCAAACTAT TGTCGTAAAT GTAAGA.CTT CAGTATTTTT TAGTTATTCG 227
WM1.12 CTGCACTACT GTCAAATTAT TGTGGTAAAT GTAAGATCTT CAGAAGCATA TG....TT.. 231
WM1.1 TAATCTTGTA AGTCCTTA.CT ATTCTTTCTGG ACTGTAATAA TGAGGGTGG. ...TCTTTCC 271
WM1.2 GATAGATGCA AGACTTTACT GTTCTTGTGG GTTGTAATCA TCATCAGGG. ...TCTTTCT 274
WM1.3 GACACATCTA AGACTTTACT ATTCTTCTGG GATGTACTCC CTCCGTTCCT AAATATTTGT 287
WM1.7 GATACATGTA AGACTTTCAC ATTCTTCTGG ACTGTAATCA TGAGGGT... ....CTTTCCC 267
WM1.10 GATACATGTA GGACTTTACT ATTCTTCTGG GTTGTAATAA TCATGAGGG. ...TCTTTCCC 285
WM1.11 TATACATGTA AGACTTTAAC ATTCTTCTAG ACTGTAATCA TGAGGGT... ....CTTTCCC 280
WM1.12 .ATA.TTGTG TGCCTTTGTCT GTACCCCCTA AA..TAAAGA TGAAGATG.. ...TCTAGCT 282


WM1.1 C.TG.AATAA TTTGAGCACT TT.CTGGATT TACAACAAGT TAGTTTATA. .TTTTTTGCA 326
WM1.2 C.TG.AATAA TTTCAACATT CTCCTTGTGGT TACGACGAGT TAATTTTCC. .TTTTGTTGCA 330
WM1.3 CTTTTTAGAG ATTTTCAAATG GA.CTTACCAC .ATATGGATG TATATAGACA TATTTTTAGAG 345
WM1.7 C.TTG.AATAA TTTTTCAACATT TTTCTTGCGGT GACAGTGAGT TAATTTTTA. .TTTTGTTGCA 323
WM1.10 CCTTG.GATAA TTTTTCAACATT TTTTCTTGCGGT GACAGTGAGT TAATTTTTCA. .TTTTGTTGCA 342
WM1.11 C.TTG.AATAG TTTTTGAACATT TTT.CTTGCGGT TGCAACGAGT TAGTTTTTTTC TTTTTTTTGCA 337
WM1.12 CTTCTGATAC ATTTATGTCCT TTACTTACTGT .....CAAAC TATTGTGGTA AATGTAAGCC 337
WM1.1 TG....ATGG GTTTGTCATC TGGACTA... .....AATTA ACTGTAGAAA TCCCTGCAAA 374
WM1.2 GG....ACGG GTTTGTCATC TGGACTA... .....AACCG GTTGTAGAAA ACCCTGCGAA 378
WM1.3 TGTA....GA TTCACTCATT TTGC.TCCGT AT.GTAGTCA CTTGTTGAAA TCTCTAGAAA 399
WM1.7 CGTACGACGG GTCTATTTCA TAGGTTGGAG TTTTTTATGTG TACGTAGCTT TAGATGTGTG 383
WM1.10 CGTACGACGG GTTTGCCATA TGGACTTGTA CTTGTACTCT GTTGTAGACA TCCCTGCAAA 402
WM1.11 TG....ATGG TTTTGTCATC TGGACTG.AA TTAGCCTGCA CTTATA.ACC TTGCTCTTAT 391
WM1.12 TTCAGTATTG TTTAGTTATT CCGATACATG TAAAACTTTA CTATTCTACT GGACTGTAAT 397
WM1.1 ACTTCAGTTT ATAATCTTGC ACT..... .TGTACTCTA TGATAC.ACT CTCTGACTTT 425
WM1.2 ACTTCAGTTC ATAAAATCT. ..... .TGTACTTATG TGACATTATT CCCTTGCTTT 426
WM1.3 GAC.AAATAT TTAGGAATGG AGGG....AG TAATAATTC.A TGAGGGTGGT CTTTCCCTGA 453
WM1.7 CAGTAGAAAT GTAA.CTGAC TGTCCTCAAG TGTTAGTATC GATTAGGAGC AAGCGTATTT 442
WM1.10 ACTTCAGTTT ATAAAGTTTG TAGCTGTAAT CTCTACTCTG TGATGTTACT CCCTTACTTT 462
WM1.11 ACTCTGTTCT GAAGAGTACG AGTGCGACTA CTTTGTTTC.. CGCTGGTTTC AAGTTCTTAT 449
WM1.12 CATGAGAGTC TTTCCCTGA. ATAATTTCAA CAGCTTTCTC CCGTGAAAGT GAGTTAATTT 456
WM1.1 ..CTGACTTG TACAATGT.. ...GTTTGGT GTTGGACCTG GGT.TTGTTT CGTAGGTTTC 477
WM1.2 ..CTGAAGGG TACGA..... ...GTACAAT GCGCGCGATG A...TTTGTA GAACTGTGTC 473
WM1.3 ..GTAATTTG AACATTTTTA GCCGTTACAA CGAGTAAGTT TTTTATTTTT TG..... 503
WM1.7 GGTTACTCTG AAAGAGAAAT A..GAGGAAT TCACTAACCA GCAATTTGGC CGCCGTTCTG 500
WM1.10 ..CTCAAGGG TACAACCGTA CAAGTACAAT GCGCACGTTG A...TTT..... 503
WM1.11 TTGTAGCTTG AAATGTAT.. GCCGTAGAAA TGTGCAATCC TCAGATGTTA ATGTCA.... 503
WM1.12 TTATTTGTTG CACTTGTATT CTGTTGTAGA AATCCCTGCA AAACTTC... 503
WM1.1 AAGTATTTAT TTGTAG.CTT GAAA...500
WM1.2 CGGTCTTTAT CTGAGTTGGA GTTGGAG500
WM1.3 503
WM1.7 CGG..... 503
WM1.10 503
WM1.11 503
WM1.12 503

The prediction of the basal promoter is supported by the identification by BLAST 2.0 blastn program (Altschul *et al.*, 1997), of an Expressed Sequence Tag (EST) from a bread wheat endosperm, 8-12 days post anthesis, cDNA library. This EST encodes the very 5' end of the coding sequence for *WM1.11*, which initiates 7 base pairs 3' to the TATA-box (**Figure 17** – A and B).

The 5' consensus sequence for the predicted *WM1* genes does not clearly identify any unique regulatory element-binding sites upstream from the predicted basal promoter. Instead many regulatory element-binding sites were predicted (**Figure 17** - B). Localisation of these regulatory element-binding sites suggests that the predicted *WM1* genes are likely to possess a collection of upstream promoter elements.

3.3.6.1.2 Prediction of poly-adenylation signals

Most RNA polymerase II transcripts are processed by 3' cleavage and polyadenylation, however the intrinsic termination of transcription is unclear. It is expected that polyadenylation signals play a crucial part in transcription termination. The 3' consensus sequence for the predicted *WM1* genes clearly possesses two poly-adenylation signals (**Figure 17**- C).

3.3.6.2 WM1 *predicted primary peptide structure*

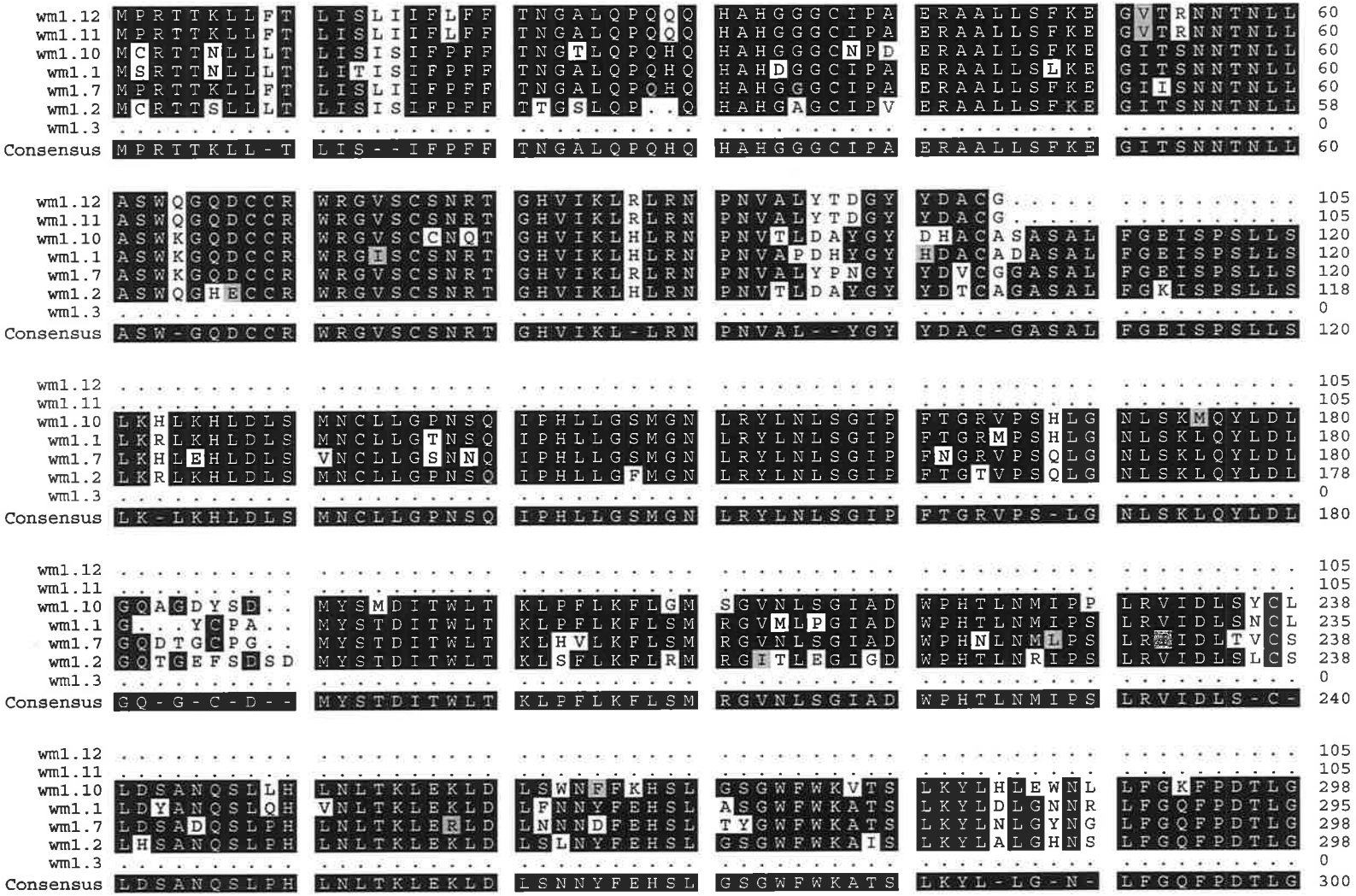
RiceGAAS predicted primary peptides for the *WM1* genes show high levels of sequence homology (**Table 7**). Primary polypeptide comparisons show that the WM1 proteins have between 70.5% and 85.9% similarity, with sequence divergence between 13.8% and 31.6%. The predicted phylogenetic relationships between WM1 predicted proteins illustrates that *WM1.3* is the most distantly related of the *WM1* gene family (**Figure 18**). Based on the number of amino acid substitution events distinguishing members of the *WM1* gene family it seems as though *WM1.1*, *WM1.10* and *WM1.12* are the most closely related. *WM1.1*, *WM1.10* and *WM1.12* can be distinguished based upon between 8 and 9 amino acid substitution events across a total of 660 amino acids (**Figure 18**).

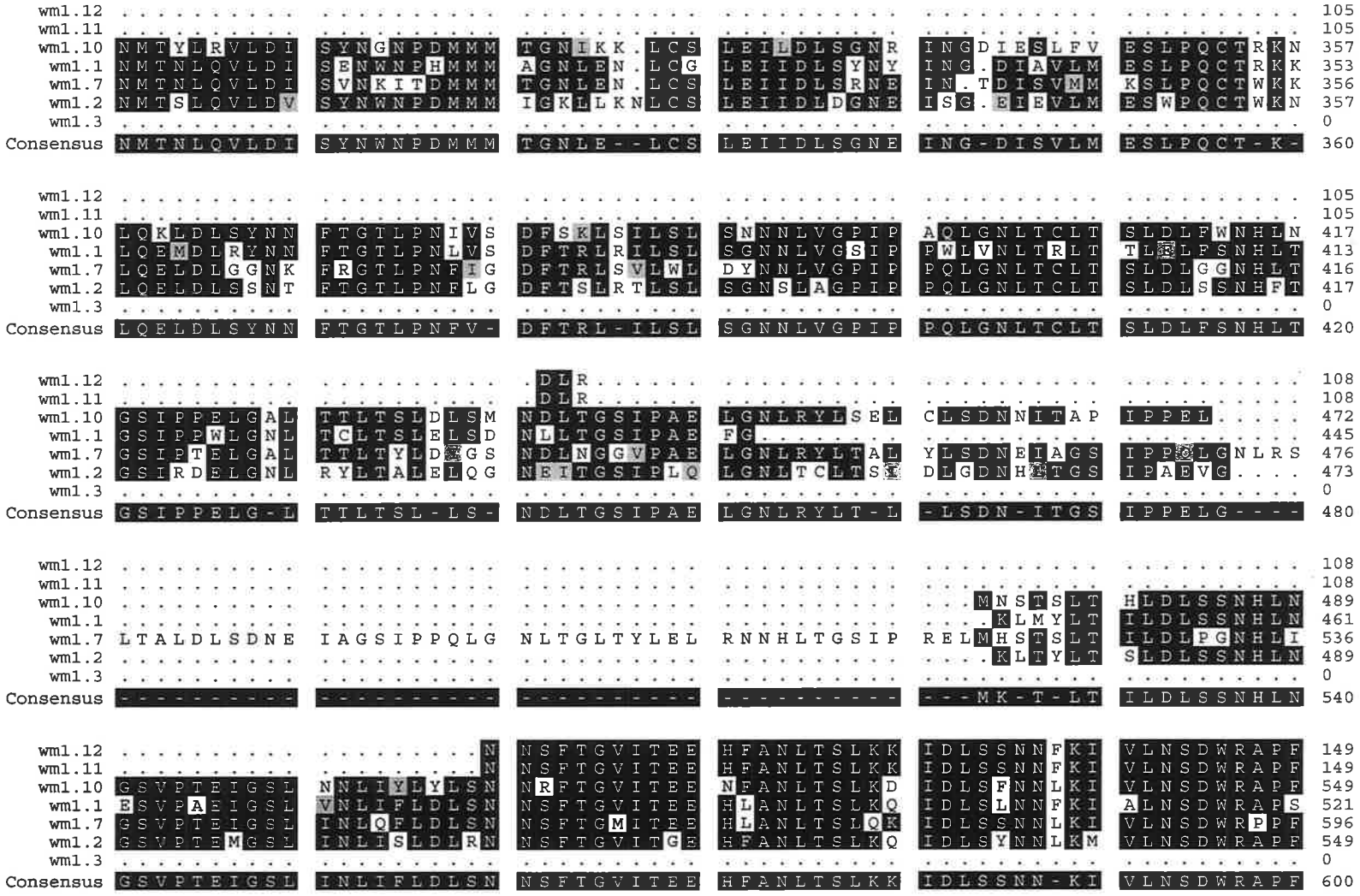
Figure 18: Relationship between the predicted primary polypeptide sequences for seven members of the *WM1* gene family.

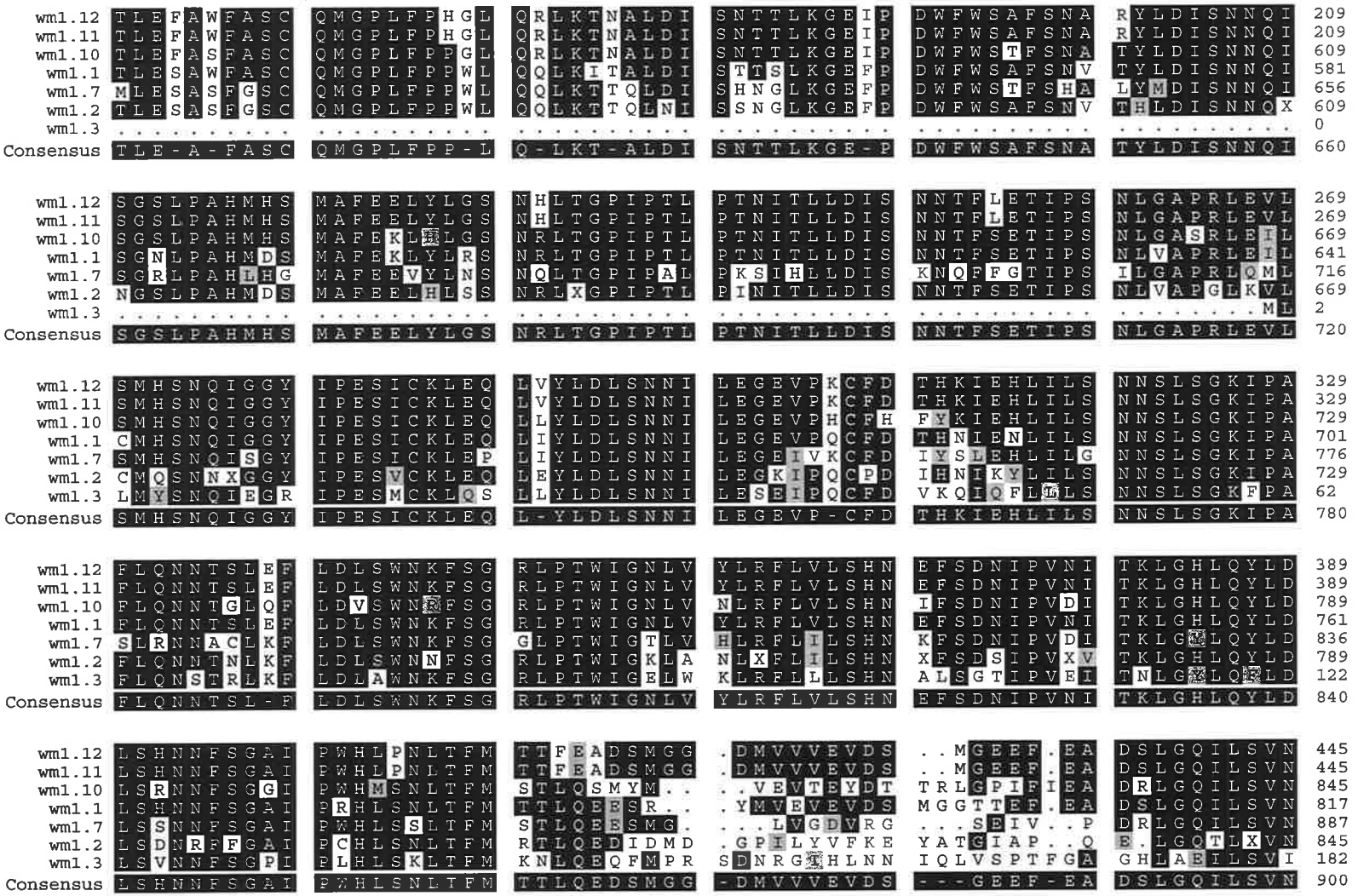
A - Prettybox alignment of the predicted primary polypeptide sequences for seven members of the *WM1* gene family.

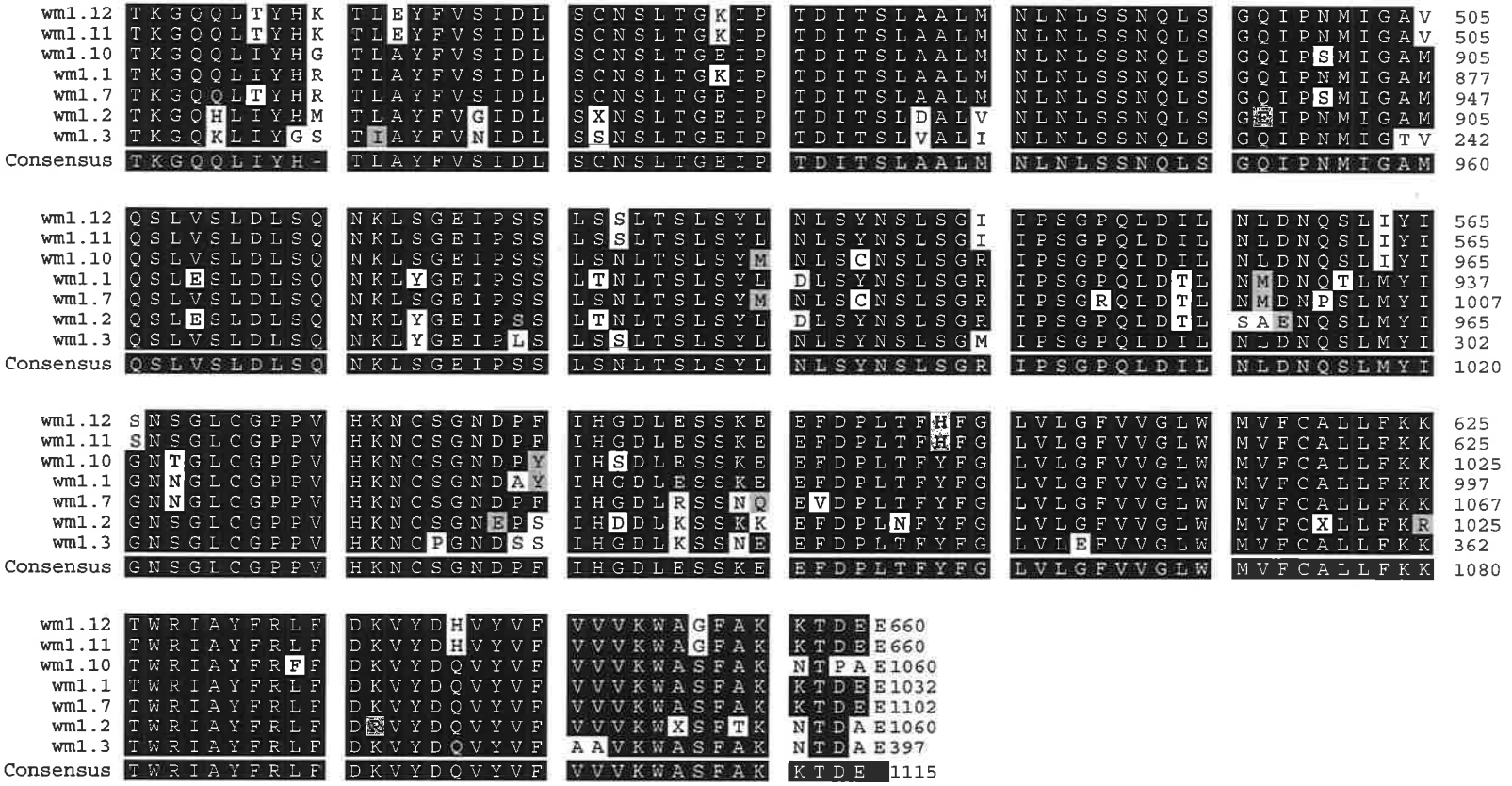
B – A rooted phylogenetic tree depicting the relationship between the predicted primary polypeptide sequences for seven members of the *WM1* gene family. Phylogentic tree describes the number of substitution events at each branch point that allow discrimination between related members of the *WM1* gene family (DNAstar Inc., 1997).

A


wm1.12 MPRTTKLLFT LISLIIFLFF TNGALQPQQQ HAHGGGCIPA ERAALLSFKE GVTRNNTNLL 60
wm1.11 MPRTTKLLFT LISLIIFLFF TNGALQPQQQ HAHGGGCIPA ERAALLSFKE GVTRNNTNLL 60
wm1.10 MCRTTNLLLT LISISIFPFF TNGTLQPQHQ HAHGGGCNPD ERAALLSFKE GITSNNTNLL 60
wm1.1 MSRTTNLLLT LITISIFPFF TNGALQPQHQ HAHDGGCIPA ERAALLSLKE GITSNNTNLL 60
wm1.7 MPRTTKLLFT LISLIIFPFF TNGALQPQHQ HAHGGGCIPA ERAALLSFKE GIISNNTNLL 60
wm1.2 MCRTTSLLLT LISISIFPFF TTGSLQP..Q HAHGAGCIPV ERAALLSFKE GITSNNTNLL 58
wm1.3 0
Consensus MPRTTKLL-T LIS--IFPFF TNGALQPQHQ HAHGGGCIPA ERAALLSFKE GITSNNTNLL 60
wm1.12 ASWQGQDCCR WRGVSCSNRT GHVIKLRLRN PNVALYTDGY YDACG 105
wm1.11 ASWQGQDCCR WRGVSCSNRT GHVIKLRLRN PNVALYTDGY YDACG 105
wm1.10 ASWKGQDCCR WRGVSCCNQT GHVIKLHLRN PNVTLDAYGY DHACASASAL FGEISPSLLS 120
wm1.1 ASWKGQDCCR WRGISCSNRT GHVIKLHLRN PNVAPDHYGY HDACADASAL FGEISPSLLS 120
wm1.7 ASWKGQDCCR WRGVSCSNRT GHVIKLRLRN PNVALYPNGY YDVCGGASAL FGEISPSLLS 120
wm1.2 ASWQGHECCR WRGVSCSNRT GHVIKLHLRN PNVTLDAYGY YDTCAGASAL FGKISPSLLS 118
wm1.3 0
Consensus ASW-GQDCCR WRGVSCSNRT GHVIKL-LRN PNVAL--YGY YDAC-GASAL FGEISPSLLS 120
wm1.12 105
wm1.11 105
wm1.10 LKHLKHLDLS MNCLLGPNSQ IPHLLGSMGN LRYLNLSGIP FTGRVPSHLG NLSKMQYLDL 180
wm1.1 LKRLKHLDLS MNCLLGTNSQ IPHLLGSMGN LRYLNLSGIP FTGRMPSHLG NLSKLQYLDL 180
wm1.7 LKHLEHLDLS VNCLLGSNNQ IPHLLGSMGN LRYLNLSGIP FNGRVPSQLG NLSKLQYLDL 180
wm1.2 LKRLKHLDLS MNCLLGPNSQ IPHLLGFMGN LRYLNLSGIP FTGTVPSQLG NLSKLQYLDL 178
wm1.3 0
Consensus LK-LKHLDLS MNCLLGPNSQ IPHLLGSMGN LRYLNLSGIP FTGRVPS-LG NLSKLQYLDL 180
wm1.12 105
wm1.11 105
wm1.10 GQAGDYSD.. MYSMDITWLT KLPFLKFLGM SGVNLSGIAD WPHTLNMIPP LRVIDLSYCL 238
wm1.1 G...YCPA.. MYSTDITWLT KLPFLKFLSM RGVMLPGIAD WPHTLNMIPS LRVIDLSNCL 235
wm1.7 GQDTGCPG.. MYSTDITWLT KLHVLKFLSM RGVNLSGIAD WPHNLNMLPS LR[illegible]IDLTVCS 238
wm1.2 GQTGEFSDSD MYSTDITWLT KLSFLKFLRM RGITLEGIGD WPHTLNRIPS LRVIDLSLCS 238
wm1.3 0
Consensus GQ-G-C-D-- MYSTDITWLT KLPFLKFLSM RGVNLSGIAD WPHTLNMIPS LRVIDLS-C- 240
wm1.12 105
wm1.11 105
wm1.10 LDSANQSLLH LNLTKLEKLD LSWNFFKHSL GSGWFWKVTS LKYLHLEWNL LFGKFPDTLG 298
wm1.1 LDYANQSLQH VNLTKLEKLD LFNNYFEHSL ASGWFWKATS LKYLDLGNNR LFGQFPDTLG 295
wm1.7 LDSADQSLPH LNLTKLERLD LNNNDFEHSL TYGWFWKATS LKYLNLGYNG LFGQFPDTLG 298
wm1.2 LHSANQSLPH LNLTKLEKLD LSLNYFEHSL GSGWFWKAIS LKYLALGHNS LFGQFPDTLG 298
wm1.3 0
Consensus LDSANQSLPH LNLTKLEKLD LSNNYFEHSL GSGWFWKATS LKYL-LG-N- LFGQFPDTLG 300


wm1.12 105
wm1.11 105
wm1.10 NMTYLRVLDI SYNGNPDMMM TGNIKK.LCS LEILDLSGNR INGDIESLFV ESLPQCTRKN 357
wm1.1 NMTNLQVLDI SENWNPHMMM AGNLEN.LCG LEIIDLSYNY ING.DIAVLM ESLPQCTRKK 353
wm1.7 NMTNLQVLDI SVNKITDMMM TGNLEN.LCS LEIIDLSRNE IN.TDISVMM KSLPQCTWKK 356
wm1.2 NMTSLQVLDV SYNWNPDMMM IGKLLKNLCS LEIIDLDGNE ISG.EIEVLM ESWPQCTWKN 357
wm1.3 0
Consensus NMTNLQVLDI SYNWNPDMMM TGNLE--LCS LEIIDLSGNE ING-DISVLM ESLPQCT-K- 360
wm1.12 105
wm1.11 105
wm1.10 LQKLDLSYNN FTGTLPNIVS DFSKLSILSL SNNNLVGPIP AQLGNLTCLT SLDLFWNHLN 417
wm1.1 LQEMDLRYNN FTGTLPNLVS DFTRLRILSL SGNNLVGSIP PWLVNLTRLT TLSLFSNHLT 413
wm1.7 LQELDLGGNK FRGTLPNFIG DFTRLSVLWL DYNNLVGPIP PQLGNLTCLT SLDLGGNHLT 416
wm1.2 LQELDLSSNT FTGTLPNFLG DFTSLRTLSL SGNSLAGPIP PQLGNLTCLT SLDLSSNHFT 417
wm1.3 0
Consensus LQELDLSYNN FTGTLPNFV- DFTRL-ILSL SGNNLVGPIP PQLGNLTCLT SLDLFSNHLT 420
wm1.12 .DLR 108
wm1.11 .DLR 108
wm1.10 GSIPPELGAL TTLTSLDLSM NDLTGSIPAE LGNLRYLSEL CLSDNNITAP IPPEL 472
wm1.1 GSIPPWLGNL TCLTSLELSD NLLTGSIPAE FG 445
wm1.7 GSIPTELGAL TTLTYLDLGS NDLNGGVPAE LGNLRYLTAL YLSDNEIAGS IPPQLGNLRS 476
wm1.2 GSIRDELGNL RYLTALELQG NEITGSIPLQ LGNLTCLTSI DLGDNHITGS IPAEVG 473
wm1.3 0
Consensus GSIPPELG-L TTLTSL-LS- NDLTGSIPAE LGNLRYLT-L -LSDN-ITGS IPPELG---- 480
wm1.12 108
wm1.11 108
wm1.10 MNSTSLT HLDLSSNHLN 489
wm1.1 KLMYLT ILDLSSNHLN 461
wm1.7 LTALDLSDNE IAGSIPPQLG NLTGLTYLEL RNNHLTGSIP RELMHSTSLT ILDLPGNHLI 536
wm1.2 KLTYLT SLDLSSNHLN 489
wm1.3 0
Consensus ---MK-T-LT ILDLSSNHLN 540
wm1.12 N NSFTGVITEE HFANLTSLKK IDLSSNNFKI VLNSDWRAPF 149
wm1.11 N NSFTGVITEE HFANLTSLKK IDLSSNNFKI VLNSDWRAPF 149
wm1.10 GSVPTEIGSL NNLIYLYLSN NRFTGVITEE NFANLTSLKD IDLSFNNLKI VLNSDWRAPF 549
wm1.1 ESVPAEIGSL VNLIFLDLSN NSFTGVITEE HLANLTSLKQ IDLSLNNFKI ALNSDWRAPS 521
wm1.7 GSVPTEIGSL INLQFLDLSN NSFTGMITEE HLANLTSLQK IDLSSNNLKI VLNSDWRPPF 596
wm1.2 GSVPTEMGSL INLISLDLRN NSFTGVITGE HFANLTSLKQ IDLSYNNLKM VLNSDWRAPF 549
wm1.3 0
Consensus GSVPTEIGSL INLIFLDLSN NSFTGVITEE HFANLTSLKK IDLSSNN-KI VLNSDWRAPF 600


wm1.12 TLEFAWFASC QMGPLFPHGL QRLKTNALDI SNTTLKGEIP DWFWSAFSNA RYLDISNNQI 209
wm1.11 TLEFAWFASC QMGPLFPHGL QRLKTNALDI SNTTLKGEIP DWFWSAFSNA RYLDISNNQI 209
wm1.10 TLEFASFASC QMGPLFPPGL QRLKTNALDI SNTTLKGEIP DWFWSTFSNA TYLDISNNQI 609
wm1.1 TLESAWFASC QMGPLFPPWL QQLKITALDI STTSLKGEFP DWFWSAFSNV TYLDISNNQI 581
wm1.7 MLESASFGSC QMGPLFPPWL QQLKTTQLDI SHNGLKGEFP DWFWSTFSHA LYMDISNNQI 656
wm1.2 TLESASFGSC QMGPLFPPWL QQLKTTQLNI SSNGLKGEFP DWFWSAFSNV THLDISNNQX 609
wm1.3 0
Consensus TLE-A-FASC QMGPLFPP-L Q-LKT-ALDI SNTTLKGE-P DWFWSAFSNA TYLDISNNQI 660
wm1.12 SGSLPAHMHS MAFEELYLGS NHLTGPIPTL PTNITLLDIS NNTFLETIPS NLGAPRLEVL 269
wm1.11 SGSLPAHMHS MAFEELYLGS NHLTGPIPTL PTNITLLDIS NNTFLETIPS NLGAPRLEVL 269
wm1.10 SGSLPAHMHS MAFEKLRLGS NRLTGPIPTL PTNITLLDIS NNTFSETIPS NLGASRLEIL 669
wm1.1 SGNLPAHMDS MAFEKLYLRS NRLTGPIPTL PTNITLLDIS NNTFSETIPS NLVAPRLEIL 641
wm1.7 SGRLPAHLHG MAFEEVYLNS NQLTGPIPAL PKSIHLLDIS KNQFFGTIPS ILGAPRLQML 716
wm1.2 NGSLPAHMDS MAFEELHLSS NRLXGPIPTL PINITLLDIS NNTFSETIPS NLVAPGLKVL 669
wm1.3 ML 2
Consensus SGSLPAHMHS MAFEELYLGS NRLTGPIPTL PTNITLLDIS NNTFSETIPS NLGAPRLEVL 720
wm1.12 SMHSNQIGGY IPESICKLEQ LVYLDLSNNI LEGEVPKCFD THKIEHLILS NNSLSGKIPA 329
wm1.11 SMHSNQIGGY IPESICKLEQ LVYLDLSNNI LEGEVPKCFD THKIEHLILS NNSLSGKIPA 329
wm1.10 SMHSNQIGGY IPESICKLEQ LLYLDLSNNI LEGEVPHCFH FYKIEHLILS NNSLSGKIPA 729
wm1.1 CMHSNQIGGY IPESICKLEQ LIYLDLSNNI LEGEVPQCFD THNIENLILS NNSLSGKIPA 701
wm1.7 SMHSNQISGY IPESICKLEP LIYLDLSNNI LEGEIVKCFD IYSLEHLILG NNSLSGKIPA 776
wm1.2 CMQSNNXGGY IPESVCKLEQ LEYLDLSNNI LEGKIPQCPD IHNIKYLILS NNSLSGKIPA 729
wm1.3 LMYSNQIEGR IPESMCKLQS LLYLDLSNNI LESEIPQCFD VKQIQFLLLS NNSLSGKFPA 62
Consensus SMHSNQIGGY IPESICKLEQ L-YLDLSNNI LEGEVP-CFD THKIEHLILS NNSLSGKIPA 780
wm1.12 FLQNNTSLEF LDLSWNKFSG RLPTWIGNLV YLRFLVLSHN EFSDNIPVNI TKLGHLQYLD 389
wm1.11 FLQNNTSLEF LDLSWNKFSG RLPTWIGNLV YLRFLVLSHN EFSDNIPVNI TKLGHLQYLD 389
wm1.10 FLQNNTGLQF LDVSWNRFSG RLPTWIGNLV NLRFLVLSHN IFSDNIPVDI TKLGHLQYLD 789
wm1.1 FLQNNTSLEF LDLSWNKFSG RLPTWIGNLV YLRFLVLSHN EFSDNIPVNI TKLGHLQYLD 761
wm1.7 SLRNNACLKF LDLSWNKFSG GLPTWIGTLV HLRFLILSHN KFSDNIPVDI TKLGWLQYLD 836
wm1.2 FLQNNTNLKF LDLSWNNFSG RLPTWIGKLA NLXFLILSHN XFSDSIPVXV TKLGHLQYLD 789
wm1.3 FLQNSTRLKF LDLAWNKFSG RLPTWIGELW KLRFLLLSHN ALSGTIPVEI TNLGWLQRLD 122
Consensus FLQNNTSL-F LDLSWNKFSG RLPTWIGNLV YLRFLVLSHN EFSDNIPVNI TKLGHLQYLD 840
wm1.12 LSHNNFSGAI PWHLPNLTFM TTFEADSMGG .DMVVVEVDS ..MGEEF.EA DSLGQILSVN 445
wm1.11 LSHNNFSGAI PWHLPNLTFM TTFEADSMGG .DMVVVEVDS ..MGEEF.EA DSLGQILSVN 445
wm1.10 LSRNNFSGGI PWHMSNLTFM STLQSMYM.. ..VEVTEYDT TRLGPIFIEA DRLGQILSVN 845
wm1.1 LSHNNFSGAI PRHLSNLTFM TTLQEESR.. .YMVEVEVDS MGGTTEF.EA DSLGQILSVN 817
wm1.7 LSSNNFSGAI PWHLSSLTFM STLQEESMG. ...LVGDVRG ...SEIV..P DRLGQILSVN 887
wm1.2 LSDNRFFGAI PCHLSNLTFM RTLQEDIDMD .GPILYVFKE YATGIAP..Q E.LGQTLXVN 845
wm1.3 LSVNNFSGPI PLHLSKLTFM KNLQEQFMPR SDNRGTHLNN IQLVSPTFGA GHLAEILSVI 182
Consensus LSHNNFSGAI PWHLSNLTFM TTLQEDSMGG -DMVVVEVDS ---GEEF-EA DSLGQILSVN 900


wm1.12 TKGQQLTYHK TLEYFVSIDL SCNSLTGKIP TDITSLAALM NLNLSSNQLS GQIPNMIGAV 505
wm1.11 TKGQQLTYHK TLEYFVSIDL SCNSLTGKIP TDITSLAALM NLNLSSNQLS GQIPNMIGAV 505
wm1.10 TKGQQLIYHG TLAYFVSIDL SCNSLTGEIP TDITSLAALM NLNLSSNQLS GQIPSMIGAM 905
wm1.1 TKGQQLIYHR TLAYFVSIDL SCNSLTGKIP TDITSLAALM NLNLSSNQLS GQIPNMIGAM 877
wm1.7 TKGQQLTYHR TLAYFVSIDL SCNSLTGEIP TDITSLAALM NLNLSSNQLS GQIPSMIGAM 947
wm1.2 TKGQHLIYHM TLAYFVGIDL SXNSLTGEIP TDITSLDALV NLNLSSNQLS GBIPNMIGAM 905
wm1.3 TKGQKLIYGS TIAYFVNIDL SSNSLTGEIP TDITSLVALI NLNLSSNQLS GQIPNMIGTV 242
Consensus TKGQQLIYH- TLAYFVSIDL SCNSLTGEIP TDITSLAALM NLNLSSNQLS GQIPNMIGAM 960
wm1.12 QSLVSLDLSQ NKLSGEIPSS LSSLTSLSYL NLSYNSLSGI IPSGPQLDIL NLDNQSLIYI 565
wm1.11 QSLVSLDLSQ NKLSGEIPSS LSSLTSLSYL NLSYNSLSGI IPSGPQLDIL NLDNQSLIYI 565
wm1.10 QSLVSLDLSQ NKLSGEIPSS LSNLTSLSYM NLSCNSLSGR IPSGPQLDIL NLDNQSLIYI 965
wm1.1 QSLESLDLSQ NKLYGEIPSS LTNLTSLSYL DLSYNSLSGR IPSGPQLDTL NMDNQTLMYI 937
wm1.7 QSLVSLDLSQ NKLSGEIPSS LSNLTSLSYM NLSCNSLSGR IPSGRQLDTL NMDNPSLMYI 1007
wm1.2 QSLESLDLSQ NKLYGEIPSS LTNLTSLSYL DLSYNSLSGR IPSGPQLDTL SAENQSLMYI 965
wm1.3 QSLVSLDLSQ NKLYGEIPLS LSSLTSLSYL NLSYNSLSGM IPSGPQLDIL NLDNQSLMYI 302
Consensus QSLVSLDLSQ NKLSGEIPSS LSNLTSLSYL NLSYNSLSGR IPSGPQLDIL NLDNQSLMYI 1020
wm1.12 SNSGLCGPPV HKNCSGNDPF IHGDLESSKE EFDPLTFHFG LVLGFVVGLW MVFCALLFKK 625
wm1.11 SNSGLCGPPV HKNCSGNDPF IHGDLESSKE EFDPLTFHFG LVLGFVVGLW MVFCALLFKK 625
wm1.10 GNTGLCGPPV HKNCSGNDPY IHSDLESSKE EFDPLTFYFG LVLGFVVGLW MVFCALLFKK 1025
wm1.1 GNNGLCGPPV HKNCSGNDAY IHGDLESSKE EFDPLTFYFG LVLGFVVGLW MVFCALLFKK 997
wm1.7 GNNGLCGPPV HKNCSGNDPF IHGDLRSSNQ EVDPLTFYFG LVLGFVVGLW MVFCALLFKK 1067
wm1.2 GNSGLCGPPV HKNCSGNEPS IHDDLKSSKK EFDPLNFYFG LVLGFVVGLW MVFCXLLFKR 1025
wm1.3 GNSGLCGPPV HKNCPGNDSS IHGDLKSSNE EFDPLTFYFG LVLEFVVGLW MVFCALLFKK 362
Consensus GNSGLCGPPV HKNCSGNDPF IHGDLESSKE EFDPLTFYFG LVLGFVVGLW MVFCALLFKK 1080
wm1.12 TWRIAYFRLF DKVYDHVYVF VVVKWAGFAK KTDEE660
wm1.11 TWRIAYFRLF DKVYDHVYVF VVVKWAGFAK KTDEE660
wm1.10 TWRIAYFRFF DKVYDQVYVF VVVKWASFAK NTPAE1060
wm1.1 TWRIAYFRLF DKVYDQVYVF VVVKWASFAK KTDEE1032
wm1.7 TWRIAYFRLF DKVYDQVYVF VVVKWASFAK KTDEE1102
wm1.2 TWRIAYFRLF DBVYDQVYVF VVVKWXSFTK NTDAE1060
wm1.3 TWRIAYFRLF DKVYDQVYVF AAVKWASFAK NTDAE397
Consensus TWRIAYFRLF DKVYDQVYVF VVVKWASFAK KTDE 1115

B

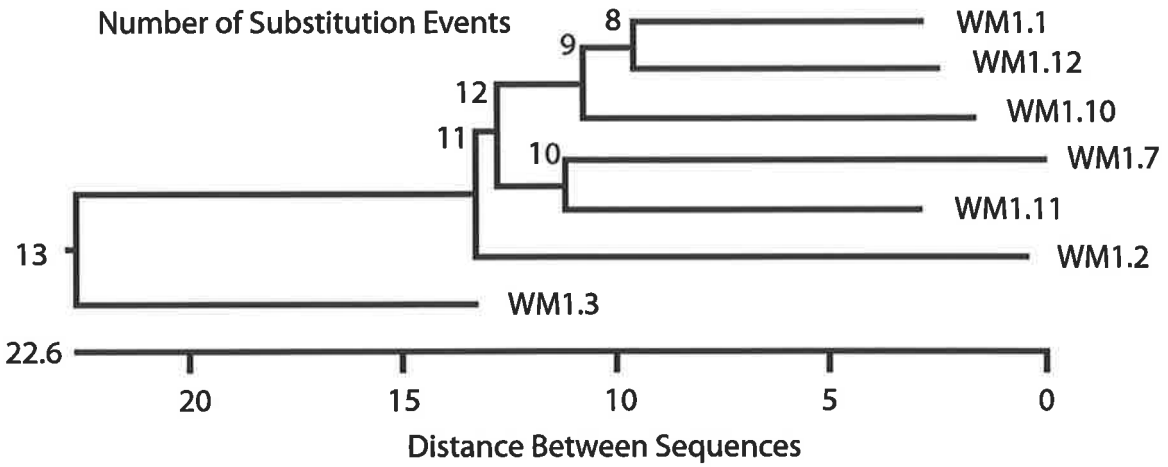

Number of Substitution Events
8
9
12
11
10
13
22.6
WM1.1
WM1.12
WM1.10
WM1.7
WM1.11
WM1.2
WM1.3
20
15
10
5
0
Distance Between Sequences

Table 7: Polypeptide sequence similarity and divergence between the predicted primary peptides of seven members of the *WM1* gene family.

Percent Divergence	Percent Similarity							
	WM1.1	WM1.2	WM1.3	WM1.7	WM1.10	WM1.11	WM1.12	
		78.7	74.6	78.8	81.9	81.0	85.9	WM1.1
	22.1		71.0	73.7	75.6	78.1	75.8	WM1.2
	28.5	31.5		72.3	73.6	70.5	73.8	WM1.3
	23.5	27.9	30.1		77.5	80.6	81.4	WM1.7
	18.5	24.6	29.6	24.9		80.6	84.4	WM1.10
	19.5	20.7	31.6	19.5	20.8		75.6	WM1.11
	13.8	25.7	30.4	20.6	15.9	21.3		WM1.12

WM1.7 and WM1.11 predicted primary peptides are also highly similar as they can only be distinguished based on 10 substitution events across 516 amino acids (**Figure 18**). **Table 8** outlines predicted peptide properties for each of the predicted WM1 primary peptides whilst **Figure 19** displays the structural features and similarity between peptides.

3.3.6.2.1 Transmembrane domains

The TMpred program (Hofmann & Stoffel, 1993) was used to make predictions of membrane-spanning regions and their orientation. The algorithm is based on the statistical analysis of TMbase, a database of naturally occuring transmembrane proteins. **Figure 19** illustrates the prediction of transmembranes within the predicted WM1 primary peptides. All predicted peptides of the *WM1* gene cluster seem to encode a highly conserved C-terminal putative transmembrane domain.

3.3.6.2.2 Leucine zippers

Several predicted WM1 proteins (WM1.1, 1.2, and WM1.10-WM1.12) have 100% identity with a zipper-like domain (**Figure 20**) as predicted by searches against the PROSITE database (Hofmann *et al.*, 1999). WM1.1, WM1.2, WM1.10 and WM1.12 each have a leucine zipper within the tandemly arrayed LRRs specifically at a position between 414 - 436 amino acids from the C-terminus. On the other hand WM1.11 has a leucine zipper at the C-terminus of the predicted peptide within the predicted C-terminal transmembrane domain.

Table 8: Properties of predicted WM1 primary peptides.

WM1 predicted primary peptide	Length (aa)	Molecular Weight (kDa)	Isoelectric Point	Charge at pH 7
WM1.1	1032	115	5.47	-23.97
WM1.2	1060	117	5.51	-22.98
WM1.3	397	44	5.89	-3.74
WM1.7	1102	121	5.84	-16.84
WM1.10	1060	118	5.82	-17.75
WM1.11	516	57	5.62	-10.20
WM1.12	660	73	5.65	-14.54

Figure 19: Diagram of the structural relationship between the predicted primary polypeptide sequences for seven members of the *WM1* gene family. Whole predicted polypeptide sequences are represented for WM1.1, WM1.2, WM1.3, WM1.7, WM1.10, WM1.11 and WM1.12. For each predicted polypeptide sequence the legend identifies structural motifs and regions of high homology.

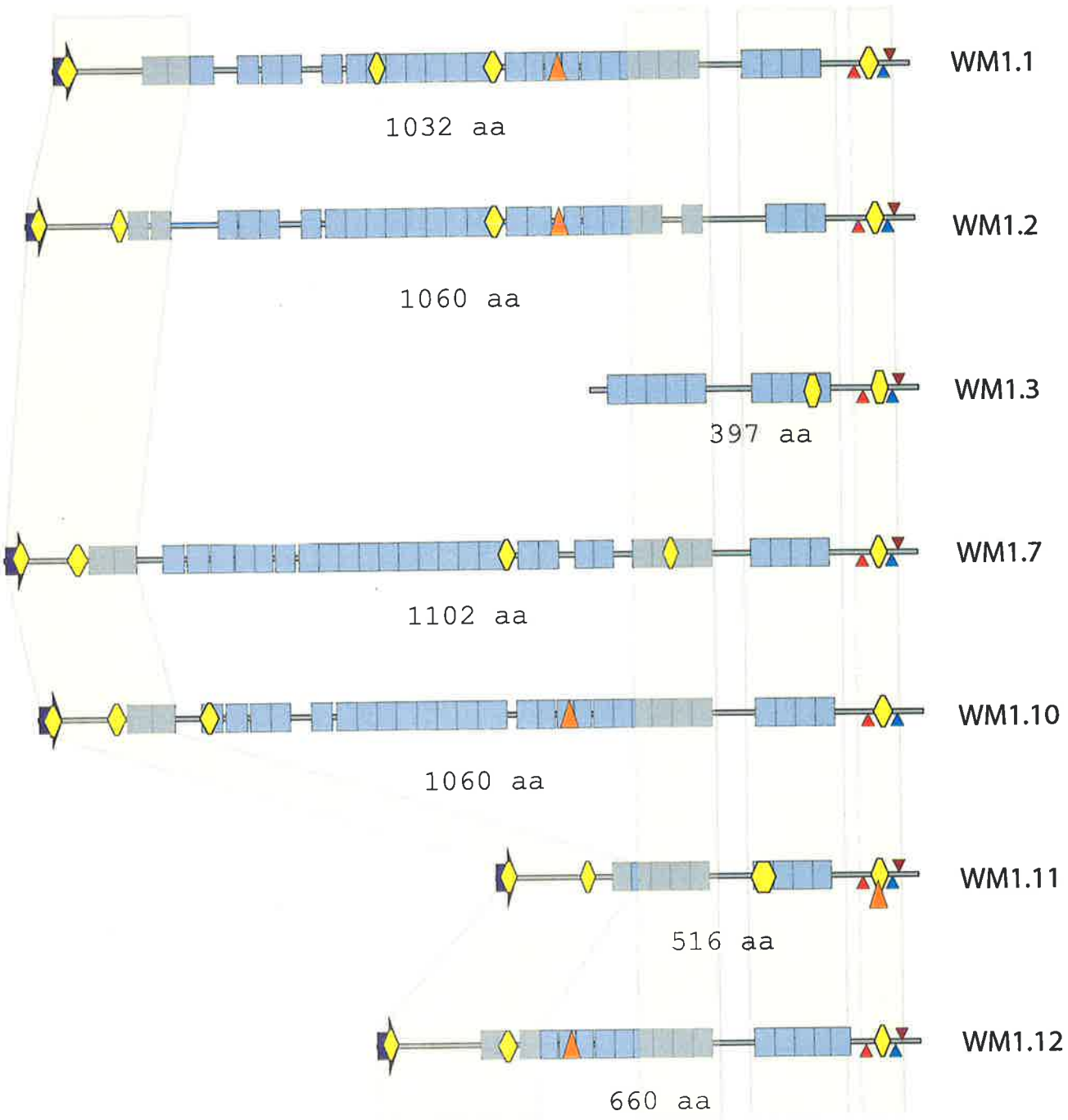
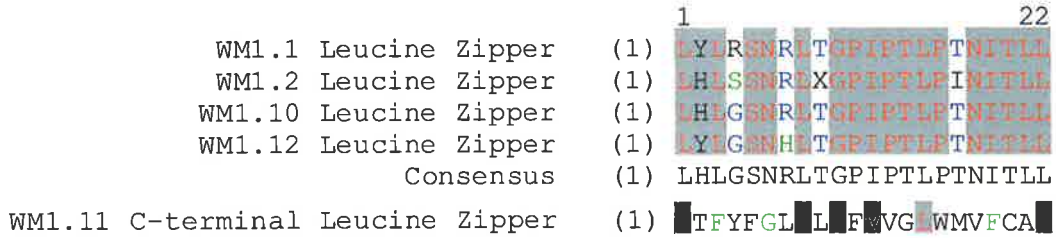

WM1.1
1032 aa
WM1.2
1060 aa
WM1.3
397 aa
WM1.7
1102 aa
WM1.10
1060 aa
WM1.11
516 aa
WM1.12
660 aa
Leucine Zipper
Leucine Rich Repeats
Signal Peptide
Transmembrane Domain
Regions of High Amino Acid Conservation
Negative Charge Cluster
Positive Charge Cluster
Endocytosis Signal

Figure 20: Structure of leucine rich repeat and leucine zipper-like motifs from within the predicted primary polypeptide sequences for members of the *WM1* gene family.

A – Leucine zipper consensus generated from the alignment of leucine zipper-like motifs from WM1.1, WM1.2, WM1.10 and WM1.12, also comparing the C-terminal leucine zipper-like motif from WM1.11.

B – *WM1* gene family leucine rich repeat consensus generated from the alignment of leucine rich repeat consensus sequences of predicted primary polypeptides WM1.1, WM1.2, WM1.3, WM1.7, WM1.10, WM1.11 and WM1.12.

A


1
22
WM1.1 Leucine Zipper (1) LYLRGNRLTGPIPTLPTNITLL
WM1.2 Leucine Zipper (1) LHLSGNRLXGPIPTLPINITLL
WM1.10 Leucine Zipper (1) LHLGGNRLTGPIPTLPTNITLL
WM1.12 Leucine Zipper (1) LYLGGNHLTGPIPTLPTNITLG
Consensus (1) LHLGSNRLTGPIPTLPTNITLL
WM1.11 C-terminal Leucine Zipper (1) TFYFGLLLFVGLWMVFCA

B


1
31
WM1.1 consensus (1) XLXXLDLSXNXLSGX------IPXX-LXXL-
WM1.2 consensus (1) XLXXLDLSXNXLXGX------IPXXXLXXL-
WM1.3 consensus (1) XLXFLDLSXNXLSGX------IPXX-XXL-
WM1.7 consensus (1) XLXXLDLSXNXLSGXXXXXXXIPXX-LXXL-
WM1.10 consensus (1) XLXXLDLSXNXLSGX------IPXXXLXXL-
WM1.11 consensus (1) XLXFLDLSXNXLSGX------IPXX-XXL-
WM1.12 consensus (1) XLXXLDLSXNXLSGX------IPSX-----
Consensus (1) XLXXLDLSXNXLSGX IPXX LXXL

3.3.6.2.3 Leucine rich repeats

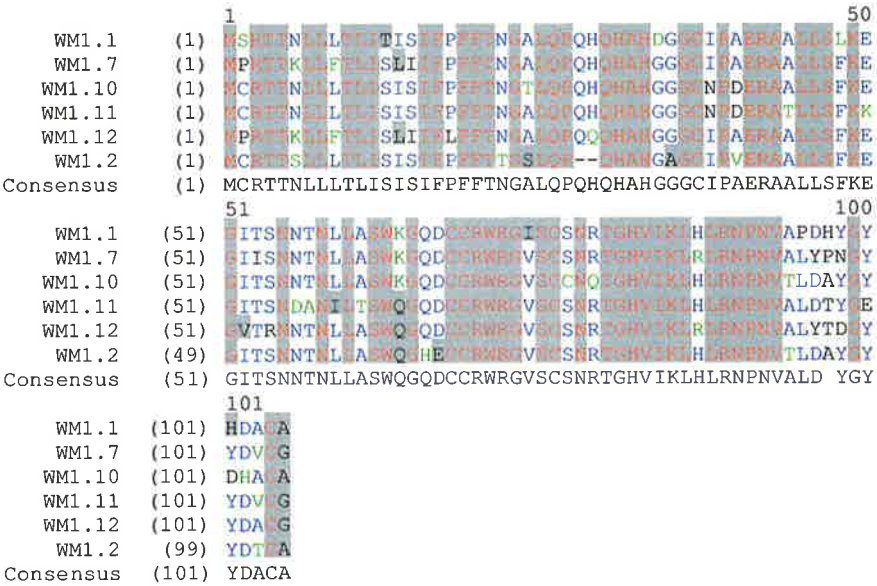
All predicted WM1 proteins have varying lengths in stretches of tandemly arrayed leucine rich repeat signatures (LRRs) as predicted by INTERPRO (Apweiler *et al.*, 2001) and BLOCKS database (Henikoff *et al.*, 2000; Henikoff *et al.*, 1999) searches (**Figure 19**). **Figure 19** displays that the number of LRRs varies between 9 and 30 for the predicted peptides of the *WM1* gene family cluster. Conservation of LRRs between the different members of the gene family is primarily within the N- and C-terminal blocks of LRRs (refer to **Figure 21**). The C-terminus of all the predicted WM1 primary peptides share 68.2% identity and 100% similarity (**Figure 21**–A), whilst the N-terminus shares 60% identity and 99% similarity (the truncated WM1.3 was excluded from the alignment) (**Figure 21**-B). This shows that there is high conservation in sequence similarity between the predicted WM1 primary peptides and most variation relates to the length of the tandem array of LRRs. However, variation is observed in the amino acid composition between individual LRRs within each of the predicted WM1 primary peptides. Typically LRRs within a predicted peptide share between 31.7% and 58.3% similarity and share between 0% and 12.5% identity. Similarity and identity between LRRs within a predicted peptide seems to decrease with the increasing length of the LRR tandem arrays. Similarity in the amino acid divergence is comparible with the Cf-2 and Cf-9 primary peptides (Dixon *et al.*, 1996; Jones *et al.*, 1994). A barley homologue of Cf-2 and Cf-9 shows similar patterns of LRR divergence within the primary peptide (data not shown).

Figure 21: Highly conserved N-terminal and C-terminal blocks of tandemly arrayed leucine rich repeats.

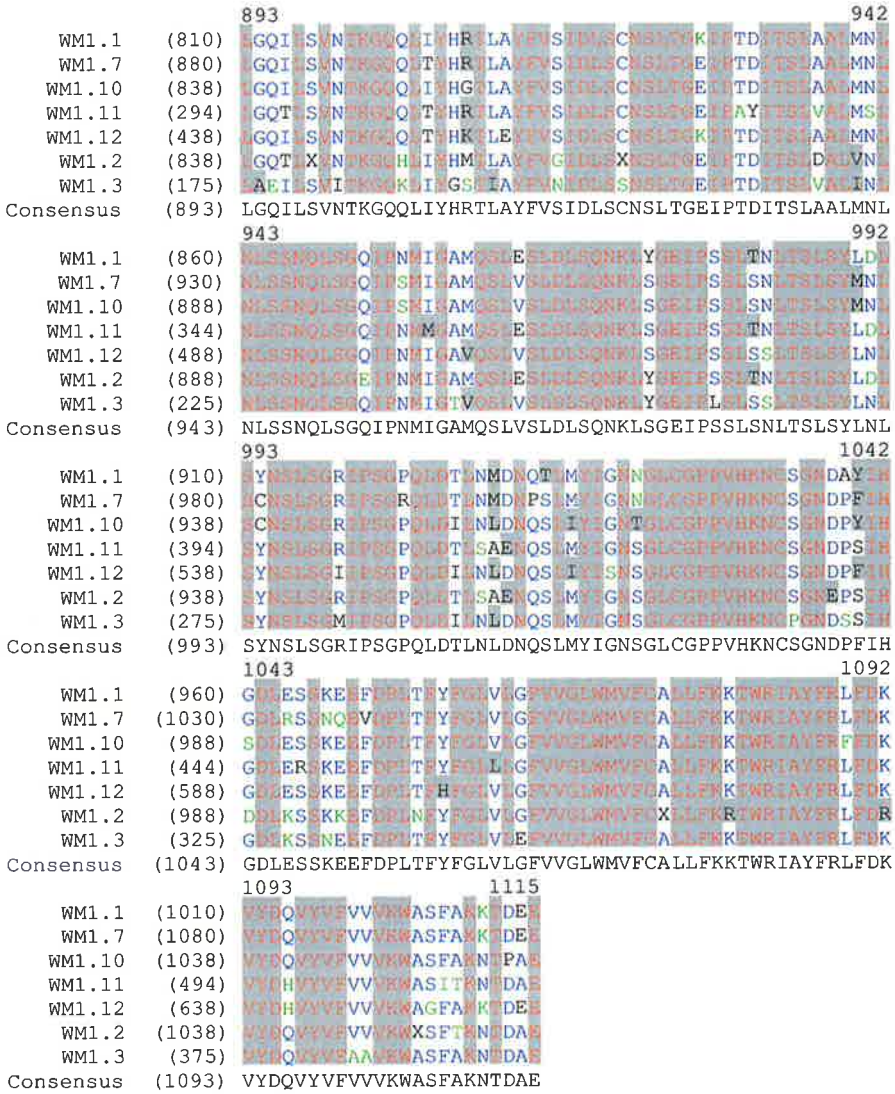
A – Alignment of the N-terminal region of predicted polypetides WM1.1, WM1.1, WM1.7, WM1.10, WM1.11 and WM1.12.

B – Alignment of the C-terminal region of predicted polypetides WM1.1, WM1.1, WM1.3, WM1.7, WM1.10, WM1.11 and WM1.12.

A


1 50
WM1.1 (1) MSHTTNLLLTTISISIPFFTNGALQPQHQHAHGGGCIPAERAALLSFKE
WM1.7 (1) MPKTTKLLFTLSLITLFFTNGALQPQQQHAHGGGCIPAERAALLSFKE
WM1.10 (1) MCRTTNLLLTSISISIPFFTNGTLQPQHQHAHGGGSNPDERAALLSFKE
WM1.11 (1) MCRTTNLLLTSISISIPFFTNGALQPQHQHAHGGGSNPDERATLLSFKK
WM1.12 (1) MPKTTKLLFTLSLITLFFTNGALQPQQQHAHGGGCIPAERAALLSFKE
WM1.2 (1) MCRTSSLLLTSISISIPFFTTSSLQR--QHAHGGACIPVERAALLSFKE
Consensus (1) MCRTTNLLLTLISISIPFFTNGALQPQHQHAHGGGCIPAERAALLSFKE
51 100
WM1.1 (51) GITSNNTNLLASWQGQDCCRWRGVSCSNRTGHVIKLHLRNPNVALDTYGY
WM1.7 (51) GIISNNTNLLASWKGQDCCRWRGVSCSNRTGRVTKLRLRNPNVALYPNGY
WM1.10 (51) GITSNNTNLLASWKGQDCCRWRGVSCCNQTGHVIKLHLRNPNTLDAYGY
WM1.11 (51) GITSNDANILTSWQGQDCCRWRGVSCSNRTGHVIKLHLSNPNVALDTYGE
WM1.12 (51) GVTRNNTNLLASWQGQDCCRWRGVSCSNRTGHVIKLRLRNPNVALYTDGY
WM1.2 (49) GITSNNTNLLASWQGHECCRWRGVSCSNRTGHVIKLHLRNPNVTLDAYGY
Consensus (51) GITSNNTNLLASWQGQDCCRWRGVSCSNRTGHVIKLHLRNPNVALD YGY
101
WM1.1 (101) HDACA
WM1.7 (101) YDVCG
WM1.10 (101) DHACA
WM1.11 (101) YDVCG
WM1.12 (101) YDACG
WM1.2 (99) YDTCA
Consensus (101) YDACA

B


893 942
WM1.1 (810) LGQILSVNTKGQQLIYHRTLAYFVSIDLSCNSLTGEIPTDITSLAALMNL
WM1.7 (880) LGQILSVNTKGQQLTYHRTLAYFVSIDLSCNSLTGEVPTDITSLAALMNL
WM1.10 (838) LGQILSVNTKGQQLIYHGTLAYFVSIDLSCNSLTGEIPTDITSLAALMNL
WM1.11 (294) LGQTLSVNTKGQQLTYHRTLAYFVSIDLSCNSLTGEIPAYITSLVALMSL
WM1.12 (438) LGQILSVNTKGQQLTYHKTLEYFVSIDLSCNSLTGKIPTDITSLAALMNL
WM1.2 (838) LGQTLXVNTKGQHLIYHMTLAYFVGIDLSXNSLTGEIPTDITSLDALVNL
WM1.3 (175) LAEILSVITKGQKLIYGSTIAYFVNIDLSSNSLTGEIPTDITSLVALINL
Consensus (893) LGQILSVNTKGQQLIYHRTLAYFVSIDLSCNSLTGEIPTDITSLAALMNL
943 992
WM1.1 (860) NLSSNQLSGQIPNMIGAMQSLESLDLSQNKLYGEIPSSLTNLTSLSYLDL
WM1.7 (930) NLSSNQLSGQITSMIGAMQSLVSLDLSQNKLSGEIPSSLSNLTSLSYMNL
WM1.10 (888) NLSSNQLSGQIPSMIGAMQSLVSLDLSQNKLSGEIPSSLSNLTSLSYMNL
WM1.11 (344) NLSSNQLSGQIPNMMGAMQSLESLDLSQNKLSGEIPSSLTNLTSLSYLDL
WM1.12 (488) NLSSNQLSGQITNMIGAVQSLVSLDLSQNKLSGEIPSSLSSLTSLSYLNL
WM1.2 (888) NLSSNQLSGEIPNMIGAMQSLESLDLSQNKLYGEIPSSLTNLTSLSYLDL
WM1.3 (225) NLSSNQLSGQIPNMIGTVQSLVSLDLSQNKLYGEIPLSLSSLTSLSYLNL
Consensus (943) NLSSNQLSGQIPNMIGAMQSLVSLDLSQNKLSGEIPSSLSNLTSLSYLNL
993 1042
WM1.1 (910) SYNSLSGRIPSGPQLDTLNMDNQTLMYIGNSGLCGPPVHKNCSGNDAYIH
WM1.7 (980) NCNSLSGRIPSGRQLDTLNMDNPSLMYIGNNGLCGPPVHKNCSGNDPFIH
WM1.10 (938) SCNSLSGRIPSGPQLDILNLDNQSLIYIGNTGLCGPPVHKNCSGNDPYIH
WM1.11 (394) SYNSLSGRIPSGPQLDTLSAENQSLMYIGNSGLCGPPVHKNCSGNDPSIH
WM1.12 (538) SYNSLSGIIPSGPQLDILNLDNQSLIYLSNSGLCGPPVHKNCSGNDPFIH
WM1.2 (938) SYNSLSGRIPSGPQLDTLSAENQSLMYIGNSGLCGPPVHKNCSGNEPSIH
WM1.3 (275) SYNSLSGMIPSGPQLDILNLDNQSLMYIGNSGLCGPPVHKNCPGNDSSIH
Consensus (993) SYNSLSGRIPSGPQLDTLNLDNQSLMYIGNSGLCGPPVHKNCSGNDPFIH
1043 1092
WM1.1 (960) GDLESSKEEFDPLTFYFGLVLGFVVGLWMVFCALLFKKTWRIAYFRLFDK
WM1.7 (1030) GDLRSSNQEVDPLTFYFSIVIGFVVGLWMVFCALLFKKTWRIAYFRLFDK
WM1.10 (988) SDLESSKEEFDPLTFYFGLVIGFVVGLWMVFCALLFKKTWRIAYFRFFDK
WM1.11 (444) GDLERSKEEFDPLTFYFGLLLGFVVGLWMVFCALLFKKTWRIAYFRLFDK
WM1.12 (588) GDLESSKEEFDPLTFHFGLVLGFVVGLWMVFCALLFKKTWRIAYFRLFDK
WM1.2 (988) DDLKSSKKEFDPLNFYFGLVLGFVVGLWMVFCXLLFKRTWRIAYFRLFDR
WM1.3 (325) GDLKSSNEEFDPLTFYFGLVLGFVVGLWMVFCALLFKKFWRIAYFRLFDK
Consensus (1043) GDLESSKEEFDPLTFYFGLVLGFVVGLWMVFCALLFKKTWRIAYFRLFDK
1093 1115
WM1.1 (1010) VYDQVYVFVVVKWASFAKKTDEL
WM1.7 (1080) VYDQVYVFVVVKWASFAKKTDEL
WM1.10 (1038) VYDQVYVFVVVKWASFAKNTPAL
WM1.11 (494) VYDHVYVFVVVKWASITKNTDAL
WM1.12 (638) VYDHVYVFVVVKWAGFAKKTDEL
WM1.2 (1038) VYDQVYVFVVVKWXSFTKNTDAL
WM1.3 (375) VYDQVYVSAAVKWASFAKNTDAL
Consensus (1093) VYDQVYVFVVVKWASFAKNTDAE

3.3.6.2.4 Signal Peptide

All WM1 predicted primary peptides, except WM1.3, possess a signal peptide. This however, does not preclude cleavage and therefore predictions from both the SignalP V2.0 (Nielsen *et al.*, 1997) and PSORT (Nakai & Kanehisa, 1992; von Heijne, 1986) (**Table 9**) programs were assessed. Results of these predictions suggest that neither WM1.1 or WM1.10 have a cleavable signal peptide, whilst WM1.2, WM1.7, WM1.11, and WM1.12 signal peptides are all likely to be cleaved.

3.3.6.2.5 Predicted subcellular localisation

PSORT program (Nakai & Kanehisa, 1992) was used to predict potential protein localisation sites in eukaryotic plant cells (summarised in **Table 9**). All predicted WM1 primary peptides revealed that the *WM1* family are all localised to plasma membranes and can be classified into three classes. The first classification is the type Ia membrane proteins which include WM1.2, WM1.7, WM1.11 and WM1.12. Secondly WM1.3 is categorised as a type Ib (N–terminus extracellular, cytosolic C-terminus) membrane protein. Finally there are the type II (N-terminus cytosolic, extracellular C-terminus) membrane proteins, which include WM1.1, and WM1.10. It should be noted that membrane topology prediction is highly dependant on the prediction of signal sequence cleavage. Manual analysis of cytoplasmic domains for all WM1 predicted peptides show that all except WM1.10 possess a YFRL motif. This matches to the mammalian consensus for the typrosine YXXϕ endosomal/lysosomal sorting signal sequences, where ϕ is an amino acid with a hydrophobic side chain, that stimulates receptor mediated endocytosis and degradation of cell-surface receptors (Letourneur & Klausner, 1992; Marks *et al.*, 1996).

Table 9: Prediction of subcellular localisation, membrane topology and signal sequence for WM1 predicted polypeptides. Predictions and assigned probabilities are based on the PSORT program (Nakai & Kanehisa, 1992) for each predicted primary peptides of WM1.1, WM1.2, WM1.3, WM1.7, WM1.10, WM1.11 and WM1.12.

Predicted Polypeptide	Signal Sequence	Membrane Topology	Subcellular localisation	Probability
WM1.1	Uncleavable N-terminal signal sequence	Type II membrane protein	Plasma Membrane	69.8 %
WM1.2	Cleavable N-terminal signal sequence	Type Ia membrane protein	Plasma Membrane	46.0 %
WM1.3	No N-terminal signal sequence	Type Ib membrane protein	Plasma Membrane	70.0 %
WM1.7	Cleavable N-terminal signal sequence	Type Ia membrane protein	Plasma Membrane	46.0 %
WM1.10	Uncleavable N-terminal signal sequence	Type II membrane protein	Endoplasmic Reticulum	64.0 %
WM1.11	Cleavable N-terminal signal sequence	Type Ia membrane protein	Plasma Membrane	46.0 %
WM1.12	Cleavable N-terminal signal sequence	Type Ia membrane protein	Plasma Membrane	46.0 %

3.3.6.2.6 Structural conservation

Figure 18 and **Figure 19** highlight the areas of conservation in amino acid sequence within WM1 putative peptides. Conservation in amino acid sequence extends to the N-terminal region of the putative signal peptide and to the blocks of tandemly arrayed LRRs at the C-terminus. Differences between the lengths of the tandemly arrayed LRRs lies mainly on the RiceGAAS prediction (**Figure 16**) of the coding sequence and the presence of several stop codons in what seems to be an original single open reading frame.

3.3.6.3 Genes homologous to WM1 *genes*

Each predicted WM1 primary peptide sequence was screened using the BLAST 2.0 blastp program (Altschul *et al.*, 1997) across the non-redundant Genbank database. Homologous proteins at the primary peptide level included the Cf-2/Cf-5-like protein from barley (*H. vulgare*), the disease resistance proteins Cf-2.1 and Cf-9 from current tomato (*Lycopersicon pimpinellifolium*), a putative disease resistance protein from *A. thaliana*, an LRR-like protein from soybean (*Glycine max*), verticillium wilt disease resistance protein from tomato Ve-2 (*Lycopersicon esculentum*), three rice (*Oryza sativa*) Cf-2-like predicted proteins and the HcrVf1 protein from apple (**Figure 22**).

Primary polypeptide comparisons show that the WM1.1 protein has between 28.8% and 52.7% similarity and 100% sequence divergence with most other type Ia plasma membrane anchored LRR-like proteins (**Table 10**). The only exception is the comparison between the barley Cf-2/Cf-9-like peptide and wheat WM1.1 peptide, which seems to have a sequence divergence of only 75.2%. The phylogeny assumed evolutionary relatedness of wheat WM1.1 with other LRR proteins and aligns peptide sequences so as to minimise evolutionary change.

Figure 22: Diagram of the structural relationship between the predicted primary polypeptide of *T. tauschii* WM1.1 and those polypeptides found to be most homologous as determined by the BLAST 2.0 blastp program (Altschul *et al.*, 1997). WM1-like peptides are from barley (*H. vulgare*), tomato (*Lycopersicon pimpinellifolium* and *esculentum*), *A. thaliana*, soybean (*Glycine max*), rice (*Oryza sativa*) and apple (*Malus floribunda*) (refer to **Figure 23**). For each predicted polypeptide sequence the legend displays structural motifs and regions of high homology. Genbank accession numbers are included in brackets.

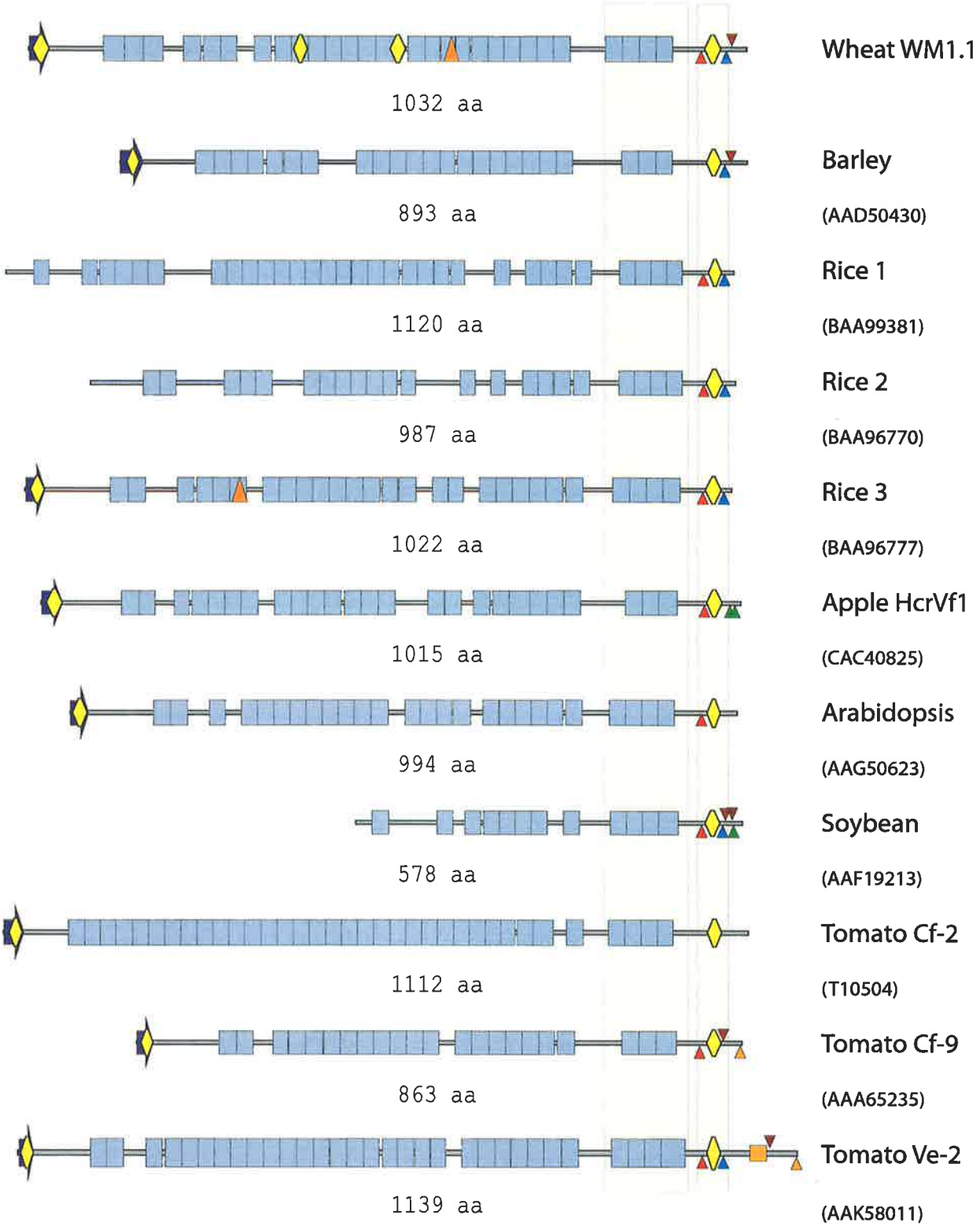

Wheat WM1.1
1032 aa
Barley
(AAD50430)
893 aa
Rice 1
(BAA99381)
1120 aa
Rice 2
(BAA96770)
987 aa
Rice 3
(BAA96777)
1022 aa
Apple HcrVf1
(CAC40825)
1015 aa
Arabidopsis
(AAG50623)
994 aa
Soybean
(AAF19213)
578 aa
Tomato Cf-2
(T10504)
1112 aa
Tomato Cf-9
(AAA65235)
863 aa
Tomato Ve-2
(AAK58011)
1139 aa
Leucine Zipper
Leucine Rich Repeats
Signal Peptide
Transmembrane Domain
PEST-like sequence
Regions of High Amino Acid Conservation
Negative Charge Cluster
Positive Charge Cluster
Tyrosine Endocytosis Signal
Endoplasmic Retential Signal
Dileucine Endocytosis Signal

Table 10: Polypeptide sequence similarity and divergence between the predicted primary polypeptide of *T. tauschii* WM1.1 and those polypeptides found to be most homologous as determined by the BLAST 2.0 blastp program (Altschul *et al.,* 1997). *WM1*-like peptides are from barley (*H. vulgare*), tomato (*Lycopersicon pimpinellifolium* and *esculentum*), *A. thaliana*, soybean (*Glycine max*), rice (*Oryza sativa*) and apple (*Malus floribunda*) (refer to **Figure 23**).

Percent Divergence / Percent Similarity											
Tomato Cf-9	*Arabidopsis*	Barley	Tomato Ve-2	Rice 1	Apple HcrVf1	Tomato Cf-2	Wheat WM1.1	Soybean	Rice 2	Rice 3	
	43.3	41.4	34.3	29.4	33.4	36.4	35.0	43.5	38.2	36.2	Tomato Cf-9
100		32.0	34.5	29.1	31.4	30.2	32.7	31.6	37.4	36.4	*Arabidopsis*
100	100		31.2	26.5	40.5	29.2	52.7	44.4	35.4	33.6	Barley
100	100	100		36.2	31.4	34.2	32.9	27.0	39.6	38.1	Tomato Ve-2
100	100	100	100		26.1	30.8	28.8	21.1	37.0	36.3	Rice 1
100	100	100	100	100		29.9	40.3	32.7	34.3	33.8	Apple HcrVF1
100	100	100	100	100	100		29.7	25.4	35.1	32.0	Tomato Cf-2
100	100	75.2	100	100	100	100		32.9	36.8	35.3	Wheat WM1.1
100	100	100	100	100	100	100	100		32.9	28.8	Soybean
100	100	100	100	100	100	100	100	100		60.7	Rice 2
100	100	100	100	100	100	100	100	100	69.2		Rice 3

Figure 23 shows that tomato Cf-9 is most closely related to the common ancestral type Ia plasma membrane anchored LRR-like protein. Based on the number of amino acid substitution events distinguishing the different LRR-like peptides it seems as though LRR-like proteins within the monocotyledonous species (barley, wheat and rice) are the most closely related.

3.3.7 Northern analysis of *WM1* genes

Northern analysis of total RNA and polyA isolated mRNA on a range of wheat tissues was unable to detect any signal using the *WM1.1* ORF PCR fragment as a probe (data not shown), even though hybridisation and washing conditions (refer to 2.2.14) for the Northern analysis was standardised to replicate the conditions as seen for Southern analysis.

3.3.8 RT-PCR of *WM1* genes

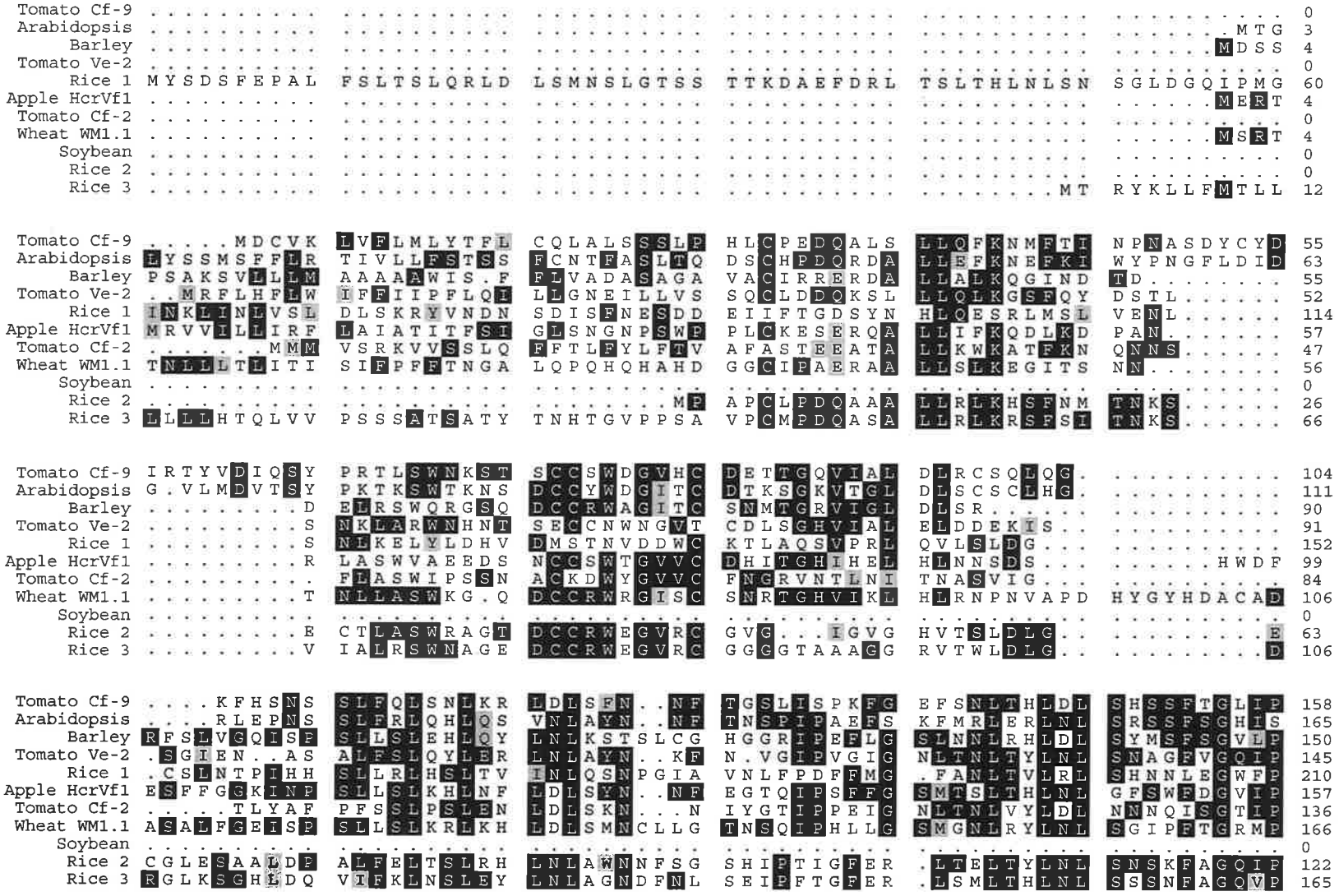
RT-PCR of *WM1.1* - *WM1.3*, *WM1.10*, and *WM1.11* was performed on 5 μg of total RNA extracted from; root tissue, young leaf tissue (3 weeks old), and spikes undergoing early meiosis (Chinese Spring, *ph2a* and *ph2b*). As most *WM1* genes do not contain introns it is difficult to discern a PCR product derived from *WM1* mRNA as opposed to a PCR product from genomic contamination. To counteract this problem DNAase treament (DNAse treated BAC DNA control) of RNA samples, PCR positive controls (*MSH7* and *Thioredoxin H*) and a genomic contamination control PCR (Glutenin Promoter) was incorporated into all RT-PCR samples (**Figure 24**). Southern hybridisation of RT-PCR products (**Figure 24**) shows that *WM1* genes are expressed in all tissues but differentially. *WM1.1* is expressed in leaf, and meiotic spikes, *WM1.11* is expressed only in meiotic spikes whilst *WM1.2*, *WM1.3*, and *WM1.10* are all expressed in root and meiotic spike tissues. The only RT-PCR products, which were clearly visible after ethidium bromide staining was from *WM1.10* in both wild type (Chinese Spring) and *ph2b* meiotic spikes. As 5 μg of total RNA was used for each sample, RT-PCR products are partially quantitative when comparisons are made within samples for individual genes.

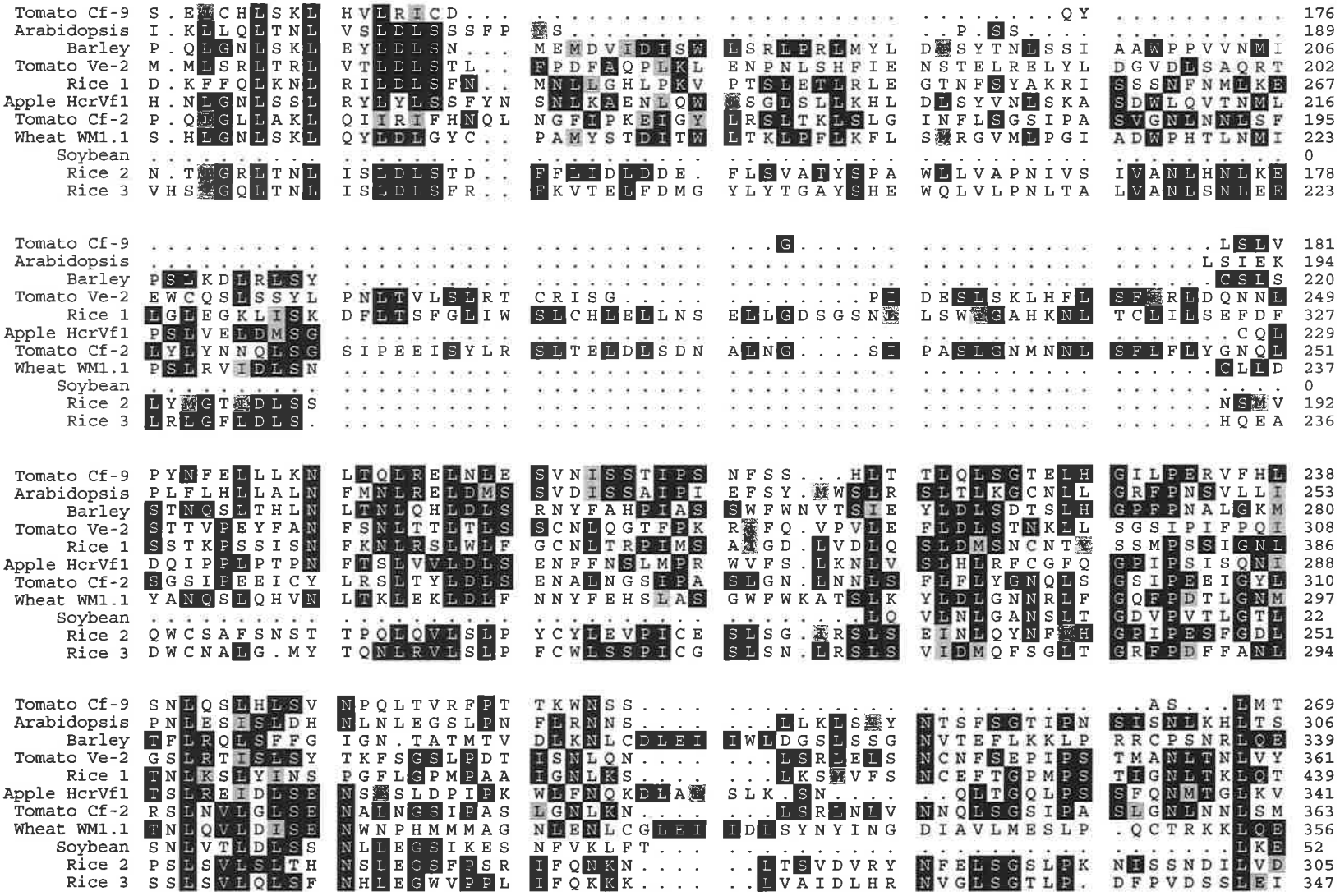
Figure 23: Relationship between the predicted primary polypeptide of *T. tauschii* WM1.1 and those polypeptides found to be most homologous as determined by the BLAST 2.0 blastp program (Altschul *et al.*, 1997). WM1-like peptides are from barley (*H. vulgare*), tomato (*Lycopersicon pimpinellifolium* and *esculentum*), *A. thaliana*, soybean (*Glycine max*), rice (*Oryza sativa*) and apple (*Malus floribunda*).

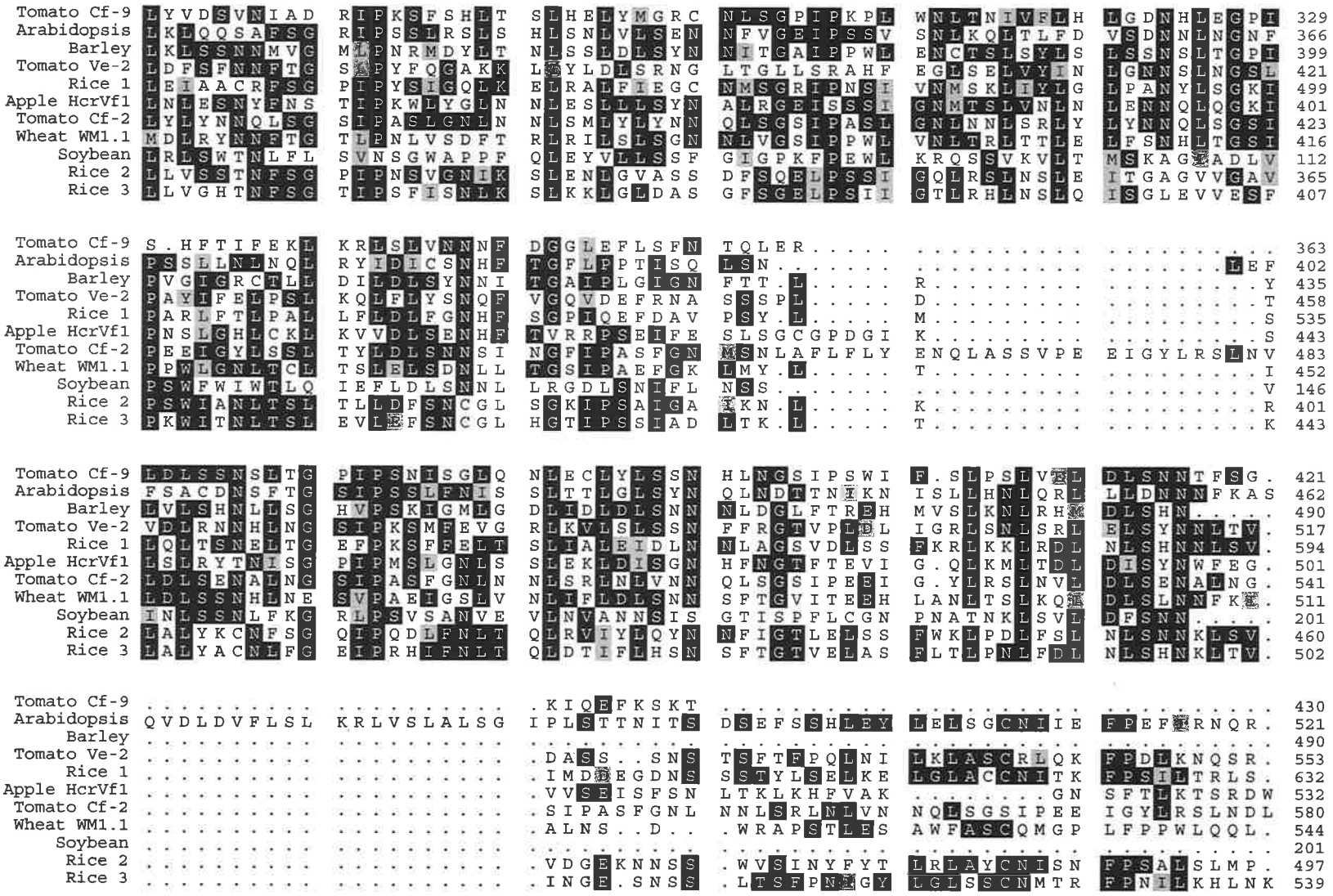
A - Prettybox alignment of the primary polypeptide sequences of LRR genes of tomato Cf-9 (Jones *et al.*, 1994), *Arabidopsis*, barley (Tokunaga *et al.*, 1999), tomato Ve-2 (Kawchuk *et al.*, 2001), rice 1, apple HcrVf1 (Vinatzer *et al.*, 2001), tomato Cf-2 (Dixon *et al.*, 1996), wheat WM1.1, soybean (de Barros *et al.*, 1999), rice 2 and 3.

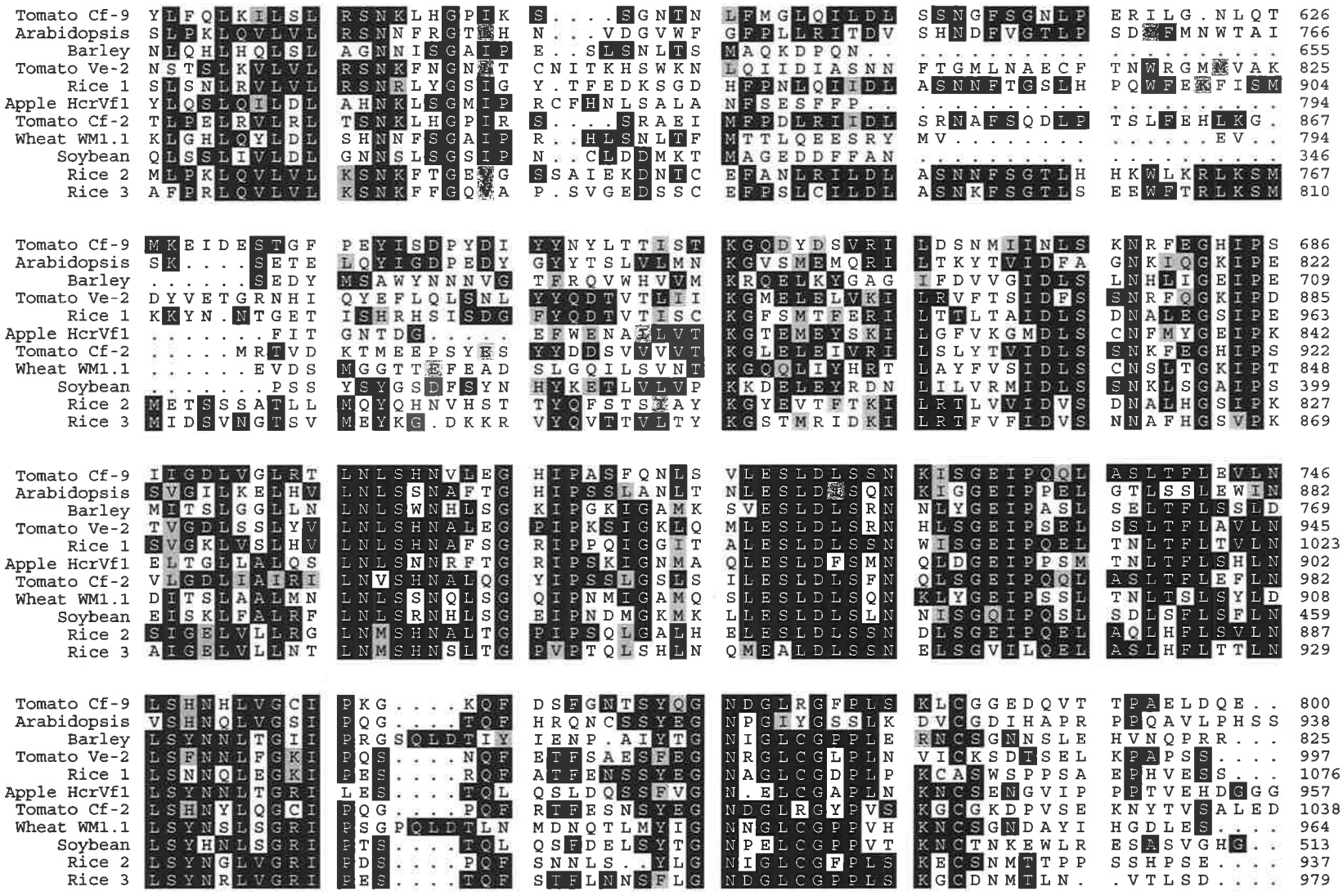
B – A rooted phylogenetic tree illustrating the relationship between the primary polypeptide sequences for all LRR-like proteins (alignment as per panel-A). Phylogentic tree illustrates the number of substitution events that allow discrimination between all LRR-like proteins analysed.

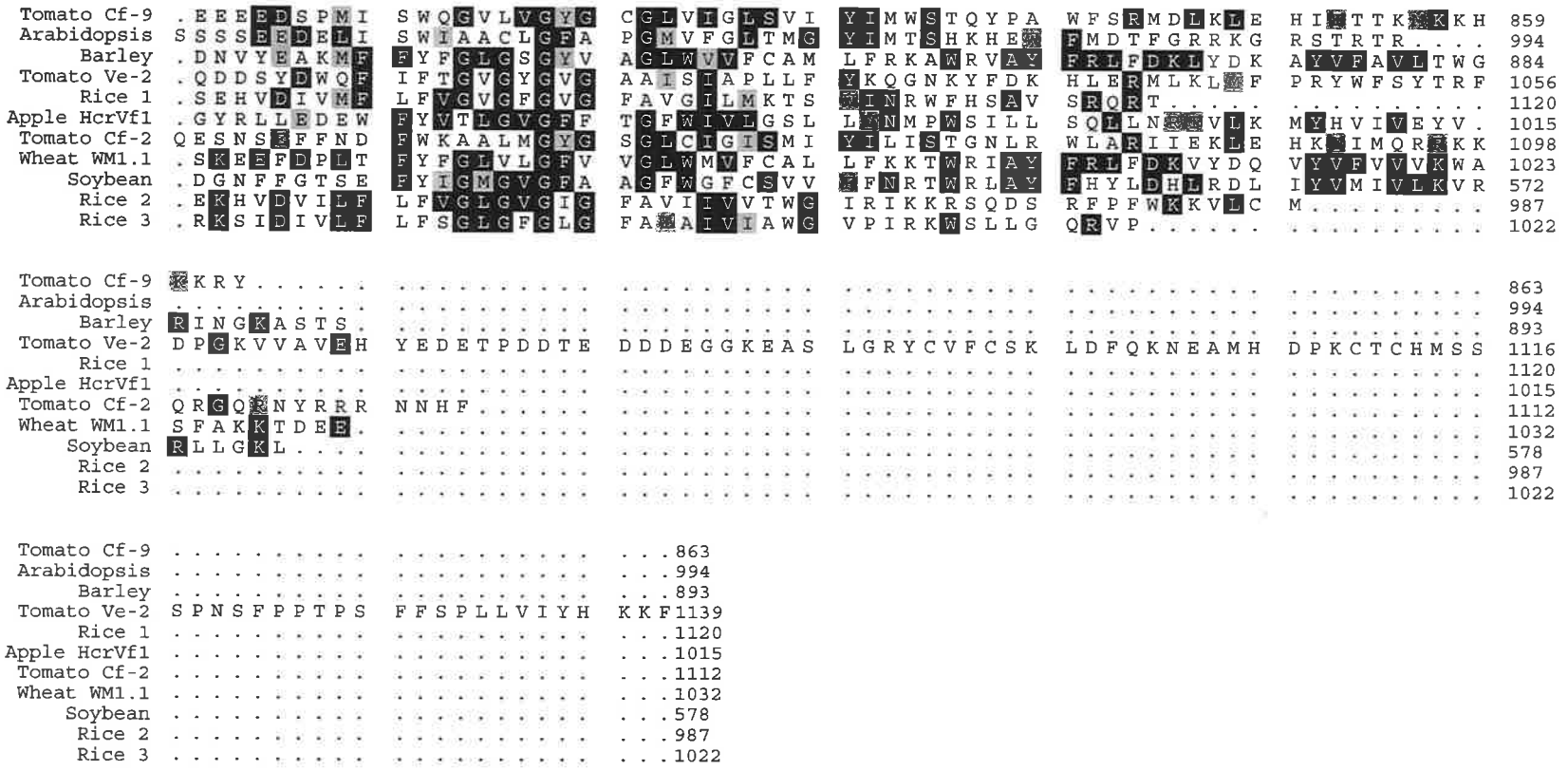
A


Tomato Cf-9
Arabidopsis
Barley
Tomato Ve-2
Rice 1
Apple HcrVf1
Tomato Cf-2
Wheat WM1.1
Soybean
Rice 2
Rice 3


Tomato Cf-9
Arabidopsis
Barley
Tomato Ve-2
Rice 1
Apple HcrVf1
Tomato Cf-2
Wheat WM1.1
Soybean
Rice 2
Rice 3


Tomato Cf-9
Arabidopsis
Barley
Tomato Ve-2
Rice 1
Apple HcrVf1
Tomato Cf-2
Wheat WM1.1
Soybean
Rice 2
Rice 3


Tomato Cf-9 626
Arabidopsis 766
Barley 655
Tomato Ve-2 825
Rice 1 904
Apple HcrVf1 794
Tomato Cf-2 867
Wheat WM1.1 794
Soybean 346
Rice 2 767
Rice 3 810
Tomato Cf-9 686
Arabidopsis 822
Barley 709
Tomato Ve-2 885
Rice 1 963
Apple HcrVf1 842
Tomato Cf-2 922
Wheat WM1.1 848
Soybean 399
Rice 2 827
Rice 3 869
Tomato Cf-9 746
Arabidopsis 882
Barley 769
Tomato Ve-2 945
Rice 1 1023
Apple HcrVf1 902
Tomato Cf-2 982
Wheat WM1.1 908
Soybean 459
Rice 2 887
Rice 3 929
Tomato Cf-9 800
Arabidopsis 938
Barley 825
Tomato Ve-2 997
Rice 1 1076
Apple HcrVf1 957
Tomato Cf-2 1038
Wheat WM1.1 964
Soybean 513
Rice 2 937
Rice 3 979


Tomato Cf-9 . E E E E D S P M I S W Q G V L V G Y G C G L V I G L S V I Y I M W S T Q Y P A W F S R M D L K L E H I T T K K H 859
Arabidopsis S S S S E E D E L I S W I A A C L G F A P G M V F G L T M G Y I M T S H K H E F M D T F G R R K G R S T R T R 994
Barley . D N V Y E A K M F F Y F G L G S G Y V A G L W V V F C A M L F R K A W R V A Y F R L F D K L Y D K A Y V F A V L T W G 884
Tomato Ve-2 . Q D D S Y D W Q F I F T G V G Y G V G A A I S I A P L L F Y K Q G N K Y F D K H L E R M L K L F P R Y W F S Y T R F 1056
Rice 1 . S E H V D I V M F L F V G V G F G V G F A V G I L M K T S I N R W F H S A V S R Q R T 1120
Apple HcrVf1 . G Y R L L E D E W F Y V T L G V G F F T G F W I V L G S L L N M P W S I L L S Q L L N V L K M Y H V I V E Y V . 1015
Tomato Cf-2 Q E S N S F F N D F W K A A L M G Y G S G L C I G I S M I Y I L I S T G N L R W L A R I I E K L E H K I M Q R K K 1098
Wheat WM1.1 . S K E E F D P L T F Y F G L V L G F V V G L W M V F C A L L F K K T W R I A Y F R L F D K V Y D Q V Y V F V V K W A 1023
Soybean . D G N F F G T S E F Y I G M G V G F A A G F W G F C S V V F N R T W R L A Y F H Y L D H L R D L I Y V M I V L K V R 572
Rice 2 . E K H V D V I L F L F V G L G V G I G F A V I I V V T W G I R I K K R S Q D S R F P F W K K V L C M 987
Rice 3 . R K S I D I V L F L F S G L G F G L G F A A I V I A W G V P I R K W S L L G Q R V P 1022

Tomato Cf-9 K R Y 863
Arabidopsis 994
Barley R I N G K A S T S 893
Tomato Ve-2 D P G K V V A V E H Y E D E T P D D T E D D D E G G K E A S L G R Y C V F C S K L D F Q K N E A M H D P K C T C H M S S 1116
Rice 1 1120
Apple HcrVf1 1015
Tomato Cf-2 Q R G Q N Y R R R N N H F 1112
Wheat WM1.1 S F A K K T D E E 1032
Soybean R L L G K L 578
Rice 2 987
Rice 3 1022

Tomato Cf-9 863
Arabidopsis 994
Barley 893
Tomato Ve-2 S P N S F P P T P S F F S P L L V I Y H K K F 1139
Rice 1 1120
Apple HcrVf1 1015
Tomato Cf-2 1112
Wheat WM1.1 1032
Soybean 578
Rice 2 987
Rice 3 1022

B

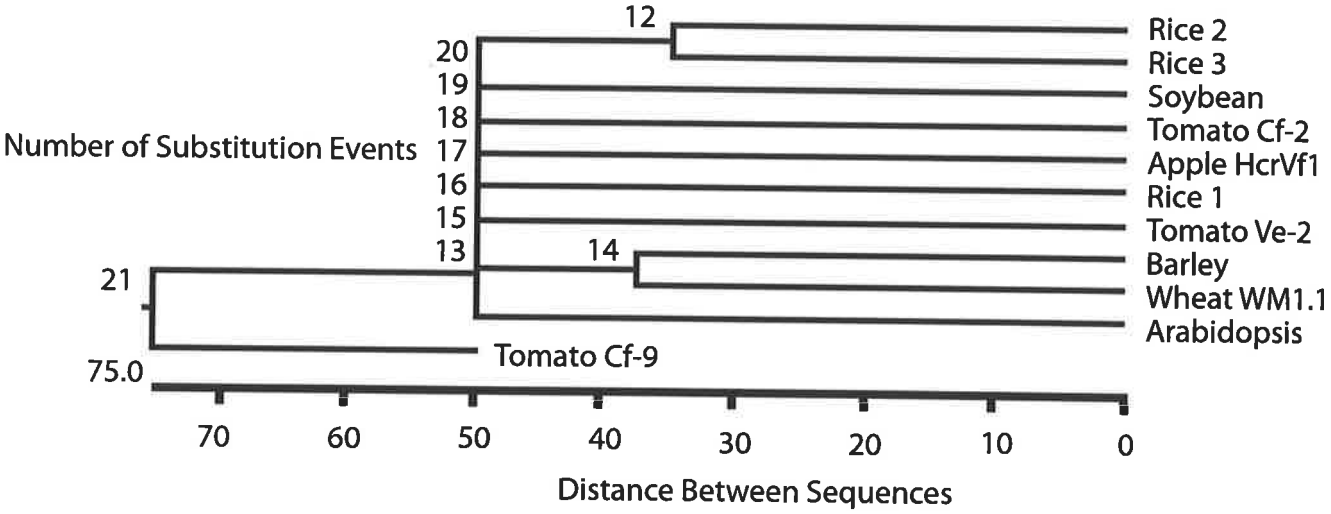
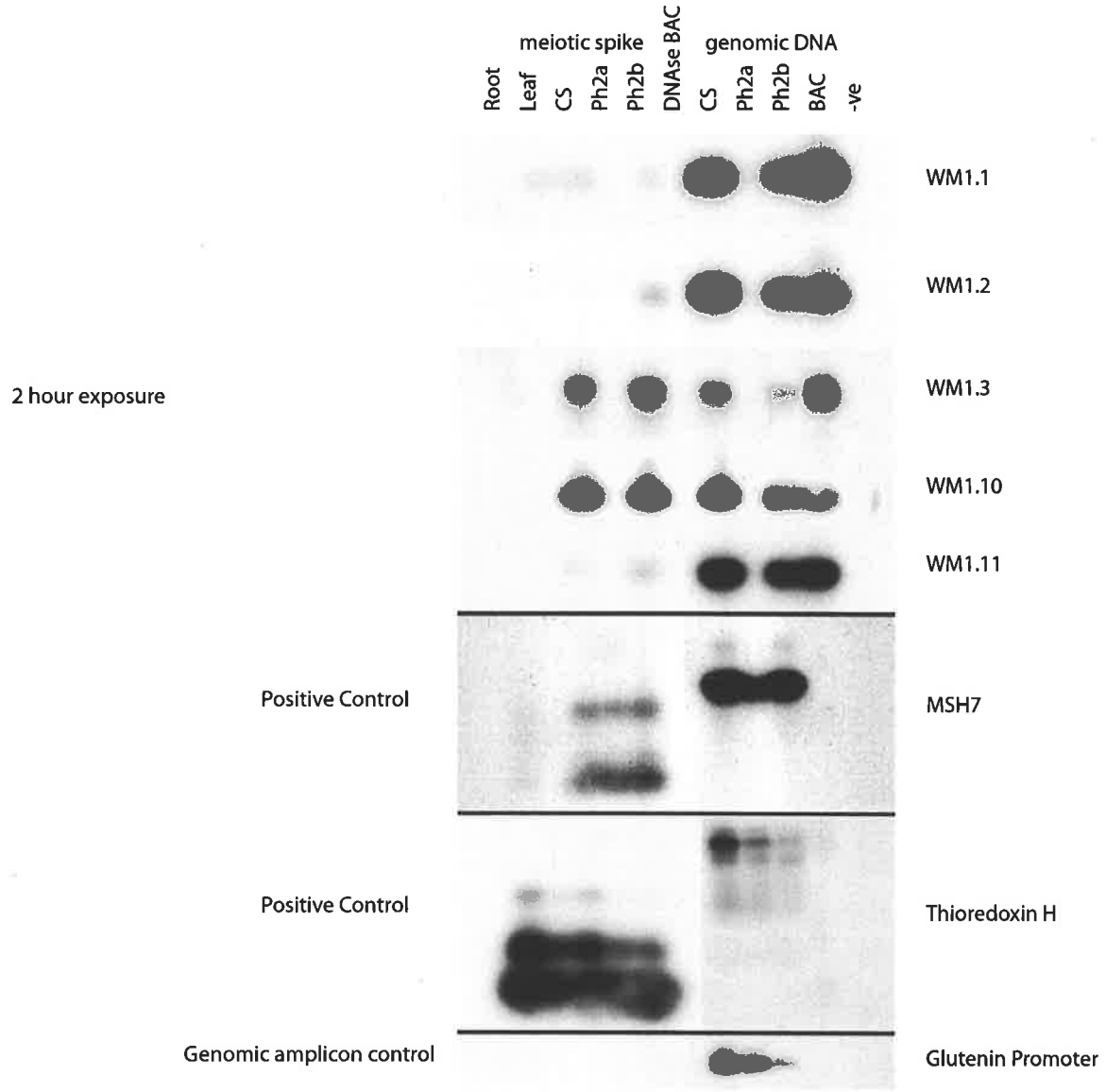
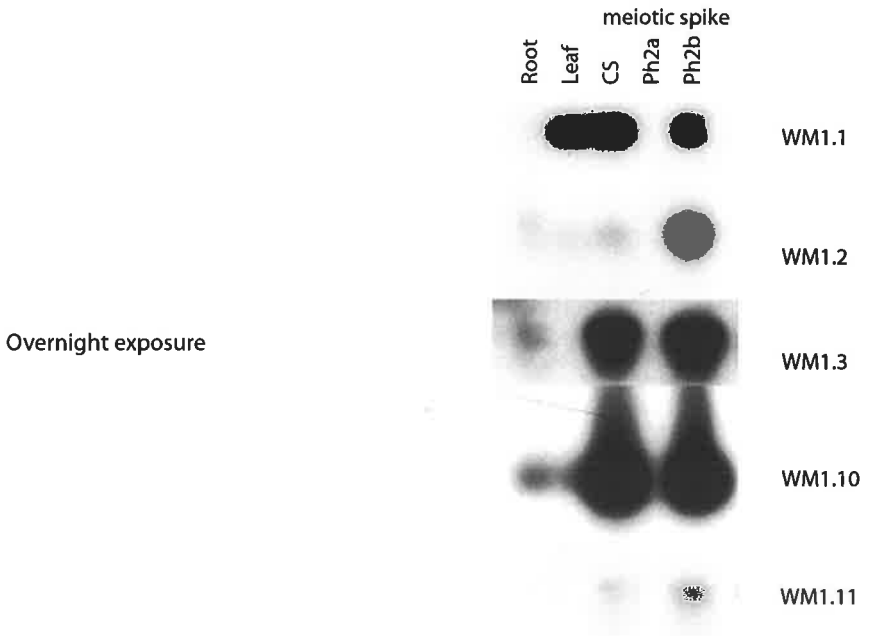

12
20
19
18
17
16
15
13
14
21
75.0
Number of Substitution Events
Rice 2
Rice 3
Soybean
Tomato Cf-2
Apple HcrVf1
Rice 1
Tomato Ve-2
Barley
Wheat WM1.1
Arabidopsis
Tomato Cf-9
70
60
50
40
30
20
10
0
Distance Between Sequences

Figure 24: Southern blot analysis of electrophoretically separated RT-PCR products of individual members of the *WM1* gene family. Positive controls include *MSH7* and *Thioredoxin H* whilst the genomic contamination control is the Glutenin promoter. A- 2 hr exposure at room temperature, B- overnight exposure at -80°C.

A


meiotic spike
DNAse BAC
genomic DNA
Root
Leaf
CS
Ph2a
Ph2b
CS
Ph2a
Ph2b
BAC
-ve
WM1.1
WM1.2
2 hour exposure
WM1.3
WM1.10
WM1.11
Positive Control
MSH7
Positive Control
Thioredoxin H
Genomic amplicon control
Glutenin Promoter

B


meiotic spike
Root
Leaf
CS
Ph2a
Ph2b
WM1.1
WM1.2
Overnight exposure
WM1.3
WM1.10
WM1.11

Results therefore suggest that *WM1* genes are preferentially expressed in floral tissues (**Figure 24**), and Northern analysis (data not shown) suggests that the overall expression level is extremely low.

3.4 Discussion

3.4.1 *WM1* genes cluster

The genes identified by hybridisation to *WM1.0* cDNA clone were named *WM1.1* to *WM1.21* (*W*heat *M*eiosis genes 1.1 to 1.21) (Ji, 1992). Using the *WM1.1* open reading frame (*WM1.1* ORF) as a probe in Southern analysis of wheat DNA the *WM1* gene family of approximately 21 genes is revealed (**Figure 7**), 14 members of the *WM1* gene family reside within the region deleted on chromosome 3DS in the wheat *ph2a* mutant (**Figure 8**). Additionally, 2 *WM1* genes are located on chromosome 3AS. However, no *WM1* genes were detected on chromosome 3BS. Sub-chromosomal assignment suggests that *WM1* genes are not represented across all chromosomes within homeogroup 3 and seem to be clustered on chromosome 3DS.

WM1 genes are not detectable by Northern analysis on either total RNA or polyA enriched RNA from wheat floral tissue. However mRNA transcripts were detected with RT-PCR, suggesting expression at very low levels. RT-PCR indicates that *WM1* genes are expressed predominantly in floral tissue, but some members have mRNA transcripts detectable in leaf and root tissues. The identification of an EST derived from *WM1.11* shows that this particular member of the gene family is likely to be expressed in endosperm 8-12 days post-anthesis. In light of the results so far, the only conclusion that can be drawn is that the *WM1* genes are differentially expressed between roots, leaves and floral tissue and are likely to be developmentally regulated.

Does the clustering of the *WM1* genes relate to developmental regulation? To what degree are the *WM1* genes clustered?

Genetic mapping of the *WM1* gene family in three barley-mapping populations Chebec x Harrington, Clipper x Sahara and Galleon x Haruna Nijo showed that

all genes co-segregate and are therefore tightly linked (refer to 5.3.1 and **Figure 37**). Consequently genetic mapping provides evidence that the *WM1* gene family clustered within 5 centimorgans.

Considering that *WM1* genes are genetically linked, Southern analysis of PFGE separated HMW DNA was utilised in determining physical linkage. Results were unable to demonstrate physical linkage between members of the *WM1* gene family (**Figure 9**). However, screening of both a barley and diploid wheat (*T. tauschii*) BAC library proved otherwise. Barley and wheat BAC clones showed that the *WM1* genes were indeed physically linked and clustering within a region of approximately 200 Kb. There are up to 11 *WM1* genes within this region (**Figure 13** and **Figure 14**).

To determine structural organisation of the *WM1* gene family three *T. tauschii* BAC clones, *TtBAC3*, *TtBAC6*, and *TtBAC11* were shotgun sequenced. Shotgun sequencing revealed the presence of seven *WM1* genes within a 270 Kb region (**Figure 19**). The seven *WM1* genes (*WM1.1-1.3*, *WM1.7*, and *WM1.10-1.12*) that were predicted in this region, are likely to be intron-less (**Figure 16**) and predominantly encode type Ia plasma membrane anchored leucine rich repeat proteins.

WM1 genes encode LRR receptor-like proteins and appear to be co-localised within a small region on chromosome 3DS of hexaploid wheat. Localisation of the *WM1* genes within such a defined region of the genome poses the question; Is this region of the genome structurally important and does this relate to the function of the *WM1* genes. Additionally, what are the likely evolutionary mechanisms that have operated on the *WM1* gene cluster?

3.4.1.1 WM1 *gene structural organisation*

WM1 genes are tandemly arranged in both a head-to-tail and tail-to-tail orientation (**Figure 15**) spanning the 270 Kb region. Flanking the *WM1* gene family cluster are multiple copies of genes encoding putative ripening-related proteins (3 copies), Bowman-Birk type trypsin inhibitor proteins (4 copies) and a single Pst19-like protein. Furthermore, 17 transposable element component

proteins with homology to hypothetical proteins of *A. thaliana*, *Oryza sativa*, and *T. aestivum* were located between the *WM1* genes. Interestingly a Pst19 protein has previously been found to be co-localised with a LRR-receptor like kinase gene cluster on chromosome $1A^mS$ of *T. monococcum* (A^m genome) (Feuillet *et al.*, 2001; Stein *et al.*, 2000; Wicker *et al.*, 2001). This cluster of genes was found whilst in the pursuit of the *Lr10* resistance gene. Further analysis may highlight an important evolutionary relationship between LRR-like genes and this *Pst19*-like sequence.

Most *WM1* genes seem to be separated by large stretches of transposable element-like sequences. Some of these repetitive sequences show homology to *BARE-1* and *Wis21-A*-like retrolelements. Preliminary analysis of the 270 Kb region shows that it contains approximately 80% repetitive sequence. It is anticipated that further analysis will allow a greater understanding of transposable element evolution and help reveal their impact on the origin of LRR gene clusters both in terms of gene expression and function.

3.4.1.2 WM1 *gene evolution*

When looking at the physical arrangement of *WM1* genes with respect to one another, similarities can be drawn to other LRR gene clusters. The presence of tandemly duplicated homologous sequences at the *WM1* locus could promote chromosomal mispairing, resulting in non-reciprocal homologue exchange or unequal crossing-over. This may induce meiotic instability of different *WM1* alleles as has been shown with alleles at the *Rp1* rust resistance locus in maize (Hulbert, 1997) and furthermore this could result in sequence homogenisation within a gene family through frequent intergenic sequence exchange. This would be desirable at developmental gene loci where conservation in LRR genes maintains integrity, which is likely to be crucial for development. On the other hand it is undesirable at resistance gene loci where novel variants in LRR genes are required to combat a constantly changing pathogen population (Thomas *et al.*, 1998).

Despite the prospect for gene homogenisation, analysis of the sequence 3' and 5' to the *WM1* coding regions shows high sequence divergence suggesting that these genes were not generated from a recent crossover or gene conversion event as was shown for several *Cf* genes (Parniske *et al.*, 1997; Parniske & Jones, 1999; Parniske *et al.*, 1999). It is anticipated that the rate of such crossing-over or gene conversion, resulting in duplication and deletion of whole *WM1* genes or blocks of LRRs with genes, is extremely low. Such changes in the *Cf-9* LRR gene resulting in loss of function are believed to occur once in 22,000 meiotic events (Parniske *et al.*, 1997).

Interestingly, the 5' sequence for *WM1.11* includes a short region with homology to the leucine zipper region of the *Yr10* (Laroche *et al.*, 2000) and *Rpm1* (Grant *et al.*, 1995) disease resistance genes (refer to **Figure 16**). This region is upstream from the predicted basal promoter region and is not a part of the coding sequence. The *WM1.11* EST derived from endosperm 8-12 days post anthesis confirms this. These results suggest that there may be an evolutionary relationship between the leucine zipper – nucleotide binding site - LRR (LZ-NBS-LRR) class of disease resistance proteins and the purely LRR receptor-like class of proteins (*WM1* gene family). Further analysis of the *WM1* gene cluster may highlight the importance of this evolutionary link in terms of function.

The developmental expression pattern, co-localisation and genic structure around the region of *WM1* genes is interesting but what is their likely function in the plant? Without transformants, the predicted peptide sequences of the *WM1* genes must be used to elucidate potential functions?

3.4.2 *WM1* gene analysis

Several gene classes, including histones (Chaboute *et al.*, 1993) and polyphenol oxidases (Newman *et al.*, 1993), have been identified in plants that generally lack introns and encode short polypeptides (< 62kDa). It appears that the *WM1* family of LRR receptor-like proteins (~112 kDa) encode the largest genes believed to be intron-less in plants.

The *WM1* primary peptides show highest homology and structural similarity with the *Cf-2* (Dixon *et al.*, 1996), *Cf-9* (Jones *et al.*, 1994), *Ve-2* (Kawchuk *et al.*, 2001), and *HcrVf1* (Vinatzer *et al.*, 2001) receptor-like disease resistance genes from plants. The WM1 predicted proteins possess a hydrophobic N-terminus, which is typical for a signal peptide. Signal peptide cleavage is likely to be important for targeting these receptor-like proteins to the plasma membrane.

3.4.2.1 Leucine rich repeats

The signal peptide precedes a long stretch of tandemly arrayed LRRs with between 9 and 30 imperfect copies of a 23 amino acid consensus [XXLXXLXLXXLDLSXNXLSGXIP]. These tandemly arrayed LRRs have been found in many proteins with diverse functions. As with many LRR-like genes, the *WM1* primary peptides have length variation in the N-termini. This region is not only heterogenous in repeat number but also in repeat length. The *Drosophila* chaoptin (Reinke *et al.*, 1988), for example, has 41 repeats, whereas the *Rec-A* like protein of *Arabidopsis* (Pang *et al.*, 1992) has only four. The LRRs are generally believed to be involved in protein-protein interaction (Gay *et al.*, 1991; Suzuki *et al.*, 1990) but the mode and specificity of the interaction is unclear. A glycine present within the consensus sequence is consistent with the expectation that they are extracellular, with its location likely to assist in the detection of a signalling ligand (Jones *et al.*, 1994; Song *et al.*, 1995). In plants, several LRR proteins have been isolated in recent years. They fall into several major classes, the receptor-kinase like proteins and the disease resistance proteins (Braun & Walker, 1996; Dangl, 1995). One of the former types, the *Xa21* gene product from rice, has been found to function as a disease resistance molecule (Song *et al.*, 1995).

Although LRR-like proteins are associated with diverse functions, a common property involves protein-protein interaction. The 3-D structure of a ribonuclease inhibitor, a protein containing 15 LRRs, has been determined (Kobe & Deisenhofer, 1993), revealing LRRs forming α/β coil structures. These 15 LRR's folded in an unusual horseshoe structure, which has the ability to clamp onto the ribonuclease. The LRRs of WM1 proteins suggest that they are

not likely to adopt the α/β coil structure as they are highly homologous to other plant LRRs which have been shown to contain residues not common in α-helices (Dixon *et al.*, 1996; Thomas *et al.*, 1997). The likely β-helical structure (Kobe & Deisenhofer, 1994) of these plant LRRs would aid the parallel stacking of the β-strand/β-turn motif, resulting in the formation of a rod-like structure with an extensive and potentially variable ligand-binding surface (Thomas *et al.*, 1998).

The greatest homology of the *WM1* primary peptides is with *Cf-2* and *Cf-9* *Cladosporium fulvum* disease resistance genes of tomato (Dixon *et al.*, 1996; Jones *et al.*, 1994). Highest homology lies in the C-terminal region of these proteins, composed of one 23 and three 24 amino acid LRRs, as well as a transmembrane domain (**Figure 22**). Variation in the amino-terminal LRRs of *Cf* genes, as with *WM1* genes, is likely to have arisen through mutation, segmental exchange between adjacent homologues within tandem arrays (either by repeated rounds of unequal exchange or by gene conversion) and duplication or deletion of complete LRR units: For example, Cf-4 differs from Cf-9 by a precise deletion of two complete LRRs (Thomas *et al.*, 1997). It has been suggested that this N-terminal region plays a role in ligand binding specificity. This has implications when assessing the potential tertiary structure of the LRRs. The predicted parallel stacking of the β-sheet arrangement of LRRs means that any given amino acid has both a horizontal context, the neighbouring amino acids within its own LRR unit, and a vertical context, amino acids in a similar position in the two flanking LRRs. Segmental exchange, deletion and duplication of whole LRRs, have limited roles in producing novel combinations of amino acids in the horizontal context, but would seem to have a more important role in producing novel combinations of amino acids in the vertical context. This is likely to have important repercussions in specificity of recognition of protein ligands (Ellis & Jones, 1998) and may prove extremely important when determining the role of individual members of the *WM1* gene family.

Between 7 and 27 sequences matching the N-glycosylation consensus sequence NX(S/T) were observed within the predicted LRR regions of the *WM1*

primary peptides (data not shown). This would suggest WM1 protein glycosylation, which may disrupt the regular tertiary structure of LRRs and subsequently influence potential ligand binding. Such suggestions have been made based on the *Cf-9* gene (Piedras *et al.*, 2000; Thomas *et al.*, 1998).

3.4.2.2 Leucine zipper

Several WM1 proteins (*WM1.1*, *WM1.2*, *WM1.10*, and *WM1.12*) have 100% identity with a leucine zipper-like domain (**Figure 19**). The leucine zipper consists of a periodic repetition of leucine residues at every seventh position over a distance covering eight helical turns. These predicted leucine zippers are embedded within two LRRs; consequently the two structures would be mutually incompatible based on secondary structural characteristics. Leucine zippers require an α-helix conformation whilst leucine rich repeats form a β-strand flanked by β-turns. The presence of multiple helix-breaking prolines within the potential leucine zipper of WM1 peptides would oppose the formation of a leucine zipper. Given the leucine zipper structure is unlikely to form, the potential for LRRs to be involved in the dimerisation process is not negated. The involvement of LRRs in heterodimerisation of such type Ia plasma membrane anchored LRR proteins is yet to be proven.

3.4.2.3 Sorting signals

A hydrophobic sequence with a predicted α-helical secondary structure, which is bordered by a negatively charged extracytoplasmic domain and a positively charged cytoplasmic domain, are typical in membrane-spanning proteins (**Figure 19**). Importantly cytoplasmic domains for all WM1 predicted peptides except WM1.10 possess a YFRL motif. This matches the mammalian consensus for the typrosine YXXϕ endosomal/lysosomal sorting signal sequences, where ϕ is an amino acid with a hydrophobic side chain that stimulates receptor mediated endocytosis and degradation of cell-surface receptors (Letourneur & Klausner, 1992; Marks *et al.*, 1996). Whether the tyrosine YXXϕ motif functions as an endocytosis signal in plants is yet to be proven experimentally. Availability of complete DNA sequence for *WM1* genes

will help in experimental confirmation of the localisation of the WM1 proteins, through GFP fusion proteins, to the cell membrane and consequently determine the function of the YXXϕ motif in receptor mediated endocytosis and degradation.

Unlike the tomato *Verticillium* resistance (*Ve-1* and *Ve-2*) peptides, the WM1 peptides do not possess Pro-Glu-Ser-Thr (PEST) sequences that are commonly found in proteins with short cytoplasmic half-lives (Rechsteiner & Rogers, 1996; Rogers *et al.*, 1986) (**Figure 22**). Additionally WM1 peptides do not possess KKX motifs that signal endoplasmic reticulum retention in plants (Benghezal *et al.*, 2000).

3.4.3 Relating developmental regulation to disease resistance

It would be interesting to investigate if *WM1* genes are upregulated in response to pathogen attack as would be expected based on the high level of peptide homology to disease resistance genes (*Cf-2*, *Cf-9*, *Ve-2*) (Kawchuk *et al.*, 2001; Thomas *et al.*, 1998) and the induced barley *Cf-2/Cf-5* like LRR gene isolated by differential display of BTH-treated plants (Tokunaga *et al.*, 1999). BTH (benzo(1,2,3)thiadiazole-7-carbothioic acid S-methyl ester) is a new class of commercial agrochemicals that indirectly control fungal plant diseases through the activation of the plant's own defences (Kessman & Nordmeyer, 1996). In dicots, BTH is thought to act as a functional analogue of the endogenous defence-signalling molecule, salicyclic acid, and induce systemic aquired resistance (Du & Klessig, 1997; Hunt & Ryals, 1996; Ryals *et al.*, 1996).

The first step in determining if *WM1* genes have a role in plant defence is to determine if any known genes for disease resistance reside on chromosome 3DS in the locality of the *WM1* gene cluster. One such example is leaf rust resistance gene 32 (*Lr32*) (Kerber, 1988). It would need to be determined whether any of the *WM1* genes are upregulated upon infection by this pathogen. It is also important to assess whether *WM1* genes are upregulated upon BTH treatment, which may suggest a role in SAR. Results of these experiments may

prove a link between developmental regulation and disease resistance as well as the possible recruitment of developmental genes for disease resistance.

3.4.4 A tentative model for developmental signal transduction

As has been shown with the *Cf* genes, WM1 peptide conservation is at the C-terminus. A model proposed by de Wit and Joosten (1999) for perception of avirulence factors (Avr9) suggests that Cf-9 is one of at least three molecules required in Cf-9-dependant signal transduction leading to a hypersensitive response (**Figure 25**). The conserved C-terminal domain of Cf-9 is proposed to interact with the HABS-Avr9 complex. HABS is known to be a high affinity binding site for Avr9 in the mesophyll cells of tomato. It was suggested that the HABS molecule might represent an LRR-receptor-like kinase similar to the *Xa21* disease resistance gene from rice (de Wit & Joosten, 1999).

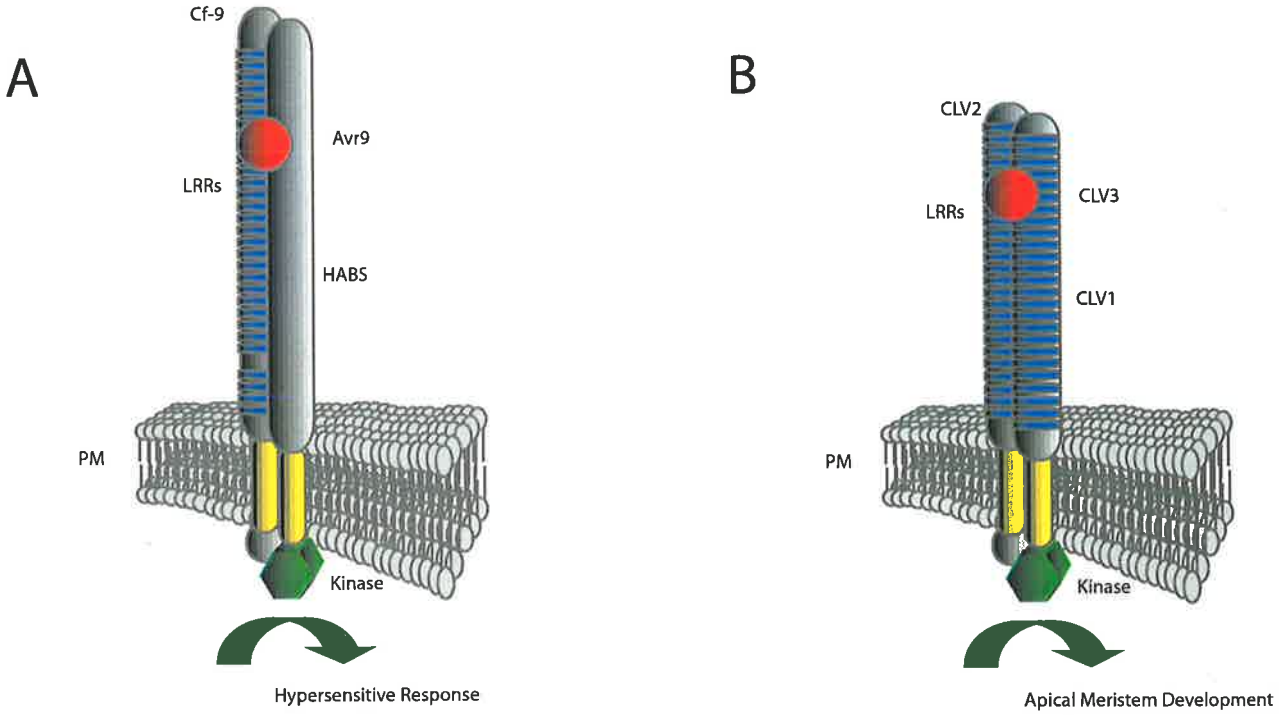
The Cf-9/Avr9 model of signal transduction shows similarities to the *CLAVATA* (*CLV*) family's role in *Arabidopsis* shoot apical meristem development (Becraft, 1998). Mutations in the *CLV* family of genes result in the apical meristem developing into a club-like structure. CLV1 is a LRR receptor like kinase (analogous to HABS) (Clark *et al.*, 1997), while CLV3 might be a ligand analogous to Avr9 (Clark *et al.*, 1995). Similarly CLV2 could be analogous to Cf-9 and consequently the *WM1* genes. A tentative model is proposed for developmental signal transduction of the *WM1* genes based on models for the *CLV* gene family and Cf-9/Avr9 interaction (**Figure 25**). However, a role for the *WM1* genes in disease resistance should not be disregarded.

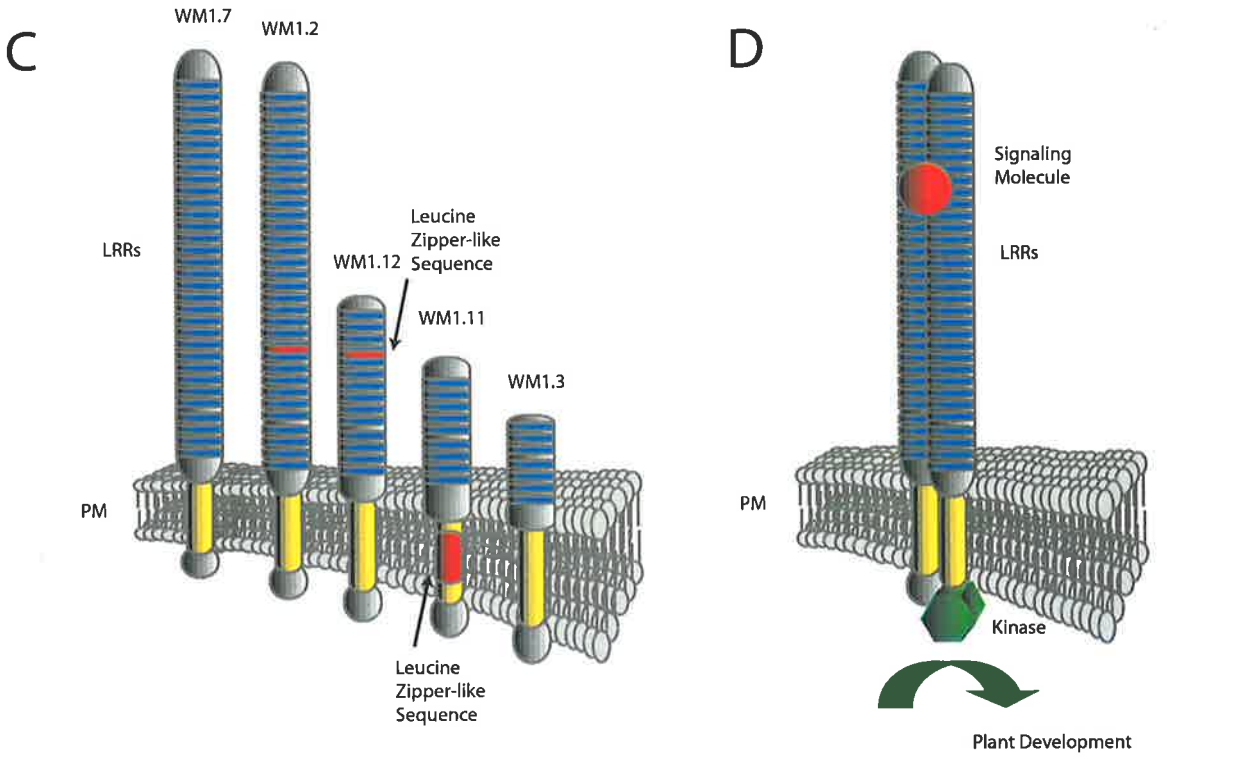
3.5 Conclusions

The *WM1* genes represent a complex gene family that seem to be developmentally regulated with low but preferential expression in floral tissues. The *WM1* genes are located on chromosome 3DS within a region of the wheat genome that is deleted in the *ph2a* mutant. Whether any of these *WM1* genes represent the *Ph2* gene(s) requires cytological examination of chromosome pairing behaviour in both "knock-out" and over expression transformants. A role

for the *WM1* genes in the perception of pathogen attack cannot be excluded and should also be investigated further.

Figure 25: A speculative model for the role of the predicted membrane bound receptor-like WM1 peptides in perception and signaling leading to regulation of floral tissue development in wheat (D). The model is based on evidence from the Cf-9/Avr9 disease perception model (A) and the model of the CLAVATA gene family's role in apical meristem development (B) (de Wit & Joosten, 1999). Membrane topology and relative predicted structure of individual members of the WM1 gene family is shown in (C).


A
Cf-9
Avr9
LRRs
HABS
PM
Kinase
Hypersensitive Response
B
CLV2
LRRs
CLV3
CLV1
PM
Kinase
Apical Meristem Development


C
WM1.7
WM1.2
LRRs
WM1.12
Leucine
Zipper-like
Sequence
WM1.11
WM1.3
PM
Leucine
Zipper-like
Sequence
D
Signaling
Molecule
LRRs
PM
Kinase
Plant Development

4 Chapter

IN PLANTA ANALYSIS OF THE MEIOTIC GENE *WM5*

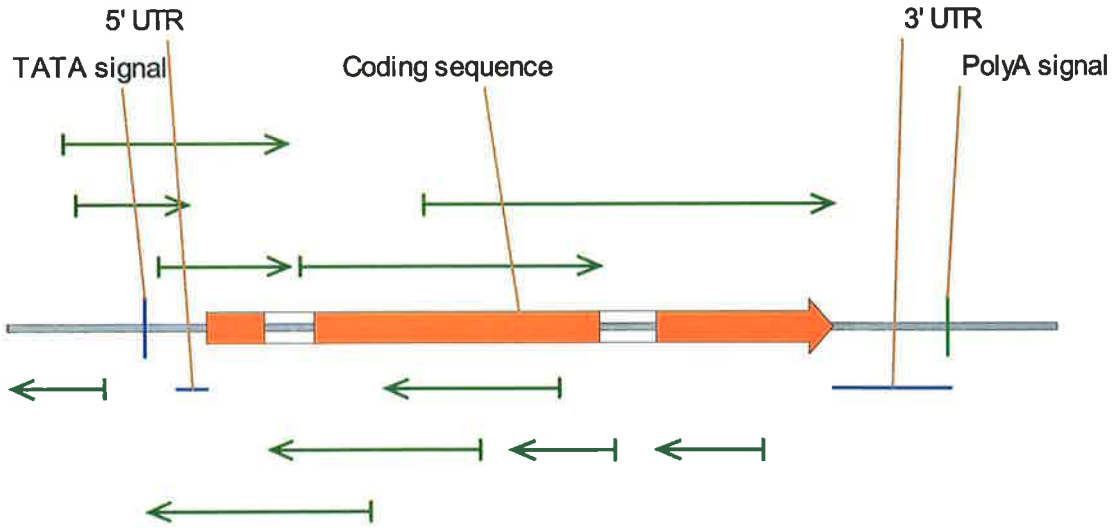
4.1 Introduction

The Wheat Meiosis 5 (*WM5*) gene was isolated by subtractive hybridisation of mRNA from pollen mother cells at pre-meiotic interphase with mRNA from immature pollen (Thomas, 1997). *WM5* is strongly and preferentially expressed during early meiosis in the flower bud and expression gradually reduces as meiosis proceeds. Lower levels of expression have also been detected in actively growing young shoot apices. The gene is not expressed in highly active mitotic tissue such as the root. Given this expression pattern, *WM5* may be involved in the establishment of and development of the shoot meristem including both the vegetative and reproductive apex. *WM5* may also have functions associated with early meiotic activity, such as chromosome pairing and recombination.

DNA sequence analysis of *WM5* (Dong *et al.*, 2001a) has shown no significant homology to known genes. **Figure 26** describes the structure of *WM5* both from the nucleotide and the polypeptide level. The deduced polypeptide of *WM5* is rich in glycine (16.3%), serine (15.1%), proline (10.8%) and alanine (12.9%) and is likely to possess two hydrophobic regions, one in the amino terminus (residues 1-25), which suggests a signal peptide for either membrane targeting and possibly for secretion of the processed protein. The second hydrophobic domain is in the carboxy-terminal region, which suggests a possible membrane anchoring function. Both regions contain a high percentage of alanine. The central region is rich in glycine, serine and proline, and is highly repeated. The repeating motifs include PSHGG, GGAYGSSPTP.

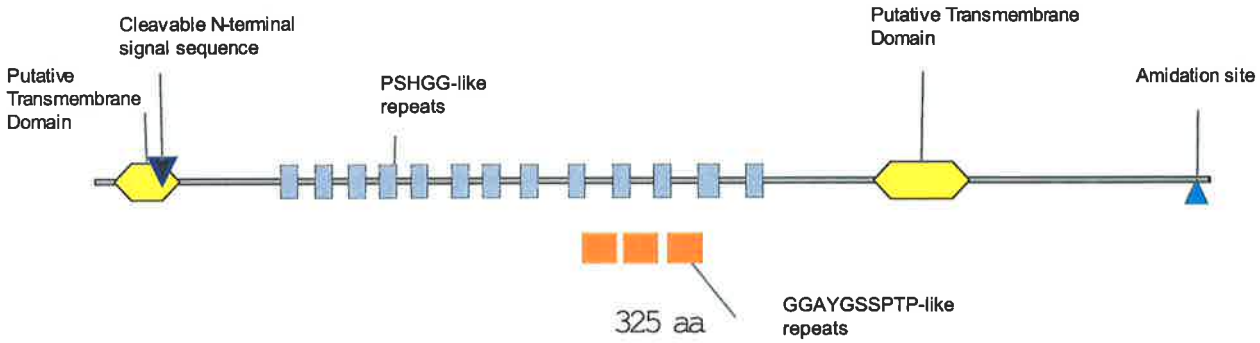
Figure 26: Diagram of the nucleotide and polypeptide sequence of *WM5* as described by (Dong *et al.*, 2001a). A - nucleotide diagram illustrates 3' and 5' untranslated regions, TATA signal, poly-adenylation signal and the coding sequence showing exon/intron boundaries. B – primary polypetide sequence highlighting putative transmembrahne domains, cleavage site for N-terminal signal sequence, PSHGG-like repeats, GGAYGSSPTP-like repeats, and a putative amidation site. The green arrows depict all possible open reading frames.

A


5' UTR
TATA signal
Coding sequence
3' UTR
PolyA signal

1965 bp

B


Cleavable N-terminal signal sequence
Putative Transmembrane Domain
PSHGG-like repeats
Putative Transmembrane Domain
Amidation site
325 aa
GGAYGSSPTP-like repeats

The overall WM5 amino acid sequence shows no significant homology with other known proteins. However, the carboxy-terminal end of WM5 protein shares homology with *Arabidopsis* PDF1 (protodermal factor actor 1) protein (Genbank accession number AF141375) (Abe *et al.*, 1999; Abe *et al.*, 2001). *Arabidopsis* PDF1 gene encodes a proline-rich protein, that is produced in the protodermal L1 layer of meristematic tissues (Abe *et al.*, 1999). Comparison of WM5 and PDF1 reveals that two proteins have very similar hydropathy plots and predicted secondary structure, which indicates that they have a similar function.

Southern analysis (Dong *et al.*, 2001a) has revealed that there is a copy of *WM5* on the short arm of each group three chromosome in the bread wheat variety Chinese Spring. Importantly, one copy is located within the region deleted in the *ph2a* mutant.

Given that *WM5* has a preferential meiotic expression pattern and is localised within the region deleted in the homeologous chromosome pairing mutant *ph2a*, it is a candidate for the *Ph2* gene. To confirm the function of *WM5* in either apical meristematic development (similar to *PDF1* of *Arabidopsis*) or homeologous chromosome pairing and recombination (as described in the *ph2a* wheat mutant), transgenic wheat lines were generated with altered expression of *WM5*.

Three different transformation systems have been successfully used for wheat.

Direct gene transfer into wheat via protoplasts (He *et al.*, 1994). This technique has been technically difficult and requires a long period in culture often resulting in the regenerants having reduced fertility and vigour.

The use of *Agrobacterium tumefaciens* (Cheng *et al.*, 1997) offers advantages of single copy transgene integration. However the time required in the production of transgenic plants is longer (2.5 – 3 months) than required for microprojectile bombardment (8 – 9 weeks).

Microprojectile bombardment is the most common transformation method used in wheat (Lörz *et al.*, 1998). The advantage of microprojectile bombardment comes from its ease of use, range of source tissue and high levels of transformation frequency relative to the other transformation techniques.

In summary, this chapter describes the use of microprojectile bombarment of wheat scutella in the production of transgenic wheat harbouring modified forms of the *WM5* gene. In addition, analysis of promoter specificity and WM5 functionality *in planta* was also investigated.

4.2 Materials and methods

4.2.1 Plant material and growth conditions

Four wheat (*Triticum aestivum* L.) genotypes were used in biolistic transformation experiments.

Spring type:

"Veery N° 5", CIMMYT variety listing (CIMMYT, Mexico);

"Combi", German variety listing 1998 (Federal variety names, Hannover, Germany);

"Pavon", German variety listing 1998 (Federal variety names, Hannover, Germany).

Winter type:

"Florida", German variety listing 1998 (Federal variety names, Hannover, Germany).

Vernalisation of winter wheat was undertaken 14 days after germination for eight weeks at 2°C in a growth incubator (9 hr, 4000 lx). Winter wheat was then placed in the glasshouse under controlled lighting and temperature (18°C/16°C Day/Night, 16 hr light at 23000-25000 lx). Fertilisation took place 8 weeks after

vernalisation. The fertiliser mixture was made up with 0.7% Wuxal Top N 12:4:6 (12% N, 4% P_2O_5, 6% K_2O, Trace elements, Aglukon, Düsseldorf, Germany).

Spring wheats were grown for 6 weeks in a phytotron (18°C/16°C Day/Night, 16 hr light at 10000-16000 lx) and then placed in the glasshouse until mature. Growth conditions and fertilisation were the same as for winter wheat.

4.2.2 Gene constructs

Microprojectile bombardment was performed using plasmid constructs as described in **Figure 27**. Design for co-bombardment of constructs is outlined in **Table 11**. According to the experimental design outlined in **Table 11**, regenerant plantlets (R_0) were labelled according to experimental co-bombardment class (eg. X, Y or Z).

4.2.3 Microprojectile bombardment

Plasmid DNA was adsorbed onto gold particles of an average size between 0.4 and 1.2 μm (Heraeus, Karlsruhe, Germany) as described previously (Becker *et al.*, 1994). The particle-DNA pellet was re-suspended in 240 μL of ethanol of which 3.5 μL of the particle-DNA suspension was spread onto the surface of the macrocarrier for each experiment.

The particle gun employed in these experiments was a PDS 1000/He gun (Biorad, München, Germany). Microprojectile bombardment details are outlined in **Table 12**.

4.2.4 Culture and selection conditions

Developing grains from the four wheat genotypes were harvested 12 to 14 days after pollination (Becker *et al.*, 1994).

Figure 27: Diagramatic representation of constructs used in biolistic transformation of wheat.

A - Antisense *WM5* (*pWM53.1*) *pBluescript KS(+)* containing antisense *WM5* cDNA (1.1 Kb) with NOS terminator (0.25 Kb) driven under 2.3 Kb of native promoter.

B - Sense *WM5* (*pWM54*) *pBluescript SK(+)* containing sense *WM5* genomic DNA with both native terminator and promoter (total 4 Kb). The third exon with a 30bp linker insertion (at *Bst EII* restriction endonuclease site) harbouring a *Bam HI* restriction endonuclease site.

C - 35S *WM5* (*pWM55.6*) *pTZ18U* containing sense *WM5* cDNA (1.1 Kb) with NOS terminator (0.25 Kb) driven under Cauliflower Mosaic Virus (*CaMV*) 35S-promoter (0.77 Kb).

D – *WM5* GUS (*pWM51*) *pBluescript KS(+)* containing β-glucuronidase (GUS) (2 Kb) with NOS terminator (0.25 Kb) driven under 2.3 Kb of *WM5* promoter.

E – *WM5* GFP (*pWM52.1*) *pBluescript KS(+)* containing green fluorescent protein gene from *pMON30049* (Pang *et al.*, 1996) (1.75 Kb) with NOS terminator driven under 1.4 Kb of *WM5* promoter.

F - *pAct1-F/Npt II pBluescript KS(+)* containing Kanamycin resistance gene *Npt II* (795bp) with NOS terminator driven under 1.25 Kb Act1-F fragment of the rice actin promoter (McElroy *et al.*, 1991).

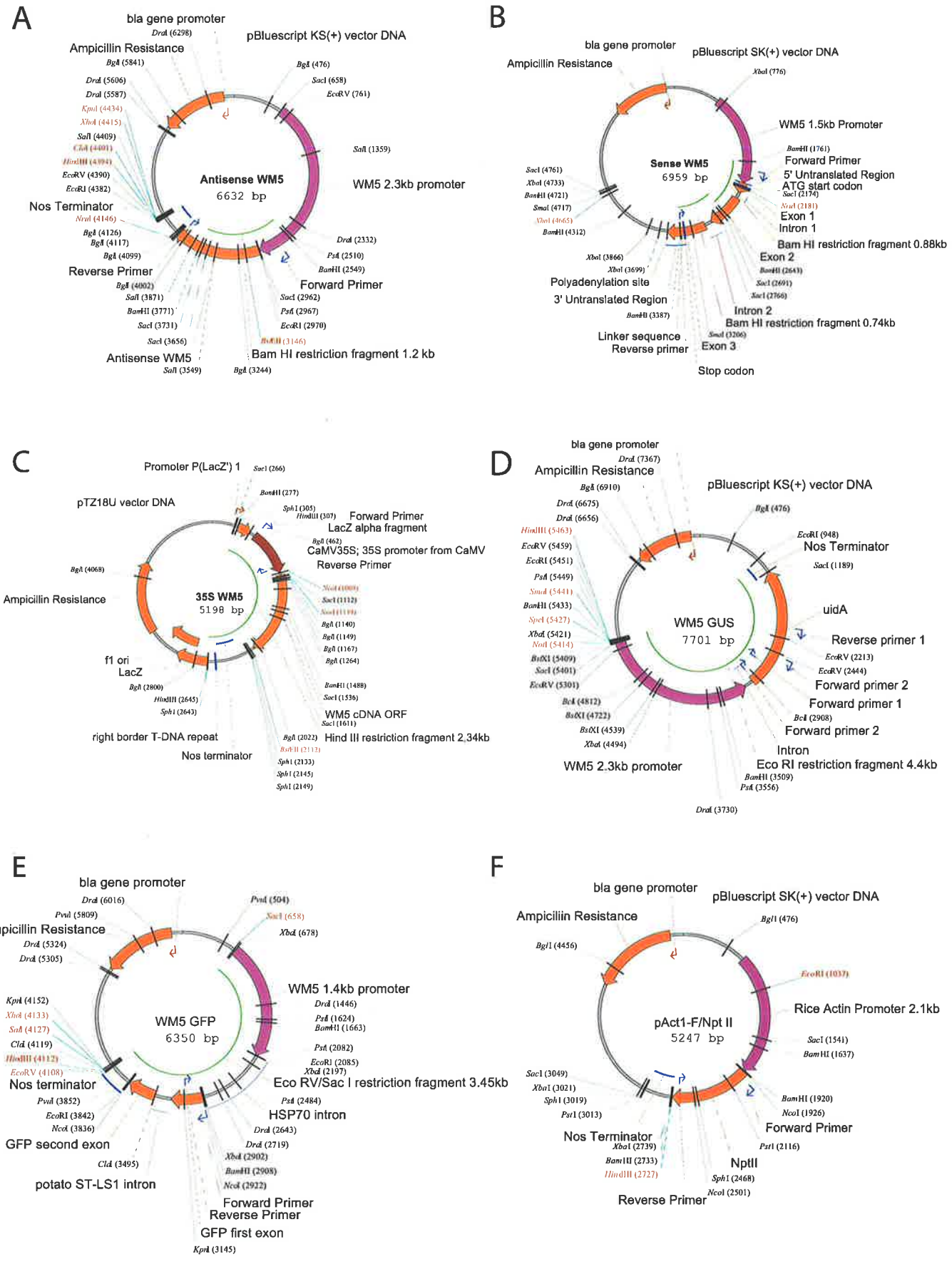

A
bla gene promoter
pBluescript KS(+) vector DNA
Ampicillin Resistance
Antisense WM5
6632 bp
WM5 2.3kb promoter
Nos Terminator
Reverse Primer
Forward Primer
Antisense WM5
Bam HI restriction fragment 1.2 kb
B
bla gene promoter
pBluescript SK(+) vector DNA
Ampicillin Resistance
Sense WM5
6959 bp
WM5 1.5kb Promoter
Forward Primer
5' Untranslated Region
ATG start codon
Exon 1
Intron 1
Bam HI restriction fragment 0.88kb
Exon 2
Intron 2
Bam HI restriction fragment 0.74kb
Exon 3
Polyadenylation site
3' Untranslated Region
Linker sequence
Reverse primer
Stop codon
C
Promoter P(LacZ') 1
pTZ18U vector DNA
Forward Primer
LacZ alpha fragment
CaMV35S; 35S promoter from CaMV
Reverse Primer
35S WM5
5198 bp
Ampicillin Resistance
f1 ori
LacZ
WM5 cDNA ORF
Hind III restriction fragment 2.34kb
right border T-DNA repeat
Nos terminator
D
bla gene promoter
Ampicillin Resistance
pBluescript KS(+) vector DNA
Nos Terminator
uidA
WM5 GUS
7701 bp
Reverse primer 1
Forward primer 2
Forward primer 1
Forward primer 2
Intron
Eco RI restriction fragment 4.4kb
WM5 2.3kb promoter
E
bla gene promoter
Ampicillin Resistance
WM5 GFP
6350 bp
WM5 1.4kb promoter
Nos terminator
Eco RV/Sac I restriction fragment 3.45kb
HSP70 intron
GFP second exon
potato ST-LS1 intron
Forward Primer
Reverse Primer
GFP first exon
F
bla gene promoter
pBluescript SK(+) vector DNA
Ampiciliin Resistance
pAct1-F/Npt II
5247 bp
Rice Actin Promoter 2.1kb
Forward Primer
Nos Terminator
NptII
Reverse Primer

Table 11: Experimental design for co-bombardment of transgene constructs into wheat scutellum.

	Co-Bombardment Class		
	X	Y	Z
Constructs	Antisense *WM5* *WM5* GUS *WM5* GFP *pAct1-F/Npt II*	Sense *WM5* *WM5* GUS *WM5* GFP *pAct1-F/Npt II*	35S *WM5* *WM5* GUS *WM5* GFP *pAct1-F/Npt II*

Table 12: Parameters used in microprojectile bombardment of scutellar tissue in wheat.

Parameters	
Distance between A: rupture disk and microcarrier B: macrocarrier and stopping screen C: stopping screen and target cells	 2.5 cm 0.8 cm 5.5 cm
Gas pressure	900-1550 psi
Partial vacuum	27 inch Hg
Particles	Gold, 0.4-1.2 μm
Particle amount per bombardment	29 μg

Immature caryopses were then surface sterilised (1% sodium hypochlorite, 0.5% Mucasol) for 20 min and washed three times with sterile distilled water. Between 20 and 30 embryos of approximately 0.8 to 1.5 mm in length (**Figure 28**- A), were aseptically excised from the caryopses and placed clustered in the centre of a 6 cm petri dish with the embryo axis embedded in the solidified MS^- media (Murashige & Skoog, 1962; Nehra *et al.*, 1994). Generally, embryos were chosen that were translucent or deposition of starch in the scutellar tissue had initiated (visually semitransparent).

Immature embryos were cultured on modified MS^- medium. MS^- media (1 Litre) was made using 200 mL MS $Macro^-$ (10x), 2 mL MS Micro (1000x), 4 mL FeNaEDTA (500x) and 60 g sucrose, pH 5.7. All media was prepared in a double concentrated form, sterilised by filtration and mixed with the same volume of Gelrite (Roth, Germany) to give a final concentration of 0.3% Gelrite (Becker, 1999).

Embryo's were pre-cultured at 26°C for 2-3 days in the dark on callus induction medium, which was a modified MS^- medium containing 2 mg L^{-1} of 2,4-D (2,4-dichlorophenoxy acetic acid). Embryos were transferred to a high osmotic potential (0.7 M Sucrose) modified MS^- medium containing 2 mg L^{-1} of 2,4-D three to four hrs before microprojectile bombardment. Embryos were left on this media for 20-24 hrs post-bombardment and then transferred to callus induction medium.

For somatic embryogenesis induction, MS^- medium was modified by the addition of 2 mg L^{-1} of 2,4-D. Somatic embryogenesis induction was carried out at 26°C in the dark for 14 days without antibiotic selection, and a further 14 days with the selection agent (150 mg L^{-1} kanamycin).

Figure 28: *In vitro* culture of wheat explants.

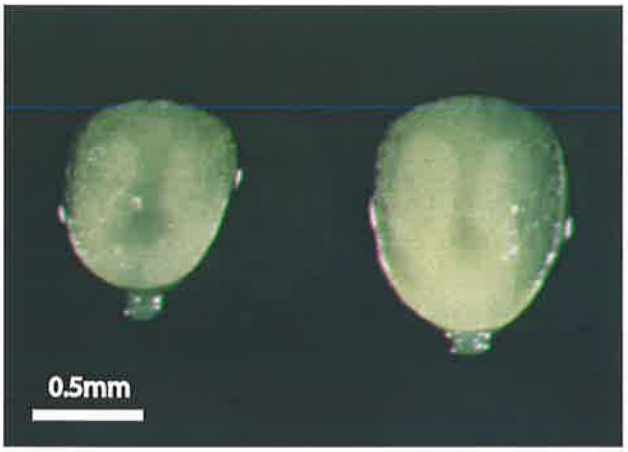
A - Size range of immature embryos dissected 10-12 days-post-anthesis for use in biolistic transformation of scutellar tissue.

B - Wheat calli showing induced shoot development with (1-selected) and without (2-unselected) kanamycin selection after 14 days on MS^- media containing 0.1 mg L^{-1} of 2,4-D and 150 mg L^{-1} of Kanamycin.

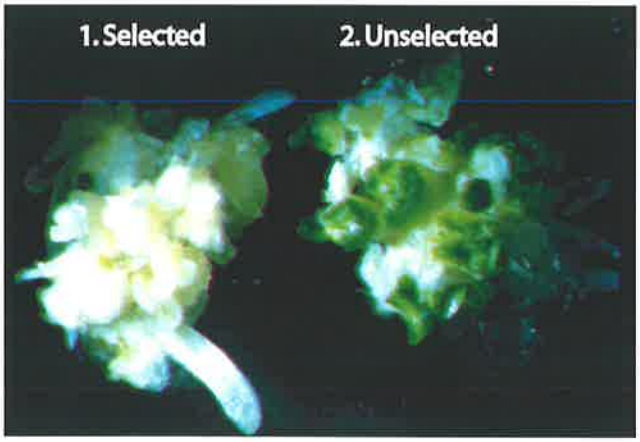
C - Regenerated plantlets from each of the co-bombardment classes grown on solidified ½ MS^- medium (0.4% agarose) without phytohormones for 2 weeks. Control- regenerated unselected shoot material, X, Y and Z – regenerated selected shoot material from each of the co-bombardment classes (**Table 11**).

D- Regenerants from the selection experiments (R_0) depicting the selection phenotype with A- unselected and B- selected by application of an aqueous solution of 2.5% Kanamycin containing 0.2% Tween 20.

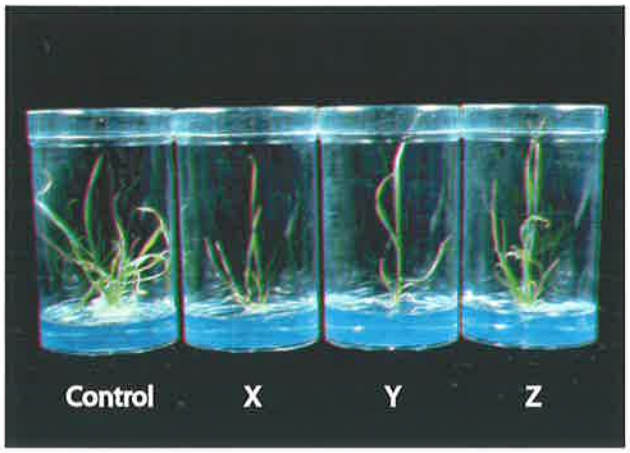
A



B



C



D

Plant regeneration was performed at 26°C under fluorescent light (3000 lx for 16 hr). Calli were transferred to MS^- medium containing 0.1 mg L^{-1} of 2,4-D and 150 mg L^{-1} of Kanamycin for 14 days to induce selectable shoot development as described in **Figure 29**. Calli were subcultured after 4 weeks if there was sufficient root initiation. Selection was continued during the plant regeneration phase. Once plantlets reached a height of approximately 1.5 cm, they were transferred to solidified ½ MS^- medium (0.3% Gelrite) without phytohormones (refer to **Figure 29**).

4.2.5 Antibiotic selection in glasshouse

After 3 weeks, developing plantlets were transferred to soil and placed in the glasshouse under growth conditions as described in Section 4.2.1. All regenerants from the selection experiments (R_0) were sprayed with an aqueous solution of 2.5% Kanamycin containing 0.2% Tween 20. One to two weeks later, the plants with at least one functional *Npt II* gene were normal and no damage could be observed, whereas the negative control plants (derived from non-bombarded cultured embryos) were bleached completely. The phenotype of the selected regenerants is depicted in **Figure 29**. Plantlets surviving selection were grown until maturity in a glasshouse according to conditions outlined in Section 4.2.1.

Ten seeds from each of the R_0 lines were planted in the glasshouse and fresh leaf material was collected for DNA extraction. However, for lines 7.X, 8.X, 9.X, 12.Y, 22.Y and 30.Z fewer than 10 seeds were available due to partial sterility.

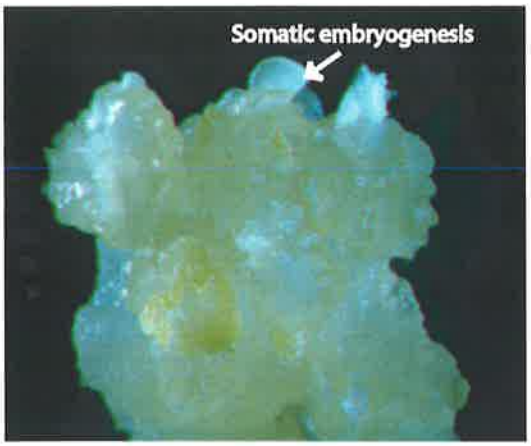
4.2.6 Histochemical GUS staining

β-glucuronidase (GUS) activity was determined histochemically using the following buffer: 0.1 mM sodium phosphate buffer, pH 7.0, containing 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 10 mM ethylenediaminetetraacetic acid, disodium salt (EDTA) and 0.1% (v/v) Triton X-100.

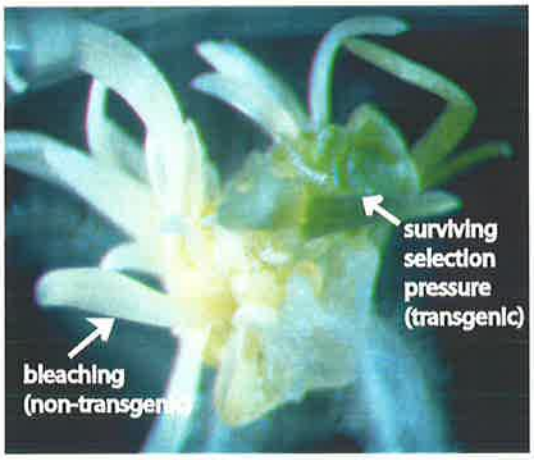
Figure 29: *In vitro* culture of wheat calli: A- 2 week old callus on callus induction media. B- 14 days of somatic embryogenesis induction of callus on modified MS^- media containing 2,4-D. C- Green shoots developing from embryogenic calli under 150 mg L^{-1} selection of kanamycin. D- Chimeric regenerated shoot under 150 mg L^{-1} selection of kanamycin.

A

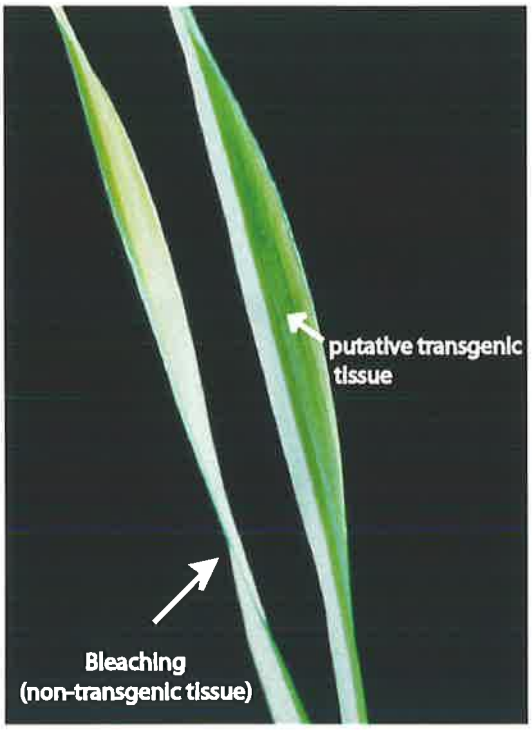
B


Somatic embryogenesis

C


surviving
selection
pressure
(transgenic)
bleaching

D


putative transgenic
tissue
Bleaching
(non-transgenic tissue)

In this buffer, 1 mg mL^{-1} 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-gluc, Nalgene, USA) was dissolved. The mixture was filter-sterilized and stored at –20°C.

Spikes from R_1 plants, containing anthers microscopically determined to be at the developmental stage of meiosis, were assayed for GUS activity. Spikes were aseptically dissected and vacuum infiltrated with staining buffer for 18-24 hrs at 37°C. Chlorophyll was subsequently extracted by incubating the spikes in a solution of 70% ethanol for 48 hrs at room temperature.

4.2.7 Microscopic detection of GFP fluorescence

GFP fluorescence of individual organs or tissue sections was monitored under a fluorescence phase contrast microscope (Zeiss Axioscope; Carl Zeiss, Jena, Germany) after excitation with light of 460- to 500-nm wavelengths. Emitted fluorescence was photographed on Kodak Extachrome 400 film, using a filter for the detection of fluorescence light at wavelengths longer than 510 nm. Anthers of GFP-expressing wheat plants were viewed with an MRC 1000 confocal laser scanning microscope (BIORAD) using blue laser excitation light (488 nm).

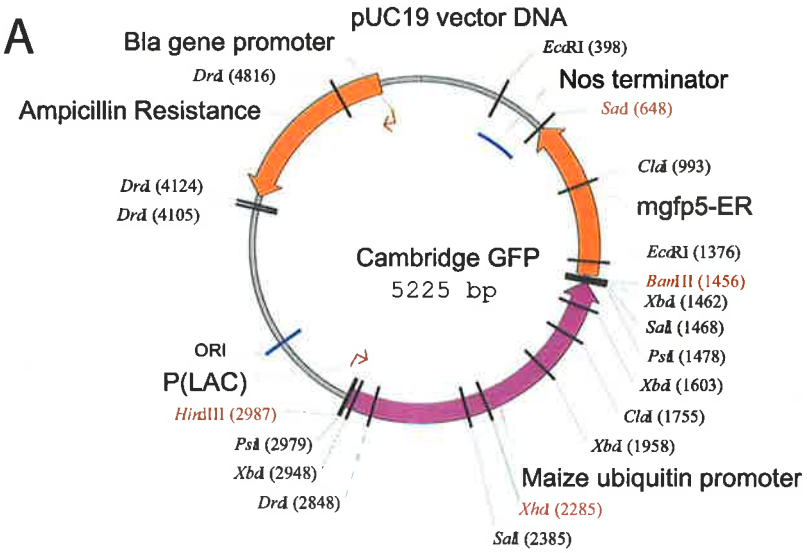
4.2.7.1 GFP *reporter gene in monocots*

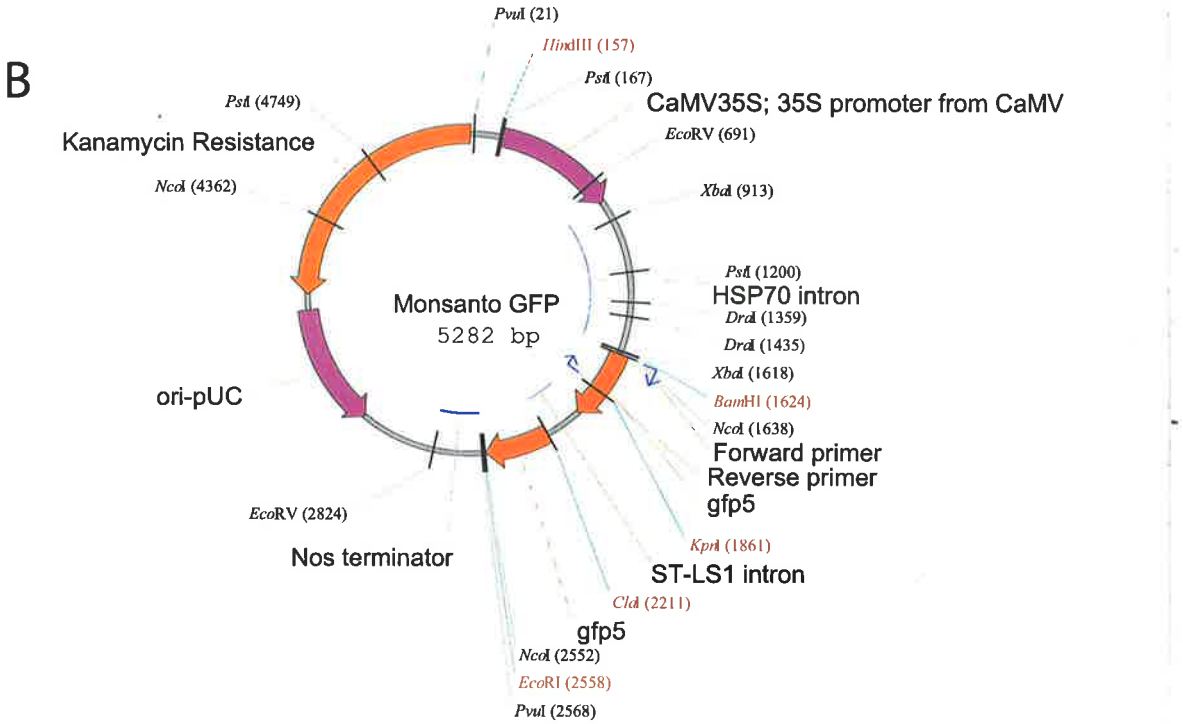
Intensity of GFP fluorescence in a transient assay on microprojectile bombarded wheat scutella (refer to Sections 4.2.3) was measured to determine differences between three different *GFP* constructs (**Figure 30**) modified for plant expression. Twenty-four hrs post-bombardment, immature scutella were visualised for GFP fluorescence.

4.2.8 PCR analysis

DNA extraction and general polymerase chain reaction protocols used in identifying transgenic wheat lines from the R_0 plantlets are outlined in Sections 2.2.4.1 and 2.2.9. PCR primers for each construct used in biolistic transformation are described in **Table 13**. Thermal cycling conditions for specific amplicons are outlined below:

Figure 30: Diagramatic representation of the constructs used in the transient bombardment assay. A - Cambridge *GFP* from *pBIN* m-gfp5-ER (Haseloff *et al.*, 1997), B- Monsanto *GFP* from *pMON30049* (Pang *et al.*, 1996) and C- Cambia *GFP* from *pCAMBIA-1302* (Hajdukiewicz *et al.*, 1994).


A
pUC19 vector DNA
Bla gene promoter
Ampicillin Resistance
DraI (4816)
DraI (4124)
DraI (4105)
EcoRI (398)
Nos terminator
SacI (648)
ClaI (993)
mgfp5-ER
EcoRI (1376)
BamHI (1456)
XbaI (1462)
SalI (1468)
PstI (1478)
XbaI (1603)
ClaI (1755)
XbaI (1958)
Maize ubiquitin promoter
XhoI (2285)
SalI (2385)
Cambridge GFP
5225 bp
ORI
P(LAC)
HindIII (2987)
PstI (2979)
XbaI (2948)
DraI (2848)


B
PvuI (21)
HindIII (157)
PstI (167)
CaMV35S; 35S promoter from CaMV
EcoRV (691)
XbaI (913)
PstI (1200)
HSP70 intron
DraI (1359)
DraI (1435)
XbaI (1618)
BamHI (1624)
NcoI (1638)
Forward primer
Reverse primer
gfp5
KpnI (1861)
ST-LS1 intron
ClaI (2211)
gfp5
NcoI (2552)
EcoRI (2558)
PvuI (2568)
PstI (4749)
Kanamycin Resistance
NcoI (4362)
Monsanto GFP
5282 bp
ori-pUC
EcoRV (2824)
Nos terminator

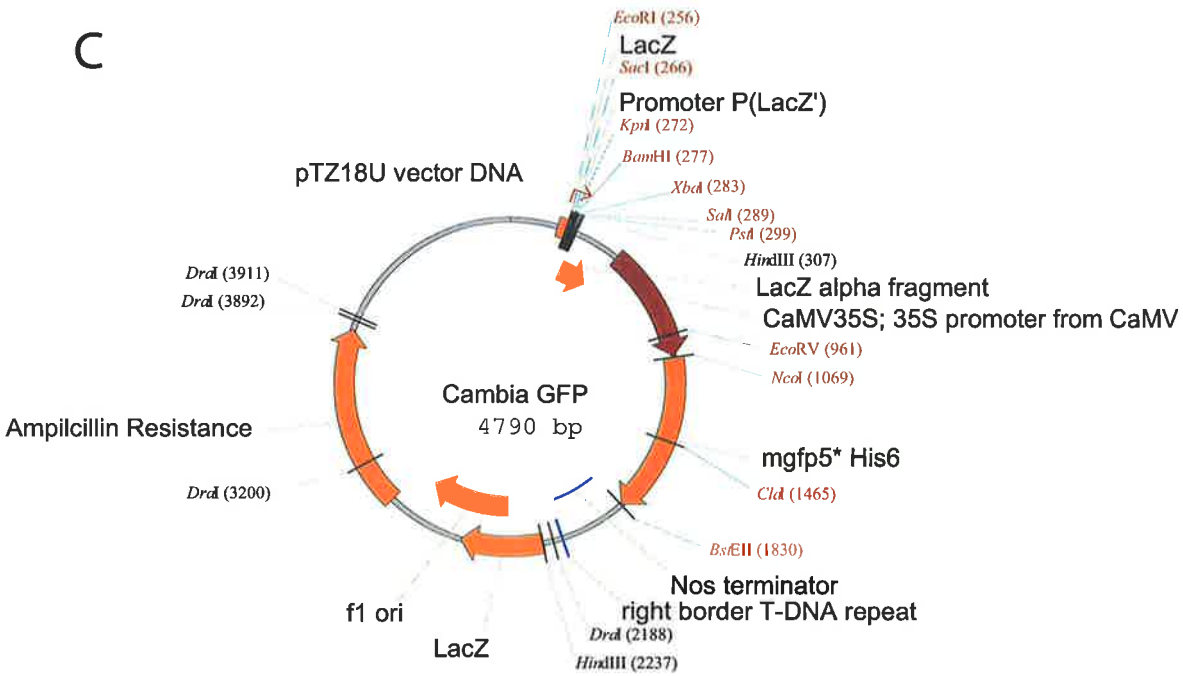

C
EcoRI (256)
LacZ
SacI (266)
Promoter P(LacZ')
KpnI (272)
BamHI (277)
XbaI (283)
SalI (289)
PstI (299)
HindIII (307)
LacZ alpha fragment
CaMV35S; 35S promoter from CaMV
EcoRV (961)
NcoI (1069)
mgfp5* His6
ClaI (1465)
BstEII (1830)
Nos terminator
right border T-DNA repeat
DraI (2188)
HindIII (2237)
pTZ18U vector DNA
DraI (3911)
DraI (3892)
Ampilcillin Resistance
DraI (3200)
Cambia GFP
4790 bp
f1 ori
LacZ

Table 13: PCR primers designed for the amplification of antisense *WM5* (p*WM53.1*), sense *WM5* (p*WM54*), 35S *WM5* (p*WM55.6*), *WM5* GUS (p*WM51*) and *WM5* GFP (p*WM52.1*), and pAct1-F/*Npt II.*

Amplicon	Forward Primer	Reverse Primer
Antisense *WM5* (1.2 Kb)	5'-CGCGCAATGAATGCCCTGGCCAA-3'	5'-CGACGGACAAGGCAAGCCAAGAT-3'
Sense *WM5* (1.5 Kb)	5'-CGCGCAATGAATGCCCTGGCCAA-3'	5'-GGGTCAGGGCTTCGGATCCACCTTGCGCCACG-3'
35S *WM5* (476 bp)	5'-AGAGGACCTAACAGAACTCGC-3'	5'-TAGAGGAAGGGTCTTGCGAAG-3'
WM5 GUS (637 bp) Primer set 1	5'-CAGGAAGTGATGGAGCATCAG-3'	5'-TCGTGCACCATCAGCACGTTA-3'
WM5 GUS (517 bp) Primer set 2	5'-CTGTAGAAACCCCAACCCGTG-3'	5'-CATTACGCTGCGATGGATCCC-3'
WM5 GFP (219 bp)	5'-GCAAGGGCGAGGAACTGTT-3'	5'-GGTACCGGGAGAAGCACTGAA-3'
pAct1-F/Npt II (700 bp)	5'-GAGGCTATTCGGCTATGACTG-3'	5'-ATCGGGAGCGGCGATACCGTA-3'

NB. Refer also to **Figure 30** for relative primer positions

Both antisense and sense *WM5* amplicons were generated by thermocycling 35 times on a MJR thermal cycler. Each cycle consisted of 1 min at 94°C for denaturation, 2 mins at 60°C for annealing and 2 mins at 72°C for extension.

GUS, *CaMV* 35S, and *Npt II* amplicons were generated under the same conditions as for antisense and sense *WM5* amplicons with the exception of a 62°C annealing temperature.

GFP amplicons were generated under the same conditions as for antisense and sense *WM5* amplicons with the exception of a 57°C annealing temperature.

4.2.9 Southern analysis

DNA extraction, Southern hybridisation and autoradiography protocols used in identifying transgenic wheat lines from the R_0 plantlets are outlined in Sections 2.2.9 through to 2.2.11. Southern analysis was used to estimate the copy number and whether there was whole construct insertion for each of the individual constructs (**Figure 27**) used bombardment experiments. All restriction endonucleases used for digestion of genomic DNA from T_0 and T_1 plantlets are given in **Table 14**.

4.2.10 Northern analysis

RNA extraction, Northern hybridisation and autoradiography protocols used in identifying transgenic wheat lines expressing transgenes are outlined in Sections 2.2.13 through to 2.2.11. Northern analysis was used to determine expression pattern for each of the individual constructs (**Figure 27**) used in the bombardment experiments.

4.3 Results

4.3.1 Plant regeneration and selection of transformants

Isolated scutella (5832) of wheat varieties Veery, Combi, Pavon and Florida were subjected to microprojectile bombardment with constructs of co-bombardment classes X, Y and Z (experimental details are outlined in **Table 19** - Appendix).

Table 14: Restriction endonucleases used for digestion of genomic DNA from T_0 and T_1 plantlets to estimate both copy number and whole transgene integration. Expected sizes of fragments yielded by enzymatic digestion are shown in parenthesis.

	Restriction Endonucleases	
Construct	Copy Number	Whole Transgene Integration
Antisense *WM5*	*Hind III*	*Bam HI* (1.22 Kb)
Sense *WM5*	*Hind III*	*Bam HI* (0.88 Kb and 0.74 Kb)
35S *WM5*	*Eco RI*	*Hind III* (2.34 Kb)
WM5 GUS	*Hind III*	*Eco RI* (4.4 Kb)
WM5 GFP	*Hind III*	*Eco RV* and *Sac I* (3.45 Kb)
pAct1-F/Npt II	*Eco RV*	Undetermined

NB. Refer also to **Figure 27** for relative primer positions

4.3.2 T_0 and T_1 analysis

4.3.2.1 PCR analysis of T_0 and T_1

The majority of transgenic plants identified by PCR amplification (**Figure 31**) were of transgenes isolated from genomic DNA of plantlets (totalling 30) that survived kanamycin selection in culture and in the glasshouse (**Table 15**). Twenty-three of the 30 regenerant lines possessed the expected 700 bp *Npt II* PCR product whilst 15 and 20 of the regenerant lines possessed the 219 bp and 637 bp PCR products for the *GFP* and *GUS* reporter genes respectively. Antisense, Sense and 35S *WM5* PCR amplicons (1.2 Kb, 1.5 Kb and 476 bp respectively) were detected in 3, 5 and 7 of the 30 regenerant lines respectively. Four of the 30 regenerent lines were shown to be non-transgenic.

Those plants identified as transgenic (T_0) were grown in the glasshouse and evaluated morphologically. All lines were observed to be morphologically similar and on further analysis, no lines were observed to be totally sterile. The majority (77%) of the transformed plants produced as many seeds as the seed-derived control plants whilst (23%) showed partial sterility.

4.3.2.2 *Southern analysis of T_0 and T_1*

Each of the 26 PCR identified transgenic T_0 wheat lines, plus the 4 lines not confirmed transgenic by PCR, were also analysed for stable integration of respective transgenes based on co-bombardment class via Southern hybridisation (**Table 15**). Southern analysis also included the T_1 lines to determine the heritability of the transgene(s). The same genomic DNA used in the PCR assay was used in the Southern analysis.

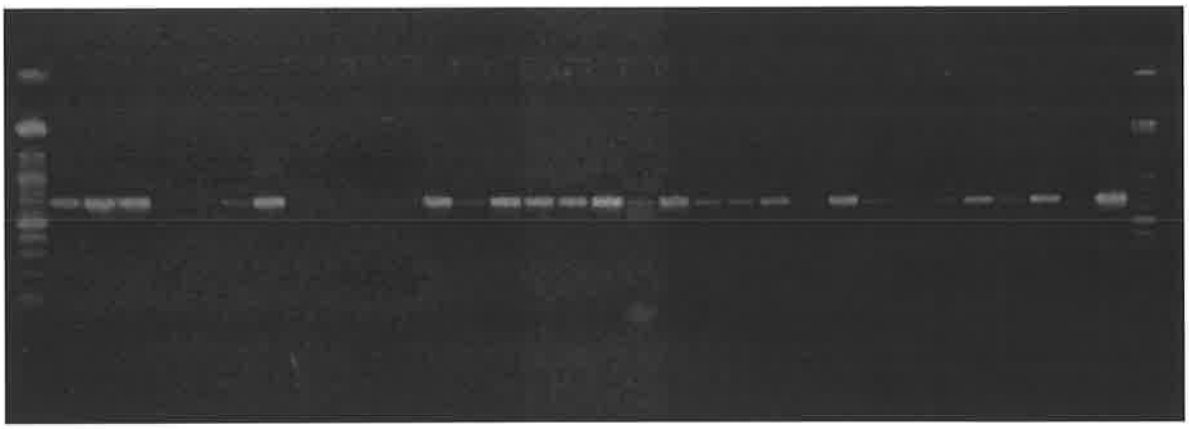
Figure 31: PCR analysis of T_0 wheat lines. Genomic DNA isolated from T_0 plants was amplified for *Npt II* selectable marker (A), *GUS* reporter (B), *GFP* reporter (C), Antisense *WM5* (D), Sense *WM5* (E) and 35S promoter (F) transgene constructs. M-1 Kb ladder, 1-30 T_0 plant labels, -ve – negative control, +ve – positive plasmid control.

A

Npt II

M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 -ve +ve M

700bp →

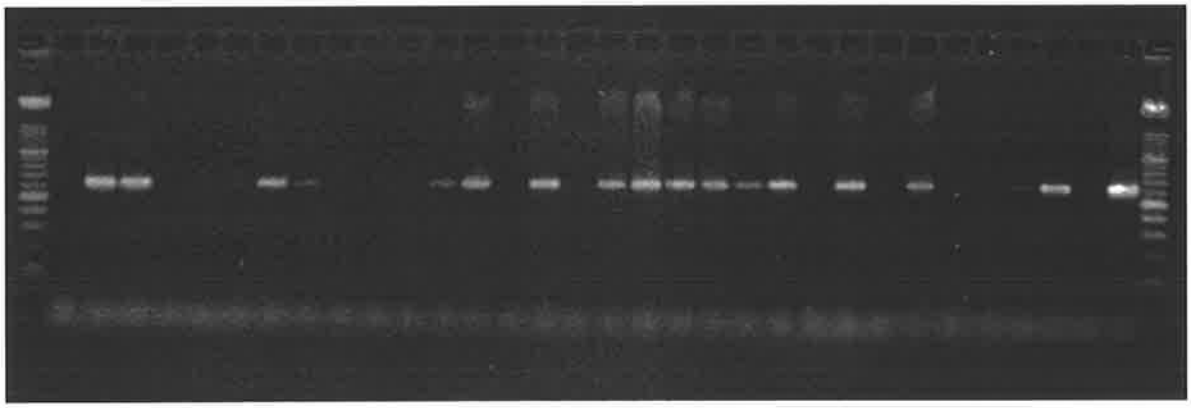


B

GUS

M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 -ve +ve M

637bp →

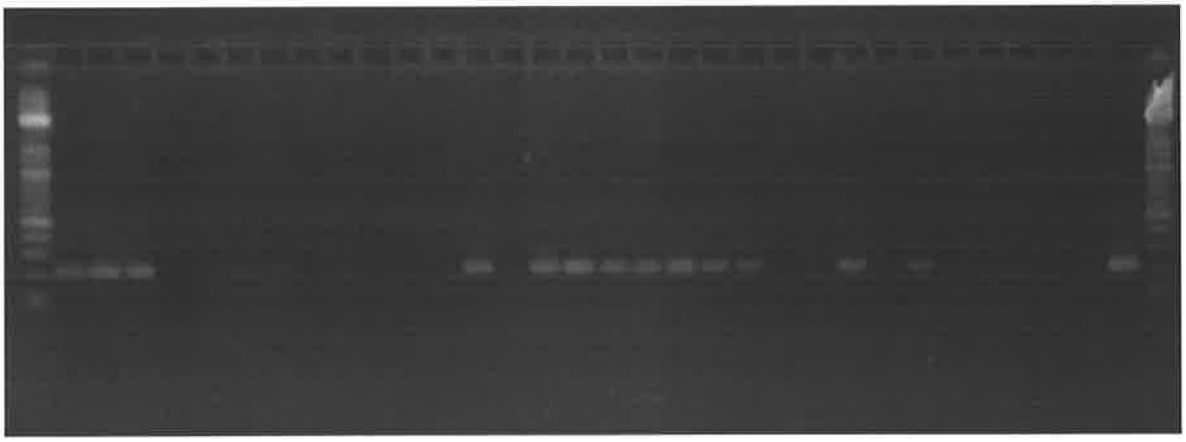


C

GFP

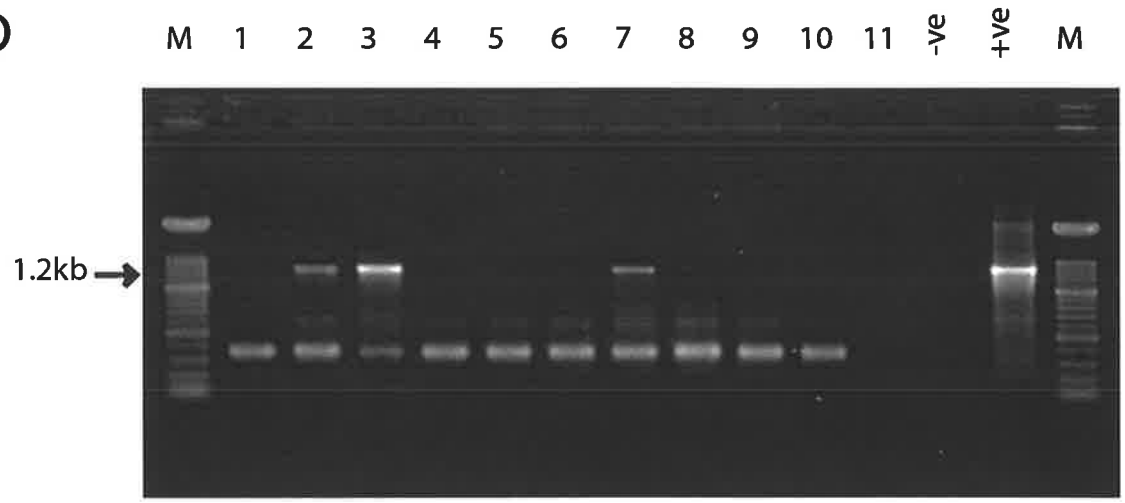
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 -ve +ve M

219bp →



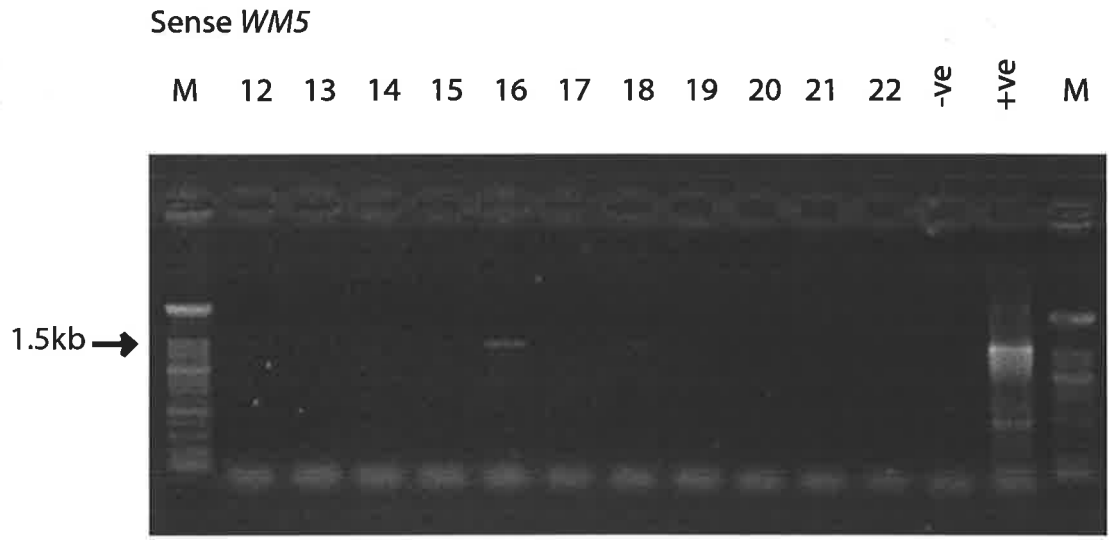
D

Antisense *WM5*



E

Sense *WM5*



F

35S *WM5*

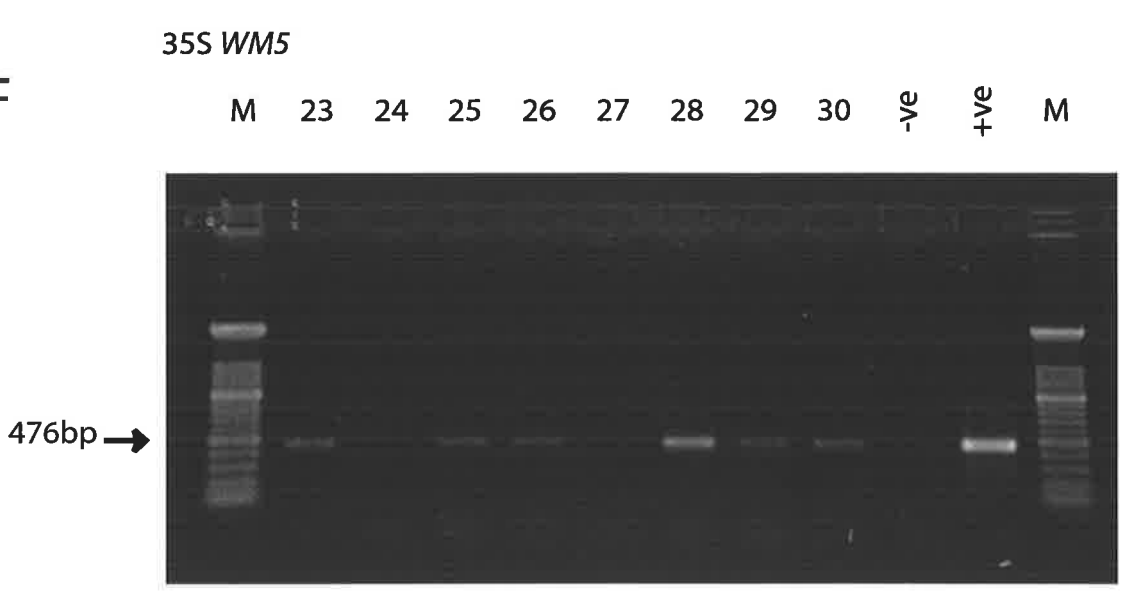


Table 15: Kanamycin selected wheat regenerants identified as being transgenic through both PCR and Southern analysis. Regenerant wheat lines are labelled 1 – 30 with a suffix of X, Y or Z depending on their co-bombardment class. "+" represents the presence of transgene, whilst "-" represents the absence of transgene.

Co-bombardment class	Variety	Bombardment Date	Plant Label	*pAct1-F/Npt II*	*WM5* GFP	*WM5* GUS	Sense *WM5*	Antisense *WM5*	35S *WM5*
Group X	Veery	26/4/99	1.X	+	+	-		-	
	Veery	06/03/95	2.X	+	+	+		+	
	Veery	16/4/99	3.X	+	+	+		+	
	Veery	31/3/99		-		-		-	
	Veery	14/5/99		-		-		-	
	Combi	18/3/99	6.X	+	+	+		-	
	Combi	22/4/99	7.X	+	+	+		+	
	Combi	22/4/99	8.X	-	-	+		-	
	Florida	26/4/99	9.X	-	-	+		-	
	Florida	26/4/99	10.X	-	-	-		-	
	Florida	26/4/99	11.X	-	-	-		-	
Group Y	Veery	18/3/99	12.Y	+	-	+	+		
	Veery	05/03/95	13.Y	+	+	+	-		
	Veery	05/03/95	14.Y	+	-	+	-		
	Veery	05/03/95	15.Y	+	+	+	-		
	Veery	09/03/95	16.Y	+	+	+	+		
	Veery	16/4/99	17.Y	+	+	+	-		
	Veery	16/4/99	18.Y	+	+	+	+		
	Veery	26/4/99	19.Y	+	+	+	+		
	Combi	13/4/99	20.Y	+	+	+	+		
	Combi	22/4/99	21.Y	+	+	+	-		
	Pavon	22/3/99	22.Y	+	-	+	-		
Group Z	Veery	26/3/99	23.Z	-	-	-			+
	Veery	16/4/99	24.Z	+	+	+			+
	Veery	16/4/99	25.Z	+	-	-			+
	Veery	29/4/99	26.Z	+	+	+			+
	Veery	03/04/95	27.Z	+	-	-			-
	Pavon	22/3/99	28.Z	+	-	-			+
	Pavon	13/3/99	29.Z	+	-	-			+
	Pavon	18/3/99	30.Z	+	-	+			+

NB. Grey overshadow are those wheat lines that were not transformed.

Figure 32 shows an example (transgenic wheat line 3.X) of the Southern analysis that was performed on all 26 transgenic lines to estimate copy number and whether there was whole transgene integration. The estimation of transgene copy number and the verification of whole transgene integration for each of the Southern confirmed transgenic lines are displayed in **Table 14**.

The pattern and intensity of hybridisation varied greatly among all the transformants and this reflects variation in the number of inserted copies of the gene (**Table 14**).

4.3.3 Transformation efficiency

Both the PCR and Southern analysis confirmed the production of 26 independent transformation events giving a transformation frequency of 0.45%. The transgenic plants produced from each of the individual microprojectile bombardment experiments are outlined in **Table 14**. Results show that there was a co-transformation frequency of 81%. Five of the 26 transformed lines had integration of a single construct type; a further 5 had integrated two different constructs; 10 had integration of three constructs whilst integration of all four different construct types occurred in only 6 of the 26 wheat lines.

4.3.4 Northern analysis

Northern analysis was used to determine the expression pattern, in the segregating T_1 transgenic wheat lines, of the *pAct1-F/Npt II*, *WM5* GFP, and *WM5* GUS transgenes. **Figure 33** shows that all wheat lines except 26.Z have a segregating pattern of *pAct1-F/Npt II* expression. The level of expression of *pAct1-F/Npt II* varied between lines and within the T_1 progeny of individual transgenic lines (17.Y and 24.Z). Analysis of *GFP* expression (**Figure 33**) showed that all lines had some level of transgene expression except lines 19.Y and 26.Z.

Table 16: Inheritance of transgenes from T_0 to 10 random T_1 progeny as identified through both PCR and Southern analysis. Transgenic lines are labelled 1 – 30 with a suffix of X, Y or Z depending on their co-bombardment class. Red boxes represent those individual wheat lines that contain *Npt II* selectable marker, green boxes the *GFP* reporter gene construct, blue boxes the *GUS* reporter gene construct, yellow boxes the Antisense *WM5* construct, purple boxes the Sense *WM5* construct and orange boxes represent those lines possessing the 35S *WM5* construct. Numbers within the boxes highlight the number of copies for each construct respectively. Light grey boxes represent lines that possess no transgenes whilst dark grey boxes represent those wheat lines with no seed available.

Co-bombardment	Variety	Plant Label	T_0	T_1 1	2	3	4	5	6	7	8	9	10
Group X	Veery	1.X											
	Veery	2.X											
	Veery	3.X											
	Combi	6.X											
	Combi	7.X											
	Combi	8.X											
	Florida	9.X											
Group Y	Veery	12.Y											
	Veery	13.Y											
	Veery	14.Y											
	Veery	15.Y											
	Veery	16.Y											
	Veery	17.Y											
	Veery	18.Y											
	Veery	19.Y											
	Combi	20.Y											
	Combi	21.Y											
	Pavon	22.Y											
Group Z	Veery	23.Z											
	Veery	24.Z											
	Veery	25.Z											
	Veery	26.Z											
	Veery	27.Z											
	Pavon	28.Z											
	Pavon	29.Z											
	Pavon	30.Z											

Figure 32: Example of Southern analysis performed on all T_0 and T_1 wheat lines. Southern analysis of T_0 wheat line 3.X including all T_1 progeny. All Southern probes were PCR amplified from each construct as per **Table 13** and **Figure 31**.

A – *GUS* probe on *Hind III* digested genomic DNA to estimate copy number of the *WM5 GUS* constuct.

B – *GUS* probe on *Eco RI* digested genomic DNA to determine whole integration of the *WM5 GUS* construct (expected 4.4 Kb restriction fragment).

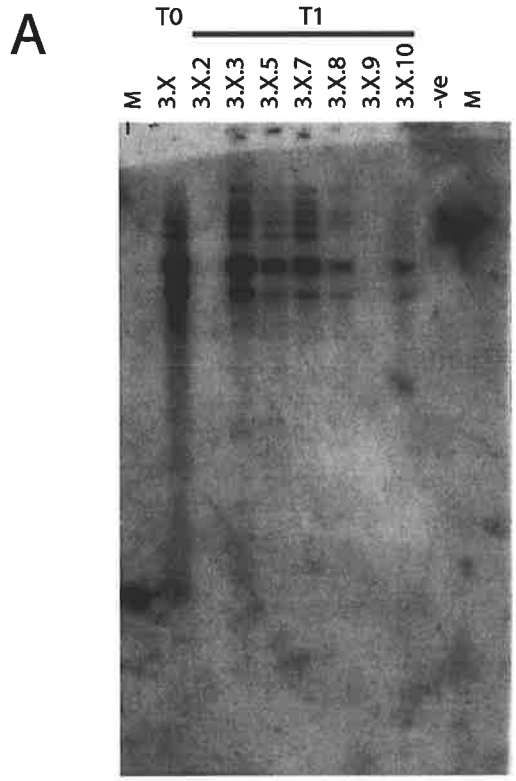
C – *WM5* probe on *Bam HI* digested genomic DNA to determine whole integration of the *Antisense WM5* construct (expected 1.22 Kb restriction fragment).

D – *GFP* probe on *Eco RV/Sac I* digested genomic DNA to determine whole integration of the *WM5 GFP* construct (expected 3.45 Kb restriction fragment).

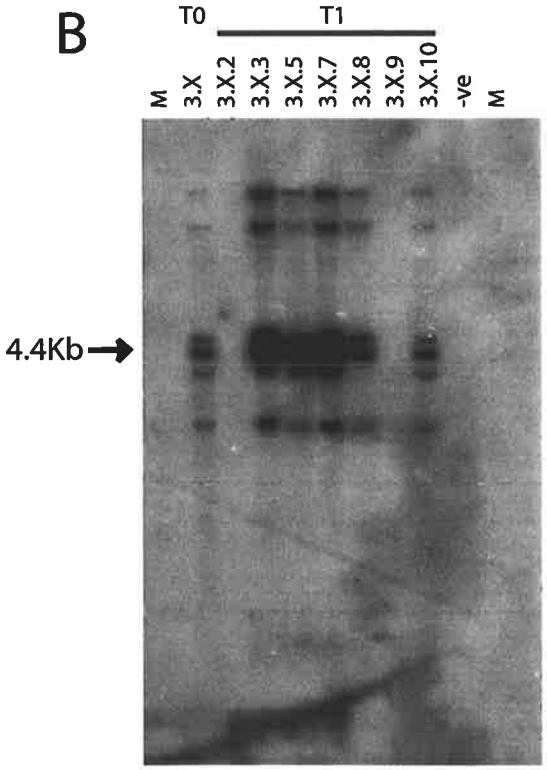
E – *Npt II* probe on *Eco RI* digested genomic DNA to estimate copy number of the *pAct1-F/Npt II* constuct

For each construct refer to restriction maps in **Figure 27**. Lanes M refers to λ DNA digested with *Bst EII* and *Sac I*. -ve refers to the negative control (untransformed regenerated wheat DNA).

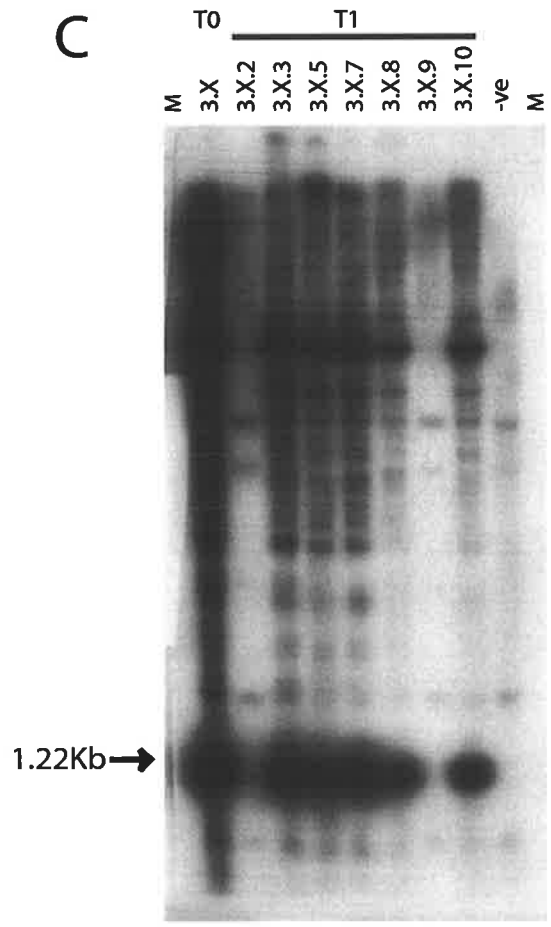
WM5 GUS Hind III



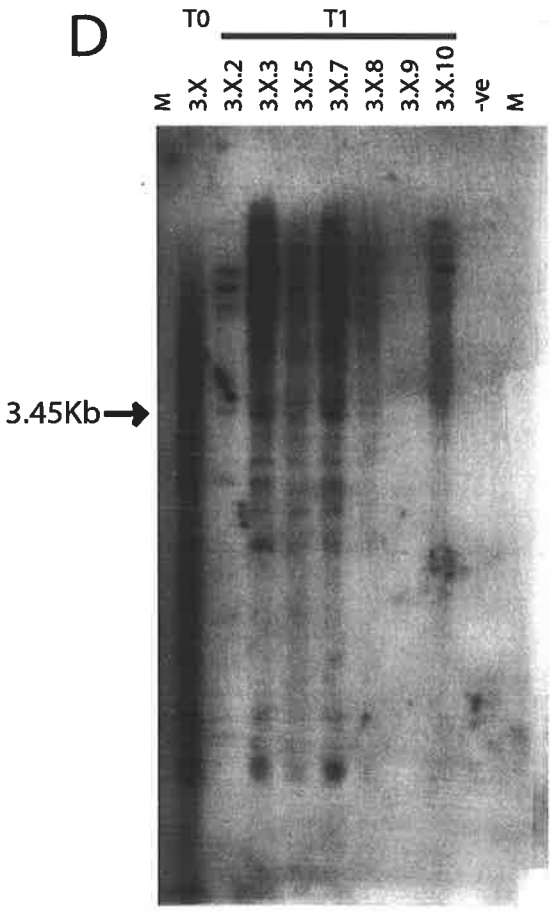
WM5 GUS Eco RI



Antisense WM5 Bam HI



WM5 GFP Eco RV/Sac I



E

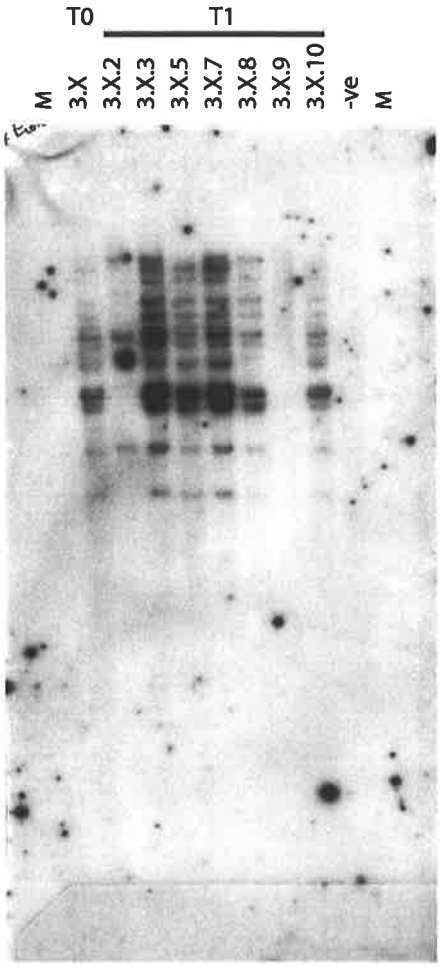
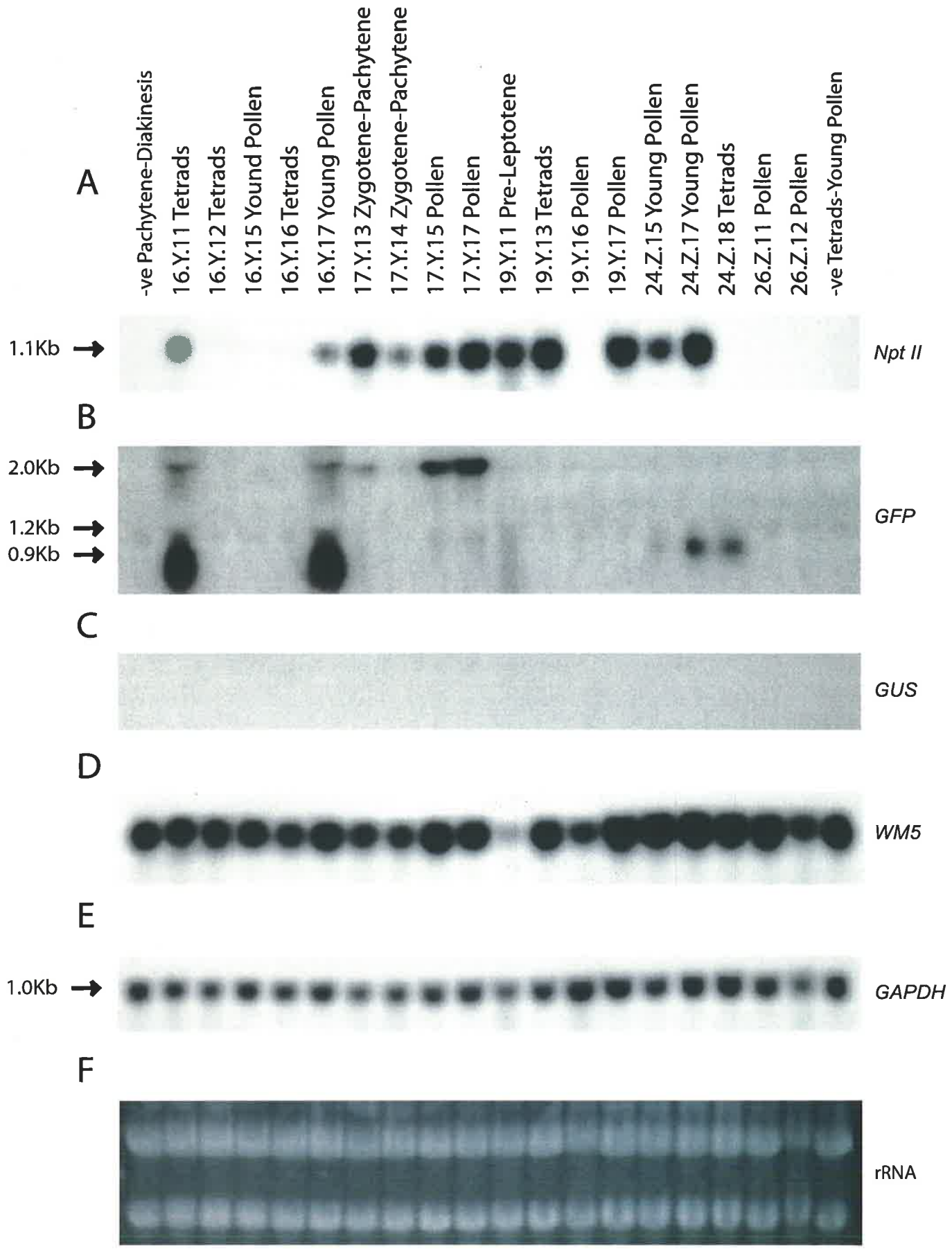


Figure 33: Northern analysis showing expression pattern of A- *Npt II*, B- *GFP* C- *GUS* (no expression), D- *WM5* and E- Glyceraldehyde-6-phosphate dehydrogenase (*GAPDH* - loading control), F- Ethidium bromide stained ribosomal RNA in the segregating T_1 transgenic wheat lines. The transgenic wheat lines analysed include 16.Y, 17.Y, 19.Y, 24.Z and 26.Z.


A
-ve Pachytene-Diakinesis
16.Y.11 Tetrads
16.Y.12 Tetrads
16.Y.15 Yound Pollen
16.Y.16 Tetrads
16.Y.17 Young Pollen
17.Y.13 Zygotene-Pachytene
17.Y.14 Zygotene-Pachytene
17.Y.15 Pollen
17.Y.17 Pollen
19.Y.11 Pre-Leptotene
19.Y.13 Tetrads
19.Y.16 Pollen
19.Y.17 Pollen
24.Z.15 Young Pollen
24.Z.17 Young Pollen
24.Z.18 Tetrads
26.Z.11 Pollen
26.Z.12 Pollen
-ve Tetrads-Young Pollen
1.1Kb
Npt II
B
2.0Kb
1.2Kb
0.9Kb
GFP
C
GUS
D
WM5
E
1.0Kb
GAPDH
F
rRNA

Furthermore Northern analysis of *WM5* showed variation in the levels of expression within the different developmental stages of the spike analysed (**Figure 33**). As for *Npt II* expression, *GFP* expression varied between lines (16.Y versus 24.Z) and within the T_1 progeny of individual transgenic lines (17.Y). *GFP* expression was only detected in those lines where the RNA was extracted from spike tissue that was older than the tetrad stage of pollen mother cell development. It is also interesting that **Figure 33** suggests that not all transgenic lines process the *GFP* transcript the same way, as line 16.Y produced an abundance of a 0.9 Kb transcript whilst line 17.Y seemed to lack a 0.9 Kb transcript and instead contained a 2 Kb transcript. This suggests that line 16.Y was able to correctly splice the *GFP* transcript whilst line 17.Y was unable to splice the two introns (refer to **Figure 27**) from the primary transcript (2 Kb).

Interestingly, no *GUS* expression was detected in the Northern analysis of the transgenic wheat lines 16.Y, 17.Y, 19.Y, 24.Z and 26.Z (**Figure 33**).

4.3.5 WM5 promoter analysis

4.3.5.1 *Analysis of the* GFP *reporter gene in monocots*

Transient expression of three modified plant *GFP* constructs in wheat scutella was analysed to determine the most suitable modified form of *GFP* for *WM5* GFP reporter construct construction.

The construct showing highest level of fluorescence in wheat scutella, given that all transformation and microscopy parameters were kept the same, was Monsanto's *GFP* (refer to **Figure 34**). Consequently this version of *GFP* was used in the construction of the *WM5* GFP reporter construct.

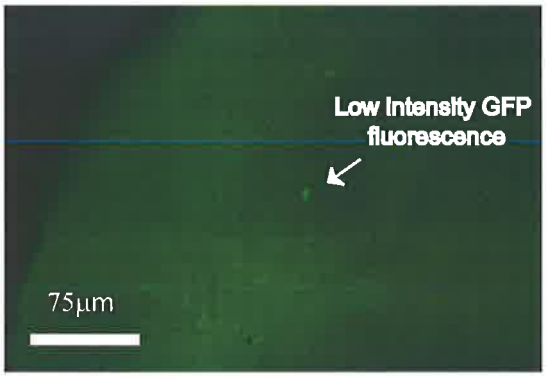
4.3.5.2 WM5 *GFP reporter analysis*

According to **Table 14** 15 of the 26 T_0 wheat lines possessed the *WM5* GFP reporter construct. Fluorescent microscopy of the T_1 progeny confirmed high levels of expression of GFP in at least 4 of 15 transgenic lines. Line 16.Y was chosen for further analysis.

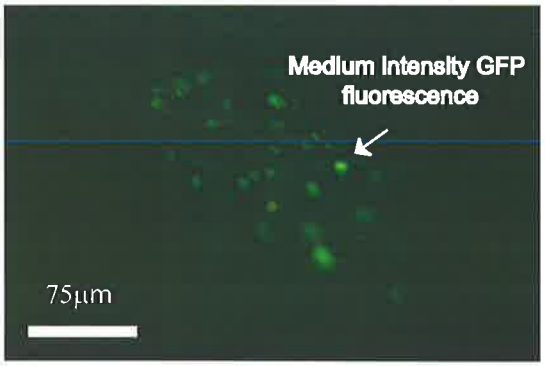
Figure 34: A. Transient expression assay of green fluorescent protein (GFP) in wheat scutella from 1- Cambridge GFP, 2- Monsanto GFP, 3- Cambia GFP and 4- *pBluescript SK(+)* control. B. Whole scutella image of transient expression assay of GFP from Monsanto GFP construct (refer to **Figure 30** for construct details).

A

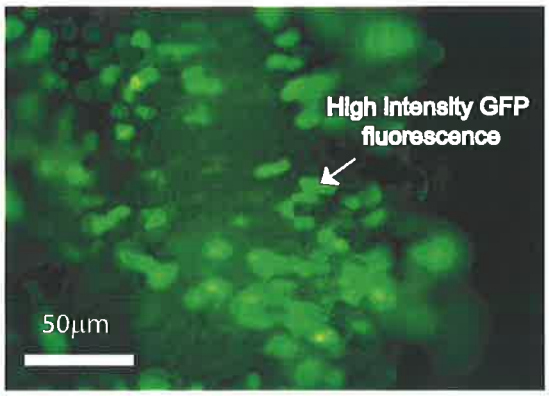
1


Low Intensity GFP
fluorescence
75μm

2


Medium Intensity GFP
fluorescence
75μm

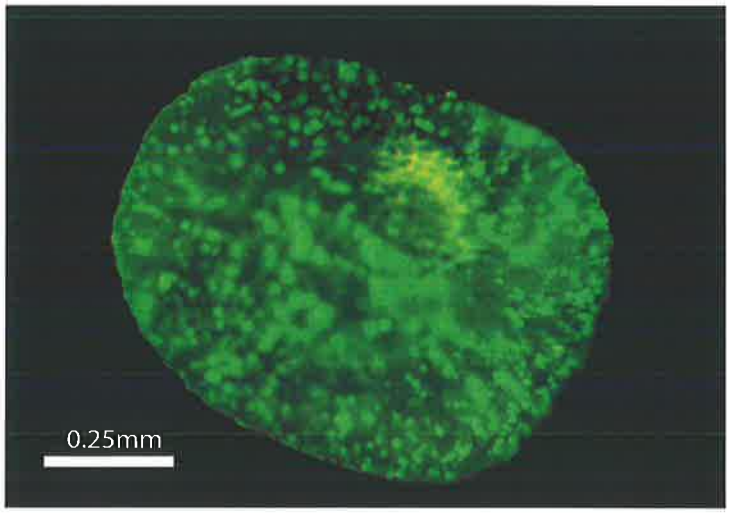
3


High Intensity GFP
fluorescence
50μm

4


No GFP
fluorescence
50μm

B


0.25mm

Fluorescence of GFP was primarily visualised in immature pollen grains (binucleate stage) from the developing anther (**Figure 35**) utilising a fluorescent dissecting microscope. Laser confocal microscopy allowed sectioning of the anther with the visualisation of GFP fluorescence. Confocal microscopy not only showed GFP fluorescence in the developing pollen, but also in the tapetal cells of the anther (**Figure 35**). Temporally, GFP fluorescence initiated at the late tetrad stage and persisted through at low levels until pollen maturation (data not shown). Fluorescent microscopy was unable to detect GFP fluorescence in female meiocytes, roots, leaves or mitotic apices.

4.3.5.3 WM5 GUS *reporter analysis*

According to **Table 14**, 20 of the 26 T_0 wheat lines possessed the *WM5* GUS reporter construct. Histochemical GUS staining of pre- to post-meiotic spikes of the T_1 progeny confirmed expression of GUS in at least 2 of 28 transgenic lines. Line 24.Z and 30.Z were chosen for further analysis.

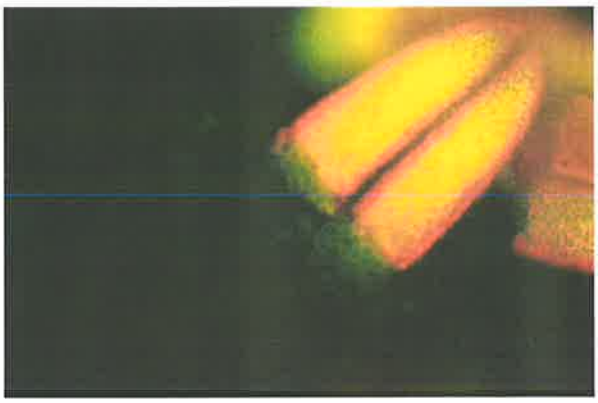
Strongest GUS expression was visualised in longitudinal strips on each of the lobes of pre-meiotic anthers (**Figure 36**). Low levels of unspecific GUS expression were detected on cut edges of all tissues examined. No GUS expression was detected in female meiocytes, roots, leaves or mitotic apices.

4.3.6 *WM5* phenotypic analysis

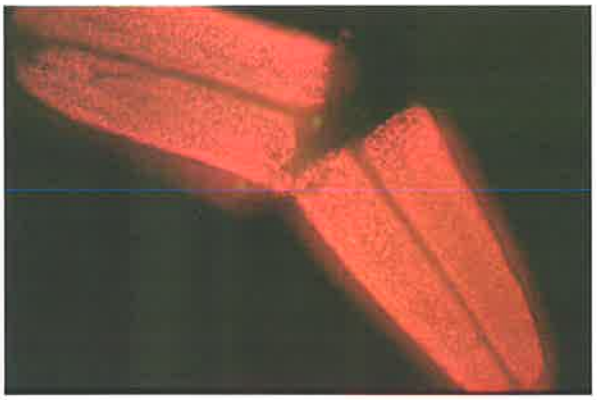
Each of the 26 analysed transgenic wheat lines were compared to the control wheat lines for significant differences in fertility, and general plant morphology. No significant differences were detected between the controls and the transgenic wheat lines.

Figure 35: Fluorescent microscopy for GFP fluorescence in the T_1 progeny of transgenic line 16.Y. A- Fluorescence of GFP in immature pollen grains from the developing anther of 16.Y.5; B- Autofluorescence from a developing anther containing immature pollen grains from Veery control; C- Laser confocal micrographs of GFP fluorescence emitted from 1) immature pollen and 2) tapetal cells (highlighted by arrows) from within a developing anther of 16.Y.5; D- Laser confocal micrographs of autofluorescence emitted from 1) immature pollen and 2) tapetal cells (highlighted by arrows) from within a developing anther of Veery control.

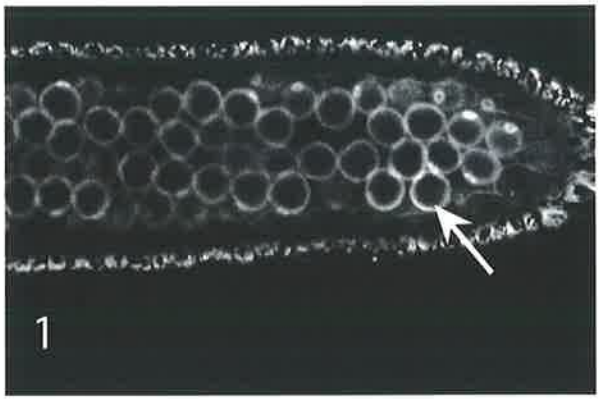
A

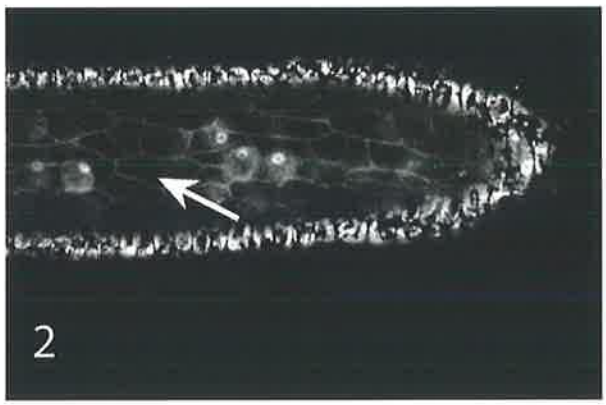


B

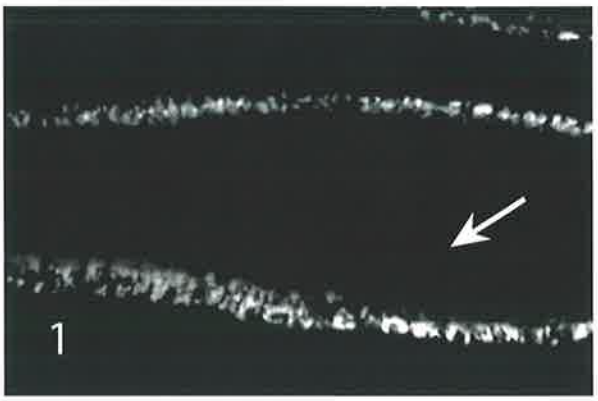


C





D



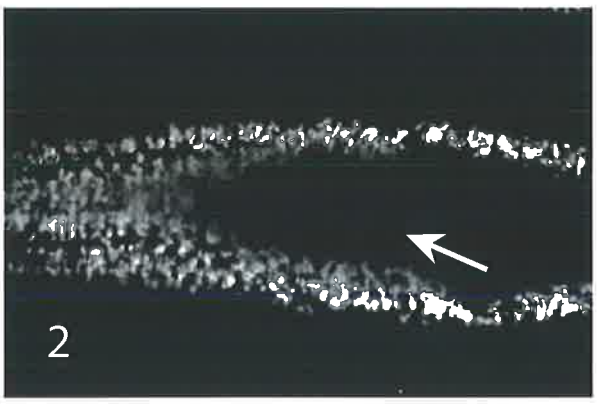
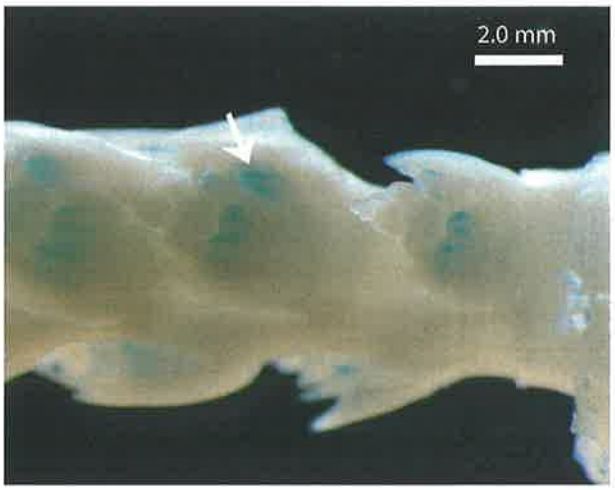
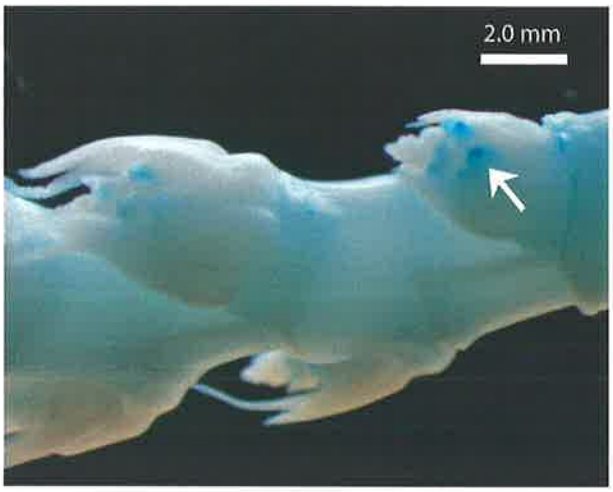


Figure 36: GUS expression of pre-meiotic spikes from T_1 progeny of 24.Z and 30.Z. A- GUS expression in anthers of transgenic wheat line 30.Z.6; B- and C- GUS expression in anthers of transgenic wheat line 24.Z.2 (highlighted by arrows); D- GUS expression absent in anthers of Veery control wheat line.

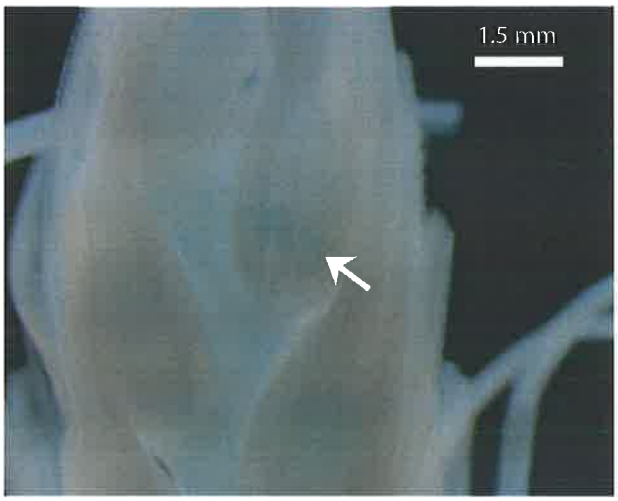
A



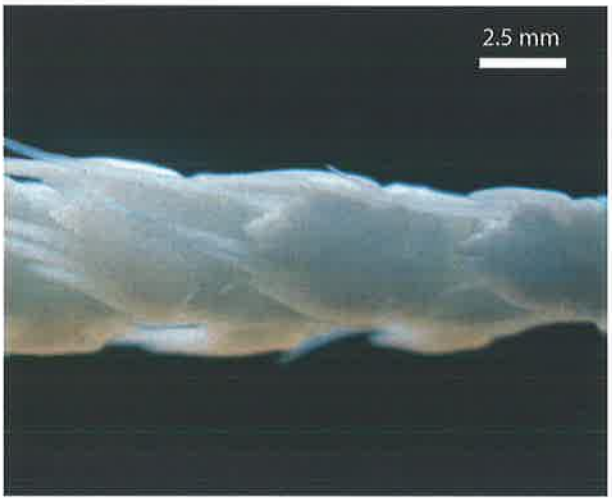
B



C



D



4.4 Discussion

This chapter describes the use of genetic transformation of wheat for functional analysis of meiotic genes.

4.4.1 Plant regeneration and selection of transformants

Of the 26 transgenic wheat lines produced, no significant phenotypic abnormalities were observed. Reduced fertility was observed in 6 of the transgenic lines but shows that it is similar to reports of the occurrence of sterility in transgenic maize, and wheat plants (Fromm *et al.*, 1990; Gordon-Kamm *et al.*, 1990; Vasil *et al.*, 1992) produced through the regeneration of both embryogenic callus and morphogenic suspension cultures. These reports attributed sterility to long-term cultures and not to the transformation procedure *per se*. However, transformation of wheat via microparticle bombardment of scutellar tissue generally results in low levels of sterility (Becker *et al.*, 1994; Nehra *et al.*, 1994) given that the culture period before embryogenesis is short (2-3 weeks).

Selection of transformed wheat tissues showed that kanamycin (150 mg L^{-1}) was an effective selection agent. Observations suggested that without two glasshouse selections with a spray solution of kanamycin (2.5% w/v) there would have been significant numbers of escapes. This is unfavourable when compared to use of the *bar* gene as a selectable marker, as fewer plantlets escape from tissue culture selection (Christou *et al.*, 1991; Gordon-Kamm *et al.*, 1990).

4.4.2 T_0 and T_1 analysis

4.4.2.1 PCR analysis of T_0 and T_1

The use of PCR to ascertain the T_0 transgenic lines is relatively quick and easy. However, during this study it was observed that some transgenic lines were not detected as a result of difficult PCR amplicons (eg. sense *WM5* construct). This does not exclude the use of PCR if rigorous conditions for the reproducible amplification of transgene constructs can be determined. However it is essential that PCR results be subsequently confirmed through Southern hybridisation.

4.4.2.2 Southern analysis of T_0 and T_1

Microprojectile bombardment commonly results in complex integration patterns (Birch & Franks, 1991). The most practical and informative genetic engineering would involve the integration of a single transgene copy to avoid potential problems of co-suppression. However multiple gene integration does not necessarily preclude lower levels of gene expression. This study generated transgenic wheat lines with both multiple transgene integration and single transgene integration. Thirty-five percent of lines contained all four different constructs (**Table 11**).

4.4.3 Transformation efficiency

Utilising microprojectile bombardment as a transformation procedure, PCR and Southern analysis confirmed transgenic wheat plants were generated at a frequency of 0.45% (refer to **Table 19** - Appendix). This is comparable to the 0.5-2.5% transformation frequency that has previously been reported in wheat (Becker *et al.*, 1994; Nehra *et al.*, 1994). Summarising the results, one transgenic wheat plant is recovered for every 224 embryos bombarded. This estimation of stable transformation frequency is substantially lower than that determined for the immature embryos of rice (one every 27 embryos (Christou *et al.*, 1991)) and maize (one every 100 embryos (Koziel *et al.*, 1993)).

Co-transformation frequency was at 81% showing that microprojectile bombardment with multiple constructs is an efficient method of delivery for more than one gene. These results are comparible to the cotransformation frequency (67-79%) seen in sugarcane callus (Bower *et al.*, 1996). Preliminary transgene segregation results based on Southern analysis of a limited number of T_1 progeny (ten plants) suggest that all three constructs were integrated at one or two independent loci within the wheat genome. It should be noted that larger T_1 and T_2 progeny populations should be analysed to confirm these preliminary conclusions. Further analysis of the T_2 progeny is likely to provide further details on the number of integration events of the transgene(s) into the wheat genome.

4.4.4 Northern analysis

This study has also demonstrated the transcriptional activity of the integrated transgenes. Northern analysis has demonstrated expression of both the *Npt II* selectable marker and *GFP*, under the transcriptional control of the 1.4 Kb *WM5* promoter fragment. Expression levels of *Npt II* and *GFP* varied between T_0 transgenic lines and within the segregating T_1 progeny.

As *WM5* is highly expressed in pollen mother cells at meiosis, it was expected that the *WM5* promoter (1.4 Kb fragment) driven GFP expression would also be seen at this developmental stage. This was not observed.

The developmental process of meiosis takes 24 hrs to complete in wheat (Bennett *et al.*, 1973) and is difficult to observe with limited numbers of tillers per plant. Limited spikes of T_1 plants from previously confirmed *pAct1-F/Npt II*, *WM5* GFP and *WM5* GUS T_0 lines were available for Northern analysis. GFP expression was detected in only 3 of the 5 analysed transgenic lines. Northern analysis suggests that the 1.4 Kb promoter fragment of *WM5* drives GFP expression from tetrad stage of development. These results are not definitive as the number of spikes analysed at varying stages was limited. Compounding this, the T_1 lines analysed were segregating for the integrated transgenes. Timing did not allow analysis of more homozygous T_2 transgenic lines.

Not all transgenic lines seemed to express GFP in the same manner. Results suggest that line 16.Y was able to correctly splice the *GFP* transcript whilst line 17.Y was unable to splice the two introns from the primary transcript (2 Kb) (**Figure 33**). However this is not the only possible explanation of the observed results. The observed banding pattern (**Figure 33**) could possibly be an abberation in the coding sequence due to multiple transgene integrations disrupting the true coding sequence for *GFP*. This is a possibility given that transgenic line 17.Y has a single copy of each of the *WM5* GFP *pAct1-F/Npt II*, and *WM5* GUS constructs present in the genome (**Table 14**).

WM5 showed variation in expression pattern in different spikes analysed (**Figure 33**). This variation is not directly attributable to the transgenes

integrated as the endogenous *WM5* gene could not be differentiated from the *WM5* transgenes in this hybridisation study. Variation in *WM5* expression was probably due to the variation known to occur in the endogenous *WM5* gene during floral development. Further analysis of these transgenic lines with transgene specific RT-PCR may help determine if both the endogenous *WM5* and the *WM5* transgenes have an altered expression level.

No *GUS* expression was detected through Northern analysis of 5 transgenic wheat lines (**Figure 33**). This result is likely to be due to either the lack of functional copies of the *WM5* GUS transgene or that the tissues in which the *GUS* transcript would be highly expressed are not represented in the RNA samples. This does not discount that *GUS* could be expressed in these tissues but is not detectable via Northern analysis. Further analysis through RT-PCR could confirm expression of the *GUS* gene in these T_1 progeny.

4.4.5 *WM5* promoter analysis

4.4.5.1 *Analysis of the* GFP *reporter gene in monocots*

GFP is a 21 kDa apo-protein that spontaneously folds into a bright-green fluorescing structure. GFP was chosen as a reporter gene in these experiments as it can be expressed and visualised in many cell types, both animal and plant. In spite of successful transient expression of the wild-type GFP in plant protoplasts (Hu & Cheng, 1995; Niedz *et al.*, 1995; Sheen *et al.*, 1995) and from virus-based vectors (Baulcombe *et al.*, 1995; Casper & Holt, 1996; Oparka *et al.*, 1995), its expression in stably transformed plants has typically yielded very faint or no green fluorescence.

Experiments were designed to test the level of fluorescence of GFP in a transient assay of three "plant-modified" *GFP* constructs in wheat scutella. Results highlighted that the construct showing highest level of fluorescence in wheat scutella, was Monsanto's GFP (refer to **Figure 34**) and was comparable with that seen in maize protoplasts and tobacco plants (Chiu *et al.*, 1996). The high level of fluorescence has been attributed to the removal of a cryptic splice site (Haseloff *et al.*, 1997) and altered codon usage (higher G/C content) than

used for human proteins. A mutation in which the serine at position 65 is modified to a threonine (S65T) plus the introduction of both the potato ST-LS1 intron into the coding region of *GFP* and the heat shock protein 70 intron before the start codon have all shown to yield higher levels of fluorescence in monocots.

4.4.5.2 WM5 GFP *reporter analysis*

Laser confocal microscopy of GFP expressing transgenic line 16.Y, showed GFP fluorescence primarily in immature pollen grains (binucleate stage) from the developing anther (**Figure 35**). These results support the observed presence of the *GFP* transcript only in those RNA samples from spikes cytologically staged to be at the tetrad stage of meiotic development (**Figure 33**). Confocal microscopy sectioning not only showed GFP fluorescence in the developing pollen, but also in the tapetal cells of the anther (**Figure 35**). GFP fluorescence was not detected in female meiocytes, roots, leaves or mitotic apices. These results are not strictly comparable to the expected expression pattern of *WM5*. Northern analysis (Dong *et al.*, 2001a) has shown the *WM5* transcript to be strongly expressed during early meiosis with expression gradually reducing as meiosis proceeds. Low levels of expression have also been detected in shoot apices. The *WM5* gene is not expressed in highly active mitotic tissue such as the root. Additionally the WM5 protein has been immuno-localised to pollen mother cells, tapetal cells and floral tissue epidermal cells (Dong *et al.*, 2001a).

Spatially GFP fluorescence was visualised in the same cells where *WM5* is known to be highly expressed (pollen mother cells and tapetal cells). However temporally, the delay of GFP expression as compared to *WM5* suggests that the 1.4 Kb fragment 5' to the *WM5* coding sequence does not contain the whole *WM5* promoter. This 1.4 Kb fragment does possess the ability to drive expression, just not at the early stages of meiosis, as expected.

The use of the *GFP* reporter gene and laser confocal microscopy illustrates the usefulness of the green fluorescent protein in determining the spatial expression pattern for genes of interest.

4.4.5.3 WM5 GUS *reporter analysis*

Histochemical GUS expression in T_1 progeny in pre- to post- meiotic spikes confirmed a similar spatial expression to GFP. GUS expression was observed in longitudinal strips on each of the lobes of pre-meiotic anthers (**Figure 36**). No GUS expression was detected in female meiocytes, roots, leaves or mitotic apices. Temporal expression of GUS in the pollen mother cells conflicted with both results of GFP fluorescence and the expected *WM5* expression pattern. GUS is predominantly expressed in pre-meiotic anthers; native *WM5* is predominantly expressed in meiotic anthers whilst GFP is predominantly expressed in post-meiotic anthers. Differences in the observed expression of both the GFP and GUS protein can possibly be attributed to the lengths of the promoter fragments transcriptionally controlling each of these genes. The *WM5* GUS possesses a 2.5 Kb fragment 5' to the *WM5* coding sequence whilst the *WM5* GFP possesses a smaller 1.4 Kb fragment 5' to the *WM5* coding sequence. Time constraints and cloning difficulties prevented the construction of *GUS* and *GFP* reporter constructs possessing the same size *WM5* promoter fragment. Nevertheless, analysis of the reporter genes from both the *WM5* GUS and *WM5* GFP transgenic lines suggest that neither the 2.5 Kb or the 1.4 Kb 5' fragments represent the whole native *WM5* promoter. Varying temporal but highly similar spatial expression patterns in reporter gene transgenic lines suggest that *WM5* is under complex transcriptional control.

Given the small number of transgenic lines showing GUS and GFP expression the impact of transgene intergration site could also be influencing observed variation in expression patterns. For correct transcriptional activity of reporter genes, as defined by the endogenous *WM5* expression, the site of integration could prove vitally important. The transgene may have to be integrated into a region of the genome that would confer meiotic expression. This could correspond to the *Ph2* region.

4.4.6 *WM5* phenotypic analysis

Of the 26 transgenic wheat lines analysed no significant phenotypic differences were seen in the T_1 progeny as compared to the controls. Further phenotypic analysis on homozygous T_2 plants needs to be performed on those transgenic lines confirmed to be expressing each of Antisense *WM5*, Sense *WM5* and 35S *WM5*. Specifically, the analysis of *WM5*'s potential influence on a *Ph2*-like phenotype, cytological examination of bivalent formation at metaphase I, needs to be examined. Additionally, homozygous T_2 lines should be crossed to *Triticum kotschyi* var. *variablis* (Sears, 1982) to examine their ability for interspecific chromosome pairing. However, these lines also need to be studied in detail with respect to apical meristematic development. Antisense *WM5* transgenic wheat lines may also show altered development of the protodermal L1 layer as has been shown with the *PDF1* gene of *Arabidopsis thaliana* (Abe *et al.*, 1999; Abe *et al.*, 2001). Such detailed analysis of the wheat transformants produced in this study may help decifer the function of the *WM5* gene.

5 Chapter

STRUCTURE OF THE *PH2* LOCUS

5.1 Introduction

Hexaploid wheat behaves strictly as a diploid during meiosis, which results in 21 bivalents at metaphase I. This phenotype is controlled by a major locus on the long arm of chromosome 5B (Okamoto, 1957; Riley & Chapman, 1958) termed *Ph1*. Given the physiological complexity of meiotic processes such as chromosome pairing and recombination, it is foreseeable that *Ph1* is not the only locus controlling wheat's strictly diploid behaviour. The *Ph2* locus, although a minor suppressor of homeologous chromosome pairing in comparison to *Ph1*, indicates that chromosome pairing is a multigenic trait. A mutation at the *Ph2* locus termed *ph2a* was originally identified in an X-ray irradiated population of hexaploid wheat during an attempt to isolate mutants of the *Ph1* locus (Sears, 1977). Current knowledge of meiotic processes is largely based on detailed studies in yeast. A large number of genes have shown importance during meiosis in yeast including chromosome pairing and recombination (reviewed in Kleckner, 1996). Phenotypic studies of the *ph1b* mutant have shown that deletion of the *Ph1* locus may affect several premeiotic and meiotic processes (Feldman, 1993; Luo *et al.*, 1996; Shaw & Moore, 1998), and that both the *Ph1* and the *Ph2* loci are unlikely to be controlled by single genes (Roberts *et al.*, 1999).

This study aims to analyse the complexity of the *Ph2* locus; is this a single gene, a multigene family or a complex of linked genes influencing different meiotic processes? Utilising a similar approach in identifying the *Ph1* gene(s) (Roberts *et al.*, 1999), this study uses comparative mapping between rice and the *Triticeae* to investigate the complexity of the *Ph2* locus.

Comparative mapping has been demonstrated in various crop plants through the use of a common set of DNA probes (Bonierbale *et al.*, 1988). Genetic maps of the *Triticeae* (*T. aestivum*, *T. tauschii*, and *Hordeum* spp.) have been compared to maps of rice, maize and oat (Vandeynze *et al.*, 1995) and have shown molecular markers on the linkage maps for these species detected with the same probes are essentially homosequential. The *Gramineae* share extensive synteny across their genomes, allowing for one species to serve as the base for comparative genomics within the family (Moore *et al.*, 1995). Within the cereals, rice presents the most tractable species for genomic applications in a cereal. Small genome size and agricultural importance make rice an ideal model species for genome analysis. The diploid rice genome consists of 431 Mb per 1C nucleus whilst the hexaploid wheat genome consists of 15,966 Mb per 1C nucleus (Arumuganathan & Earle, 1991). In comparison to other cereals, rice contains half the DNA content of sorghum (760 Mb), 17% and 8.8% of the DNA content of maize (2, 504 Mb) and barley (4,873 Mb) respectively. The major difference in genome size is due to ploidy, however some difference can be attributed to wheat possessing approximately 80% repetitive sequences compared to 50% in rice (Bennett & Smith, 1991). The selection of rice as a model organism can also be ascribed to extensive genetic and molecular resources. High-density rice maps have recently allowed contig generation from BAC libraries in preparation for map based cloning (Yang *et al.*, 1997). Extensive comparative mapping between grasses (Vandeynze *et al.*, 1995) demonstrated that the genes present on rice chromosome 1 are in a similar order to those on wheat homoeogroup 3 (Smilde *et al.*, 2001). This makes rice the logical choice for molecular characterisation of the *Ph2* locus.

5.2 Materials and methods

5.2.1 Plant material

5.2.1.1 Genetic stocks

The following hexaploid wheat (*Triticum aestivum*; 2n = 6x =42; genome AABBDD) lines were used; seeds of euploid wheat (*Triticum aestivum* cv. Chinese Spring) were obtained from Dr Ken Shepherd, Waite Agricultural

Research Institute. The Chinese Spring mutant lines *ph2a* and *ph2b* were obtained from Prof. Moshe Feldman, Plant Genetics Institute, Israel.

5.2.1.2 Mapping populations

Three double haploid populations derived of barley were used for mapping. These three mapping populations consisted of 150 lines from crosses between the Australian cultivars Chebec x Harrington, Galleon x Haruna Nijo (Dr S. Logue, Waite Agricultural Research Institute) and Clipper x Sahara (Dr A. K. R. M. Islam, Waite Agricultural Research Institute).

5.2.2 RFLP analysis

A total of 26 RFLP markers were screened for their ability to identify polymorphisms between Chinese Spring and the meiotic mutants. These DNA clones were obtained through the Australian *Triticeae* Mapping Initiative and from IPK, Gaterslaben (Germany). The majority of these markers have been mapped in wheat and barley and their chromosomal locations identified (Marino *et al.*, 1996; Nelson *et al.*, 1995a; Nelson *et al.*, 1995b; Nelson *et al.*, 1995c; Vandeynze *et al.*, 1995). The clones were selected on the basis of their availability and map location. The *WM1* family was also mapped with respect to these clones in order to ascertain the relative position of the gene family on the chromosome arm and in position within the deletion.

5.2.2.1 Southern blot analysis

Hybond N+ membranes (Amersham, Australia) prepared from restriction endonuclease digested DNA from each of the double haploid plants (including parental lines), of the three barley mapping populations were kindly supplied by the Barley National Molecular Marker Program (Waite Institute, Adelaide). Protocols for the preparation of these membranes are outlined in Sections 2.2.9 to 2.2.12.

5.2.2.1.1 WM1, WM3, WM5 *and* TaMSH7 *gene family RFLP mapping*

PCR primers were designed based on partial genomic sequence for *WM1.1* (Ji, 1992; Ji & Langridge, 1994). The whole putative coding sequence for *WM 1.1* was amplified from Chinese Spring genomic DNA using primers *WM1.1* ORF in

Table 1, whilst the whole cDNA for Wheat Meiosis 3 (abbreviated *WM3* (Letarte, 1996)), *WM5* (Dong *et al.*, 2001a; Thomas, 1997) and *T. aestivum* mismatch repair protein homologue 7 (abbreviated *TaMSH7* (Dong *et al.*, 2001b)) were used as probes for Southern analysis. *WM1.1* ORF, *WM3*, *WM5*, and *TaMSH7* probes were screened across Chebec x Harrington, Clipper x Sahara and Galleon x Haruna Nijo barley mapping population membranes as supplied by the Barley National Molecular Marker Program. Mappable polymorphisms for *WM3* and *TaMSH7* were only detected in the Galleon x Haruna Nijo (Dong *et al.*, 2001b; Letarte, 1996), whilst for *WM5;* polymorphisms were only detected in the Clipper x Sahara mapping population (Dong *et al.*, 2001a).

5.2.2.1.2 RFLP probes for determining ph2a *deletion size*

Total genomic DNA of wheat lines Chinese Spring, *ph2a* and *ph2b* were digested with *Bam HI*, *Dra I*, *Eco RI*, *Eco RV* or *Hind III* according to protocols outlined in Section 2.2.10. RFLP probes obtained through the Australian *Triticeae* Mapping Initiative and from IPK, Gaterslaben (Germany) were hybridised to Southern blots of these wheat lines. A total of 14 probes were screened across these wheat lines (**Table 17**).

5.2.2.2 Genetic analysis

Data obtained from scoring segregation patterns of the *WM1* gene family, *WM5*, *TaMSH7* and *WM3* RFLP markers were analysed with MapManager QT (Version 8.0) software (Manly, 1993) using the Kosambi mapping function (Lander *et al.*, 1987).

5.2.3 Comparative mapping between barley, wheat and rice

Putative meiosis genes *WM5*, *TaMSH7*, *WM3* and the *WM1* gene family have been localised to the region deleted in the *ph2a* mutant. Using these markers in combination with RFLP probes (Australian *Triticeae* Mapping Initiative) localised both inside and outside the *ph2a* deletion, Graingenes database was screened for comparative genetic maps between the grasses. Utilising comparative anchor probes (Vandeynze *et al.*, 1995) that flanked the *ph2a* deletion region in wheat, the syntenous *Ph2* regions were determined in barley, wheat and rice.

5.2.3.1 YAC, PAC and BAC contig development

A total of 19 rice YACs located on rice chromosome 1 region comparable to the region deleted in *ph2a* on wheat chromosome 3DS were selected from the Rice Genome Program (Japan) to identify a partial physical map (http://rgp.dna.affrc.go.jp/). P1 artificial chromosome (PAC) and BAC clones were further identified that represented a partial physical map of the rice chromsome 1 region comparable to the *Ph2* locus. All clones were derived from rice variety Nipponbare. Sequence data for each of the 53 identified PAC/BAC clones was available through the Rice Genome Program. The RiceGAAS software (http://rgp.dna.affrc.go.jp/) was used for visualisation of the predicted coding sequences.

5.2.3.2 Identification of wheat ESTs from PAC and BAC clones

Each of the 53 identified PAC/BAC clones was used in a BLAST 2.0 blastn (Altschul *et al.*, 1997) search of the public EST database for wheat ESTs (International *Triticeae* EST Corporative) that had an expectation $\leq 1xE^{-50}$.

5.2.3.3 Structural analysis of putative meiosis genes

The RiceGAAS software (http://rgp.dna.affrc.go.jp/) was used for predicting the intron-exon structure of each of the rice predicted coding sequences. Putative cDNAs were translated into the primary peptide (VectorNTI Suite Version 6.0) and compared to various motif databases including the Conserved Domain Database v1.51 (Altschul *et al.*, 1997), INTERPRO (Apweiler *et al.*, 2001), and the PROSITE database (Hofmann *et al.*, 1999). Finally, annotations on DNA sequence and peptide sequence were compiled through the use of VectorNTI Suite Version 6.0 software (Informax, Inc. USA).

5.2.3.4 Analysis of rice and barley meiosis related phenotypic traits

RiceGenes and GrainGenes databases were screened for both mutant lines and Quantitative Trait Loci (QTLs) that correlate with phenotypes associated with meiosis. All phenotypic traits on the databases likely to be associated with meiosis were screened for their map localisation to respective *Ph2* homologous regions in rice and barley.

Table 17: Restriction fragment length polymorphism probes (molecular markers) for determining *ph2a* deletion size.

Molecular Marker	Source	Type	Repetition[1]	Location with respect to *ph2a* deletion
ABC171	Barley	cDNA	Lc	Outside
ABG471	Barley	cDNA	Lc	Outside
ABG396	Barley	Unknown	Unknown	Outside
ABG57	Barley	cDNA	Unknown	Outside
AWBMA15	Barley	cDNA	Sc	Outside
ABG460	Barley	cDNA	Lc	Undetermined
BCD15	Barley	cDNA	Lc	Outside
KSUA3	Wheat	Unknown	Unknown	Undetermined
PSR1316	Wheat	Unknown	Unknown	Outside
PSR929	Wheat	Unknown	Lc	Outside
PSR598	Wheat	cDNA	Unknown	Inside
PSR1196	Wheat	Unknown	Unknown	Inside
PTAG683	Wheat	Unknown	Lc	Outside
WG405	Wheat	Genomic	Lc	Outside

[1] Sc – single copy, Lc – low copy

5.3 Results

5.3.1 RFLP probes for determining *ph2a* deletion size

The homologous region corresponding to the region deleted in the *ph2a* mutant of wheat was identified in the three-barley mapping populations (Chebec x Harrington, Clipper x Sahara and Galleon x Haruna Nijo) using molecular markers (genomic DNA or cDNA clones – refer to **Table 17**) (**Figure 37**).

This region was also determined in barley using four different molecular markers for putative meiotic genes known to reside within the region deleted in the *ph2a* mutant. These molecular markers code for the *WM1* gene family (Ji & Langridge, 1994), *WM5* (Thomas, 1997), *TaMSH7* (Dong *et al.*, 2001b) and *WM3* (Letarte, 1996) genes. **Figure 38** shows that the molecular marker order in wheat, barley and rice (3DS, 3HS, rice chromosome 1, respectively) is syntenous across the region defined by the *ph2a* mutant in wheat. Putative wheat meiosis genes are ordered on rice chromosome 1 as follows; *WM1* gene family, *TaMSH7*, *WM5* and *WM3* with the *WM1* gene family located distal and *WM3* proximal to the centromere. This order is consistent for all four markers across three barley-mapping populations (**Figure 37**).

5.3.2 Comparative mapping between barley, wheat and rice

Utilising the molecular markers found to be located within the region deleted in the *ph2a* mutant, regions syntenous to *Ph2* regions were determined in both wheat and rice. The syntenous *Ph2* regions in rice, wheat and barley are outlined in **Figure 38**. The maximum estimated genetic size of the homologous region in barley that is deleted in the *ph2a* wheat mutant is 15 cM whilst the minimum genetic size is 6.5 cM. Exact size was not determined due to the lack of molecular markers close to the deletion breakpoints in the wheat *ph2a* mutant.

5.3.2.1 YAC, PAC and BAC physical map

Comparative mapping using grass anchor probes (Vandeynze *et al.*, 1995) highlighted 19 rice YACs located on the rice chromosome 1 region homologous to the region deleted in *ph2a* on wheat chromosome 3DS.

Figure 37: The region deleted in the *ph2a* mutant of wheat delimited in the three-barley mapping populations Chebec x Harrington, Clipper x Sahara and Galleon x Haruna Nijo. The data from each of the mapping populations is compiled into the consensus map (Langridge *et al.*, 1995).

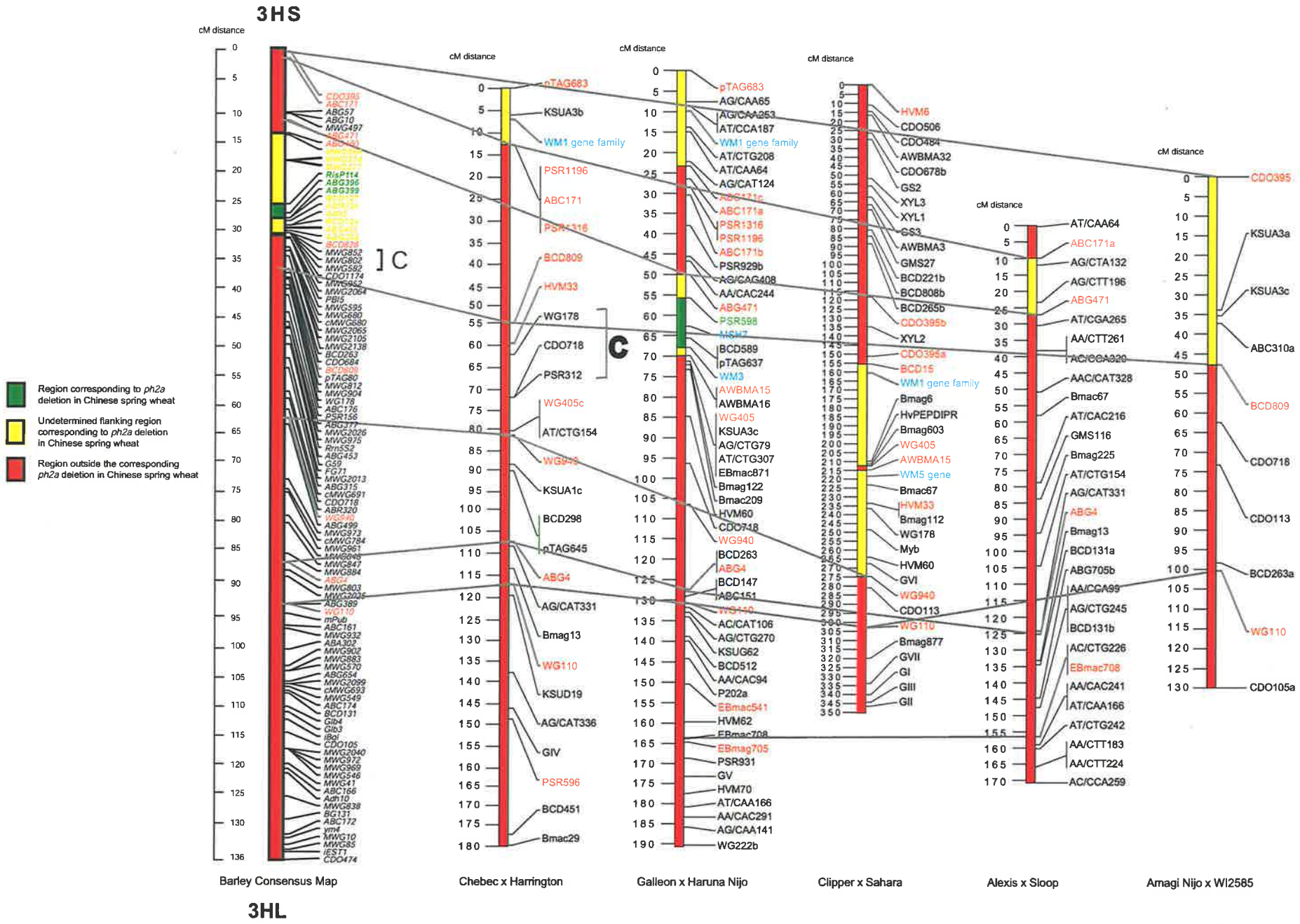

3HS
cM distance
Region corresponding to ph2a deletion in Chinese spring wheat
Undetermined flanking region corresponding to ph2a deletion in Chinese spring wheat
Region outside the corresponding ph2a deletion in Chinese spring wheat
]C
C
Barley Consensus Map
Chebec x Harrington
Galleon x Haruna Nijo
Clipper x Sahara
Alexis x Sloop
Amagi Nijo x WI2585
3HL

Figure 38: The syntenous *Ph2* regions in rice, wheat and barley (Ahn *et al.*, 1993; Langridge *et al.*, 1995; Nelson *et al.*, 1995c)

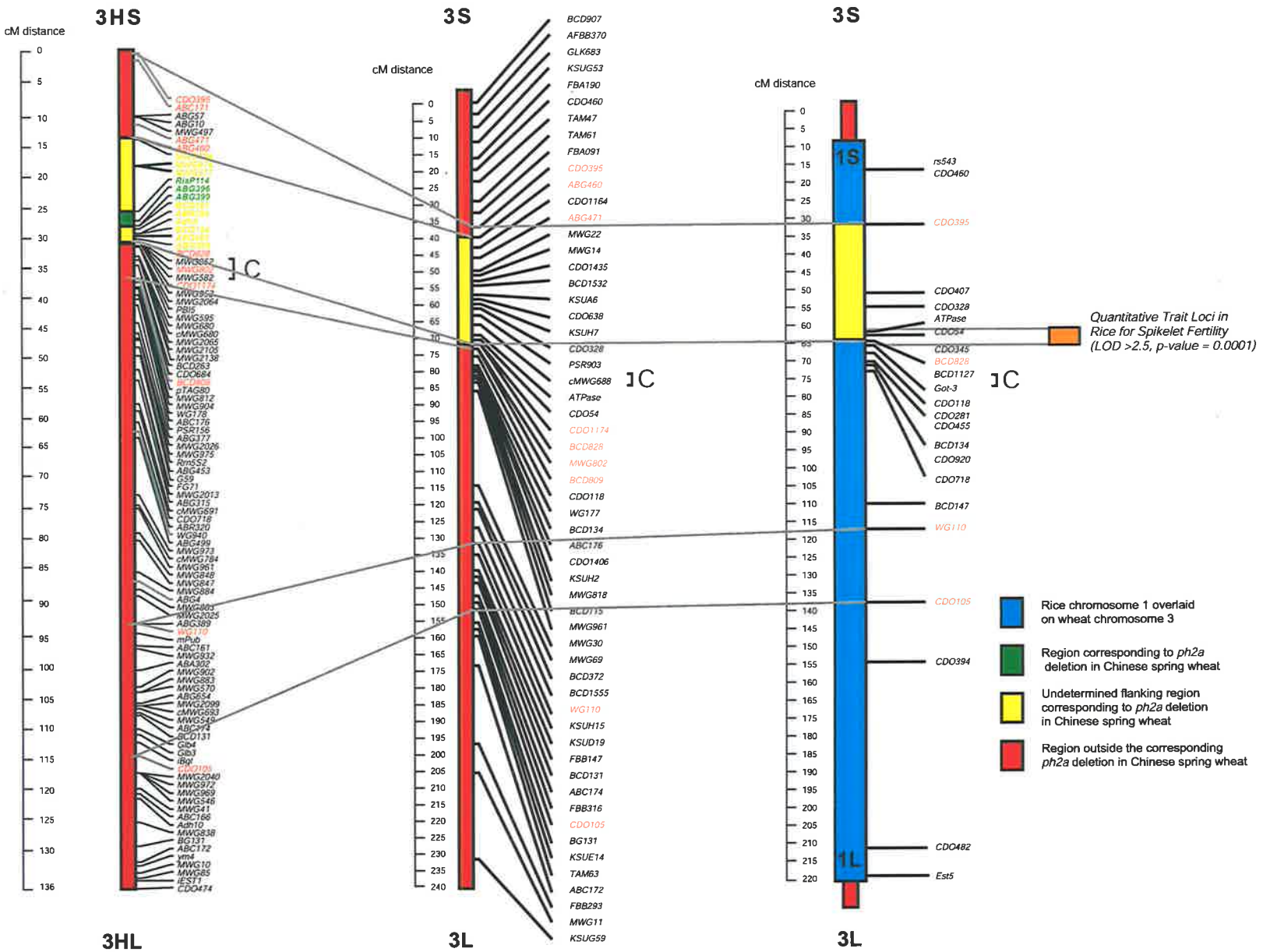

3HS
cM distance
1C
3HL
Barley Consensus Map
3S
cM distance
1C
3L
Wheat Consensus Map
3S
cM distance
1S
1L
1C
3L
Rice Chromosome 1 Overlaid on Wheat Chromosome 3
Quantitative Trait Loci in Rice for Spikelet Fertility (LOD >2.5, p-value = 0.0001)
Rice chromosome 1 overlaid on wheat chromosome 3
Region corresponding to ph2a deletion in Chinese spring wheat
Undetermined flanking region corresponding to ph2a deletion in Chinese spring wheat
Region outside the corresponding ph2a deletion in Chinese spring wheat

Figure 39 depicts the physical alignment of the YAC clones with respect to the rice genetic map (Harushima *et al.*, 1998; Saji *et al.*, 2001). From all the putative coding sequences in the region homologous to the *ph2a* deleted region in rice, the most likely meiotic or meiosis related genes are listed in **Table 18** and displayed in **Figure 39**.

5.3.2.2 Identification of wheat ESTs from PAC and BAC clones

Although the rice sequence data highlights potential coding sequences, there is no evidence that these genes are present or even expressed in wheat. To address this problem the rice PAC/BAC clones were used to identify wheat ESTs. BLAST 2.0 blastn (Altschul *et al.*, 1997) search results highlighted 280 wheat ESTs from the 5.5 Mb of rice PAC/BAC sequence. This 5.5 Mb of rice sequence contains approximately 900 putative coding sequences. The 280 wheat ESTs were then screened across the non-redundant database at NCBI via BLAST 2.0 tblastx (Altschul *et al.*, 1997) to give an indication of homologous proteins, which were subsequently, examined for homology to known meiotic or meiosis related genes (**Table 18**).

5.3.2.3 *Structural analysis of putative meiosis genes*

The structure of predicted rice coding sequences that show homology to known meiosis related genes are displayed in **Figure 40** to **Figure 45**.

5.3.2.4 Analysis of rice and barley meiosis related phenotypic traits

Both the RiceGenes and GrainGenes database were screened for QTLs in the region of *Ph2* that could be associated to a meiotic gene effect. Only one such significant QTL (LOD >2.5, p-value = 0.0001) was found localised to the delimited homologous *Ph2* region on rice chromosome 1. Interestingly the QTL was for spikelet fertility and is illustrated in **Figure 38**. No other mapped phenotypic traits related to meiosis were found localised to the short arm of rice chromosome 1. However, a gene termed *msg5* (male sterile 5) is located on the short arm of barley chromosome 3. It is not known if this gene resides in the homologous *ph2a* deletion region.

Figure 39: Physical alignment of the YAC clones with respect to the rice genetic map (Harushima *et al.*, 1998; Saji *et al.*, 2001), highlighting the estimated physical size of the homologous *Ph2* region in rice with the physical positions of the putative meiotic or meiosis related genes.

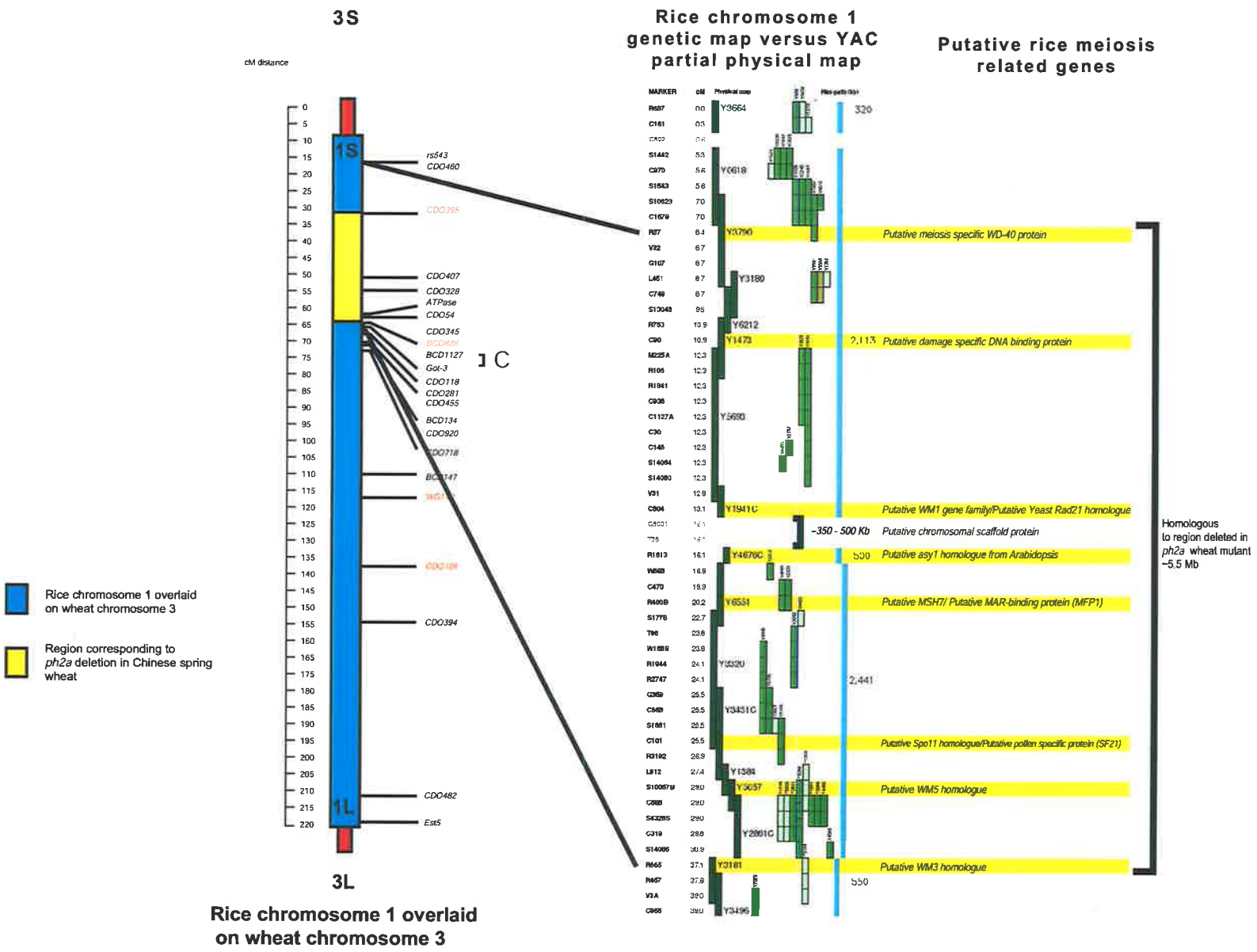

3S
Rice chromosome 1 genetic map versus YAC partial physical map
Putative rice meiosis related genes
cM distance
1S
1L
rs543
CDO460
CDO395
CDO407
CDO328
ATPase
CDO54
CDO345
BCD1127
Got-3
CDO118
CDO281
CDO455
BCD134
CDO920
CDO718
CDO394
CDO482
Est5
Y3664
Y0618
Y3790
Y3180
Y6212
Y1473
Y5690
Y1941C
Y4676C
Y6551
Y3431C
Y1384
Y3057
Y2861C
Y3181
Y3496
320
2,113
2,441
550
Putative meiosis specific WD-40 protein
Putative damage specific DNA binding protein
Putative WM1 gene family/Putative Yeast Rad21 homologue
~350 - 500 Kb
Putative chromosomal scaffold protein
Putative asy1 homologue from Arabidopsis
Putative MSH7/ Putative MAR-binding protein (MFP1)
Putative Spo11 homologue/Putative pollen specific protein (SF21)
Putative WM5 homologue
Putative WM3 homologue
Homologous to region deleted in ph2a wheat mutant ~5.5 Mb
Rice chromosome 1 overlaid on wheat chromosome 3
Region corresponding to ph2a deletion in Chinese spring wheat
3L
Rice chromosome 1 overlaid on wheat chromosome 3

Table 18: Putative meiosis related genes identified from rice PAC and BAC sequence along with cereal ESTs.

Gene	tblastx Expectation value	Homologue(s)	Putative Function	Reference
Wheat putative *Asy1* EST	$3xE^{-45}$ $3xE^{-06}$	*Arabidopsis* asynaptic mutant 1 Yeast HOP1 protein.	Chromosome synapsis via synaptonemal complex formation.	(Caryl *et al.*, 2000)
Wheat *TaMSH7* coding sequence	0 $3xE^{-58}$	*Arabidopsis* AtMSH6-2 protein Yeast MSH6 protein.	DNA mismatch repair.	(Dong *et al.*, 2001b)
Wheat *WM5* coding sequence	$2xE^{-66}$	Rice hypothetical protein.	Meiotic expression pattern. Function not determined.	(Dong *et al.*, 2001a)
Wheat *WM1* gene family (coding sequence for *WM1.1*)	0 $3xE^{-80}$	Barley Cf2/Cf5 disease-like resistance protein homolog. Tomato Cf2.1 disease resistance protein.	Low level meiotic expression pattern. Function not determined.	(Ji, 1992; Ji & Langridge, 1994)
Wheat *WM3* coding sequence	$2xE^{-18}$ $6xE^{-18}$	Sugar Beet non-specific lipid transfer protein precursor. Cotton lipid transfer protein precursor.	Meiotic expression pattern. Function not determined.	(Letarte, 1996)
Rice putative damage-specific DNA binding protein coding sequence	0 $2xE^{-25}$	*Arabidopsis* damage-specific DNA binding protein. Human damage-specific DNA binding protein.	Lesion recognition of UV-light damaged DNA. Potential role in DNA repair.	(Abramic *et al.*, 1991; Hirschfeld *et al.*, 1990; Protic *et al.*, 1989)
Rice putative *ScII*-like coding sequence	0 $3xE^{-11}$	Putative *Arabidopsis* coding sequence. Chicken chromosomal scaffold protein.	Chromosome structural protein important for chromosome condensation.	(Saitoh *et al.*, 1995)
Wheat putative *WD40* protein EST	$5xE^{-67}$ $3xE^{-32}$	*Arabidopsis* WD40 protein. Rat WD40 protein.	Potential role in metaphase-anaphase transition and recombination.	(Townsley & Ruderman, 1998) (Ito *et al.*, 2001)
Wheat putative MAR-binding protein EST	$4xE^{-18}$ $3xE^{-23}$ $4xE^{-22}$	*Arabidopsis* MAR-binding protein MFP1. Tobacco MAR-binding protein MFP1. Tomato MAR-binding protein MFP1.	Chromatin structural protein important for interaction between chromatin and the nuclear matrix. Particularly related to gene expression.	(Harder *et al.*, 2000)

Figure 40: Putative structure of rice Asy1-like protein. A – Genomic sequence, B – Predicted cDNA highlighting corresponding cereal ESTs, C – Diagramatic representation of the primary peptide sequence depicting putative structural motifs, D – Comparison of peptide sequence diagrams highlighting commonality in areas of conserved structural motifs (boxed), E - Alignment of primary peptide sequence structural motifs within areas of conservation as determined by the Conserved Domain Database v1.51 at NCBI.

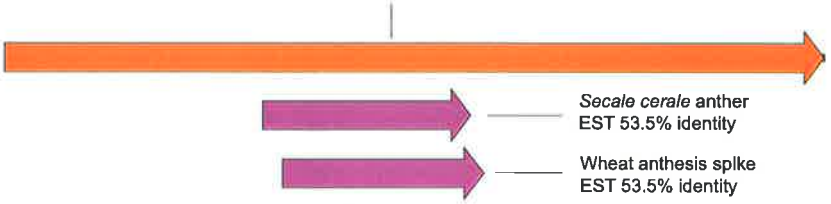
A

3887 bp

Intron exon structure of rice putative Asy1 protein

B

Rice putative coding sequence for Asy1-like protein



Secale cerale anther EST 53.5% identity

Wheat anthesis splke EST 53.5% identity

2295 bp

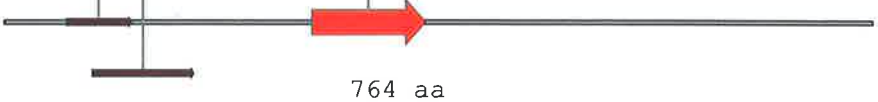
C

Rice putative primary peptide structure for Asy1-like protein

Arginine rich region

Proline rich region

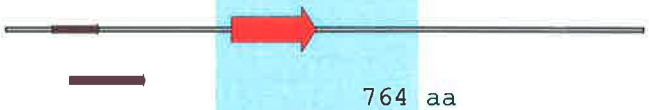
HORMA domain



764 aa

D

Rice putative primary peptide structure for Asy1 protein



764 aa

Arabidopsis thaliana primary peptide structure asynaptic mutant 1 protein

506 aa

Arabidopsis thaliana primary peptide structure for a hypothetical protein

Leucine Zipper

1552 aa

Saccharomyces cereviseae primary peptide structure for HOP1 protein

Zinc Finger

605 aa

Kluyveromyces lactis primary peptide structure for HOP1 protein

581 aa

Human primary peptide structure for MAD2, Spindle Assembly Checkpoint Protein

187 aa

E

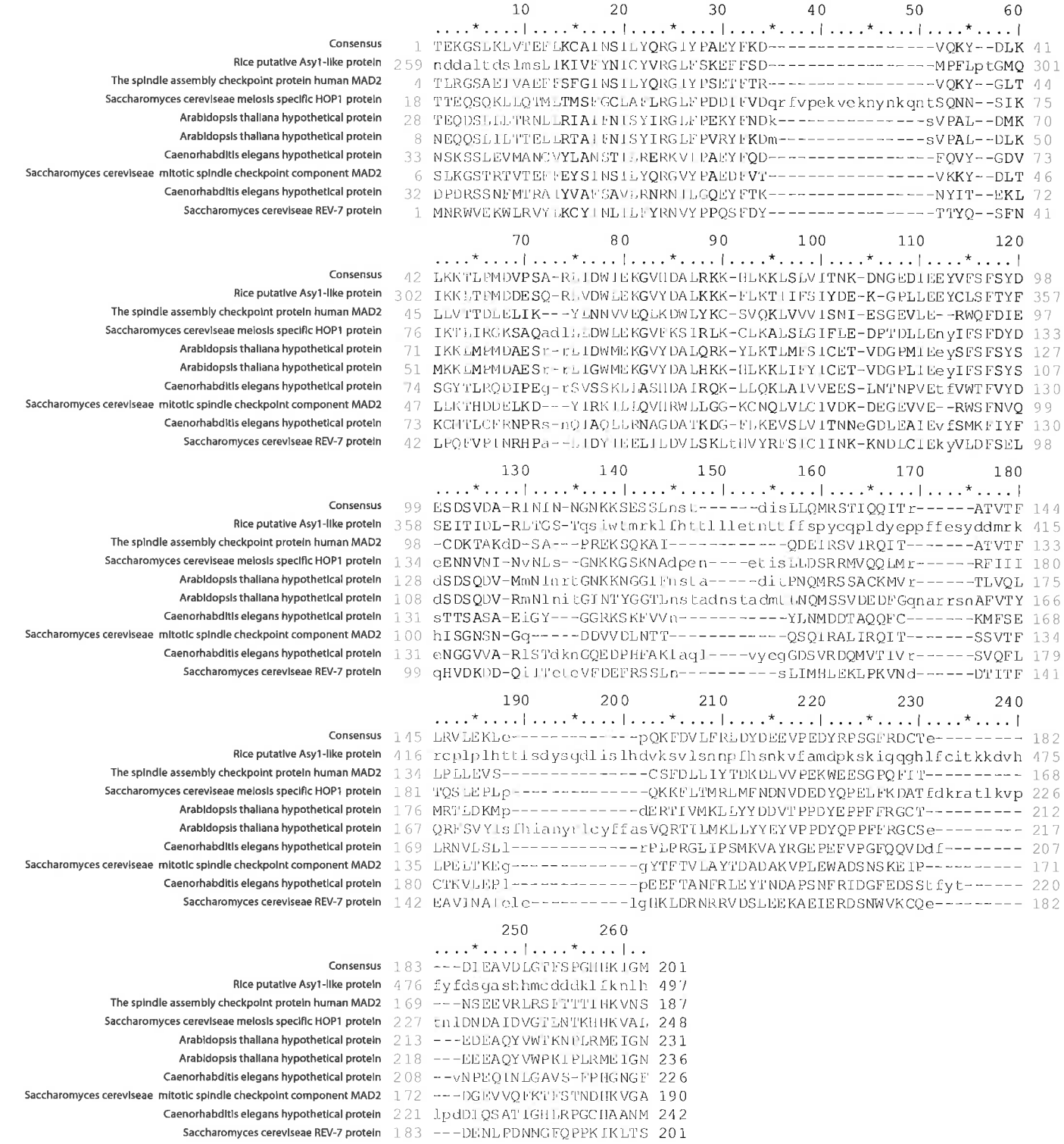
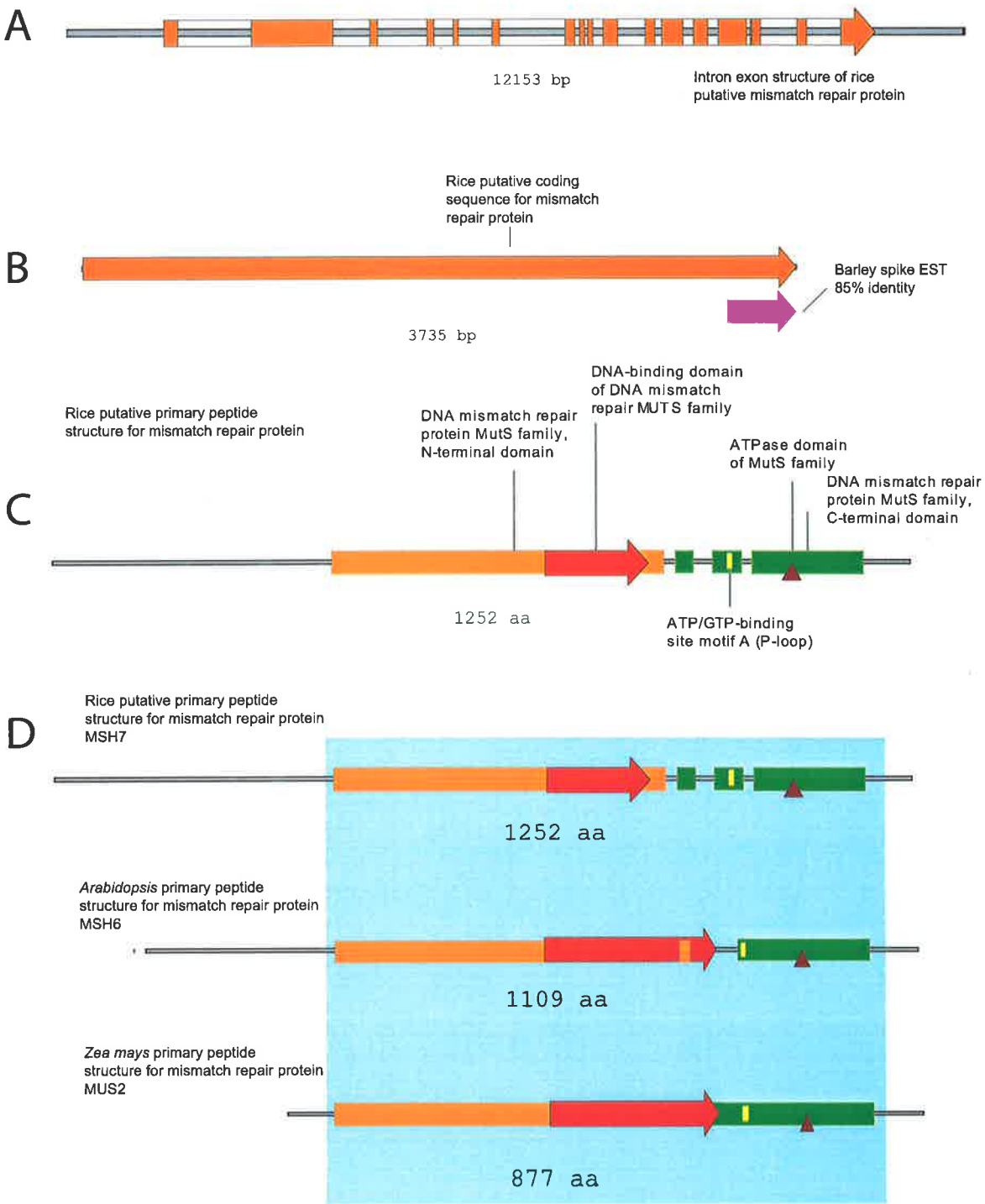

Consensus
Rice putative Asy1-like protein
The spindle assembly checkpoint protein human MAD2
Saccharomyces cerevisiae meiosis specific HOP1 protein
Arabidopsis thaliana hypothetical protein
Arabidopsis thaliana hypothetical protein
Caenorhabditis elegans hypothetical protein
Saccharomyces cerevisiae mitotic spindle checkpoint component MAD2
Caenorhabditis elegans hypothetical protein
Saccharomyces cerevisiae REV-7 protein

Figure 41: Putative structure of rice TaMSH7-like protein. A – Genomic sequence, B – Predicted cDNA highlighting corresponding cereal ESTs, C – Diagramatic representation of the primary peptide sequence depicting putative structural motifs, D – Comparison of peptide sequence diagrams highlighting commonality in areas of conserved structural motifs (boxed), E - Alignment of primary peptide sequence within areas of high conservation as displayed in D.


A
12153 bp
Intron exon structure of rice putative mismatch repair protein
B
Rice putative coding sequence for mismatch repair protein
Barley spike EST 85% identity
3735 bp
C
Rice putative primary peptide structure for mismatch repair protein
DNA mismatch repair protein MutS family, N-terminal domain
DNA-binding domain of DNA mismatch repair MUTS family
ATPase domain of MutS family
DNA mismatch repair protein MutS family, C-terminal domain
1252 aa
ATP/GTP-binding site motif A (P-loop)
D
Rice putative primary peptide structure for mismatch repair protein MSH7
1252 aa
Arabidopsis primary peptide structure for mismatch repair protein MSH6
1109 aa
Zea mays primary peptide structure for mismatch repair protein MUS2
877 aa

E

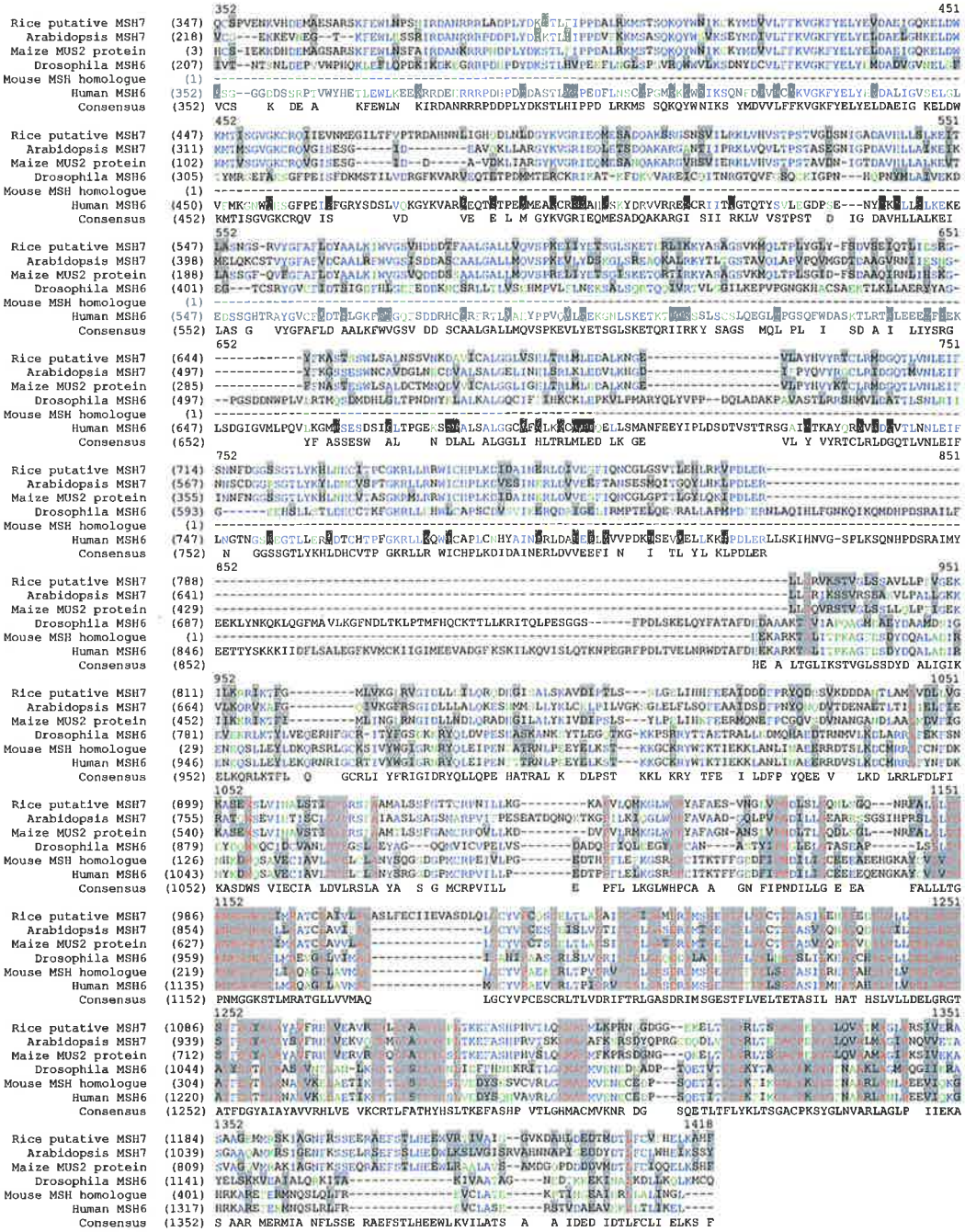
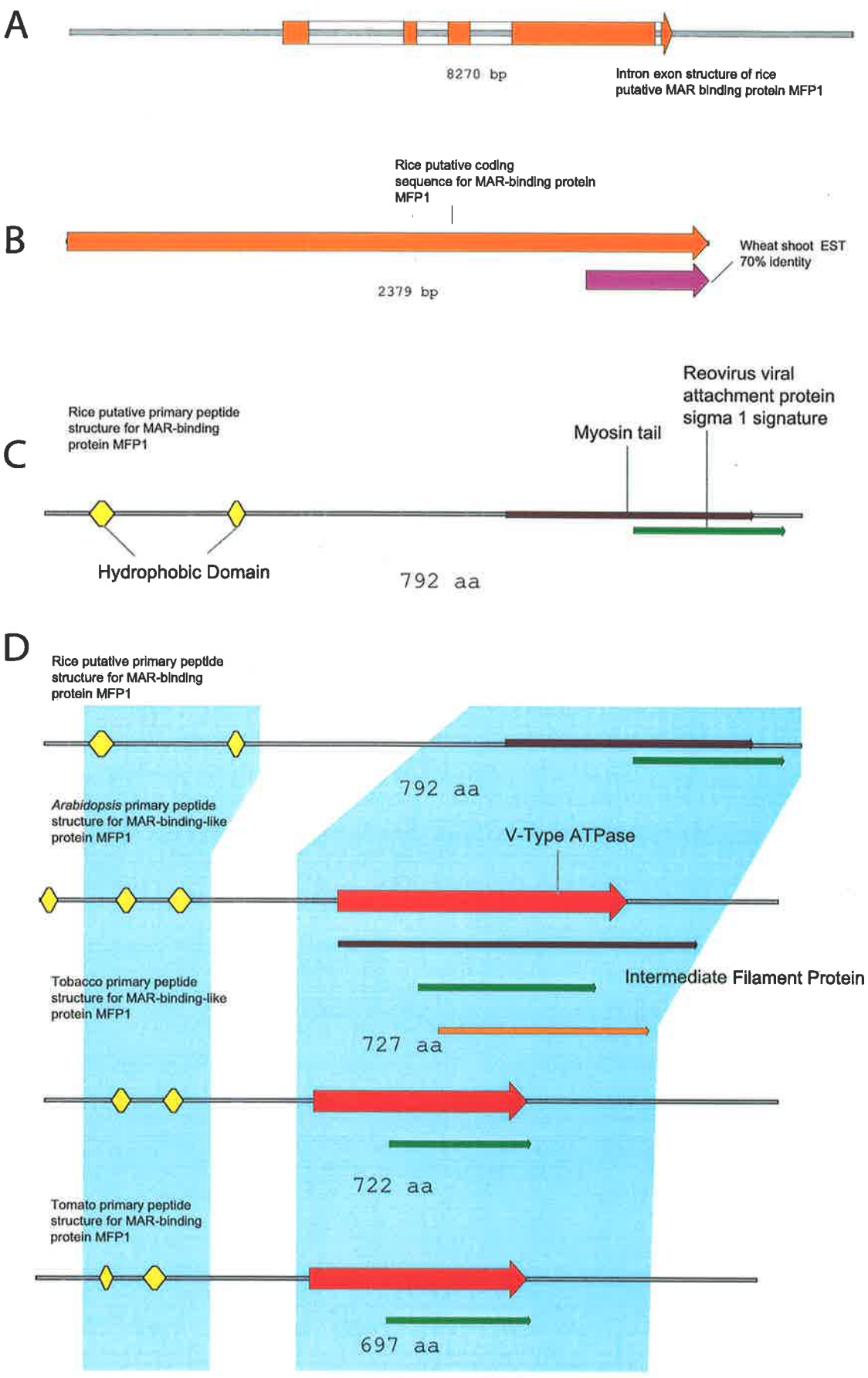

Rice putative MSH7
Arabidopsis MSH7
Maize MUS2 protein
Drosophila MSH6
Mouse MSH homologue
Human MSH6
Consensus

Figure 42: Putative structure of rice MFP1-like protein. A – Genomic sequence, B – Predicted cDNA highlighting corresponding cereal ESTs, C – Diagramatic representation of the primary peptide sequence depicting putative structural motifs, D – Comparison of peptide sequence diagrams highlighting commonality in areas of conserved structural motifs (boxed), E - Alignment of primary peptide sequence within areas of high conservation as displayed in D.


A
8270 bp
Intron exon structure of rice putative MAR binding protein MFP1
B
Rice putative coding sequence for MAR-binding protein MFP1
Wheat shoot EST 70% identity
2379 bp
C
Rice putative primary peptide structure for MAR-binding protein MFP1
Reovirus viral attachment protein sigma 1 signature
Myosin tail
Hydrophobic Domain
792 aa
D
Rice putative primary peptide structure for MAR-binding protein MFP1
792 aa
Arabidopsis primary peptide structure for MAR-binding-like protein MFP1
V-Type ATPase
Intermediate Filament Protein
Tobacco primary peptide structure for MAR-binding-like protein MFP1
727 aa
722 aa
Tomato primary peptide structure for MAR-binding protein MFP1
697 aa

E

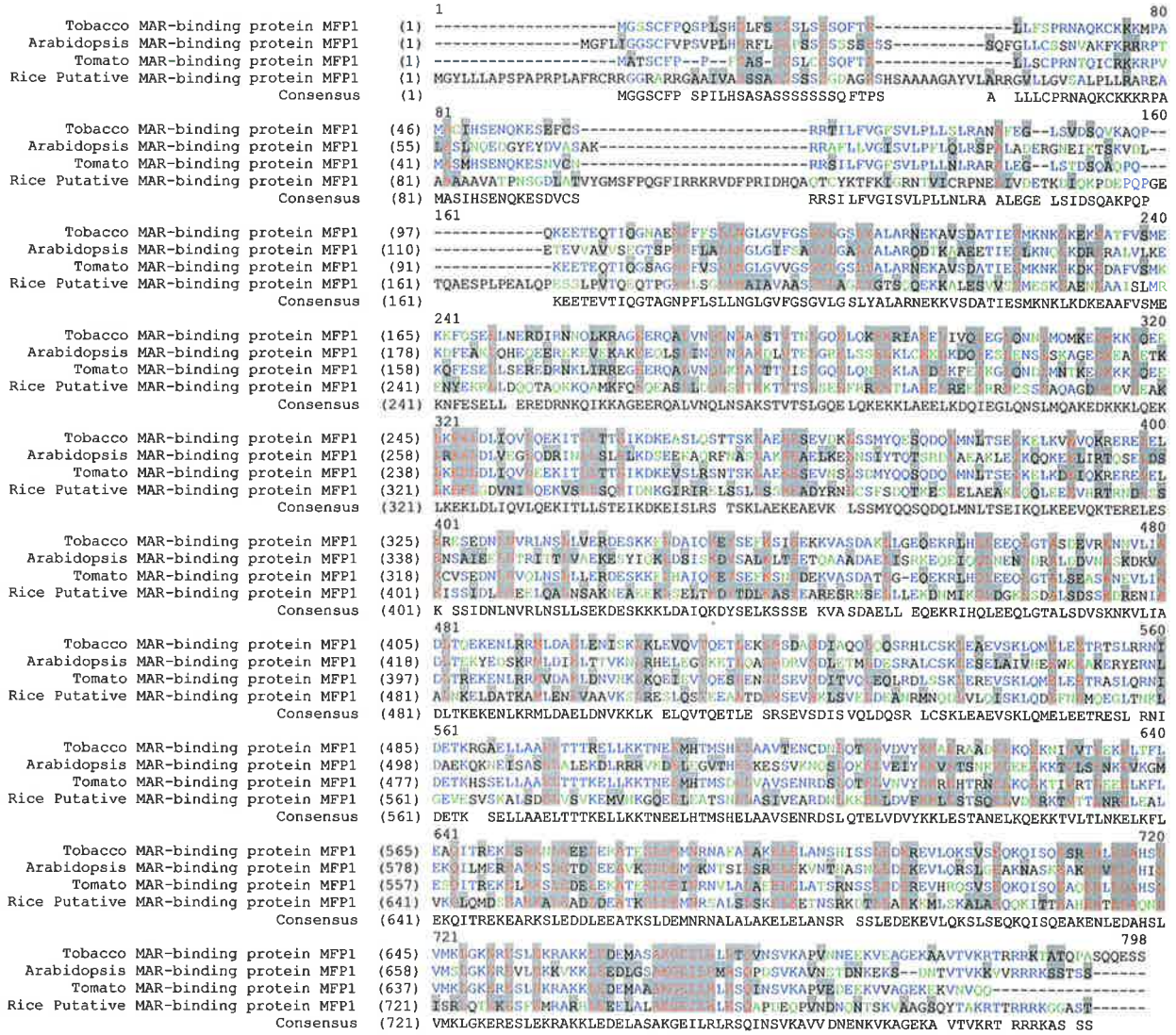

Tobacco MAR-binding protein MFP1
Arabidopsis MAR-binding protein MFP1
Tomato MAR-binding protein MFP1
Rice Putative MAR-binding protein MFP1
Consensus

Figure 43: Putative structure of rice ScII-like protein. A – Genomic sequence, B – Predicted cDNA highlighting corresponding cereal ESTs, C – Diagramatic representation of the primary peptide sequence depicting putative structural motifs, D – Comparison of peptide sequence diagrams highlighting commonality in areas of conserved structural motifs (boxed), E - Alignment of primary peptide sequence within areas of high conservation as displayed in D.

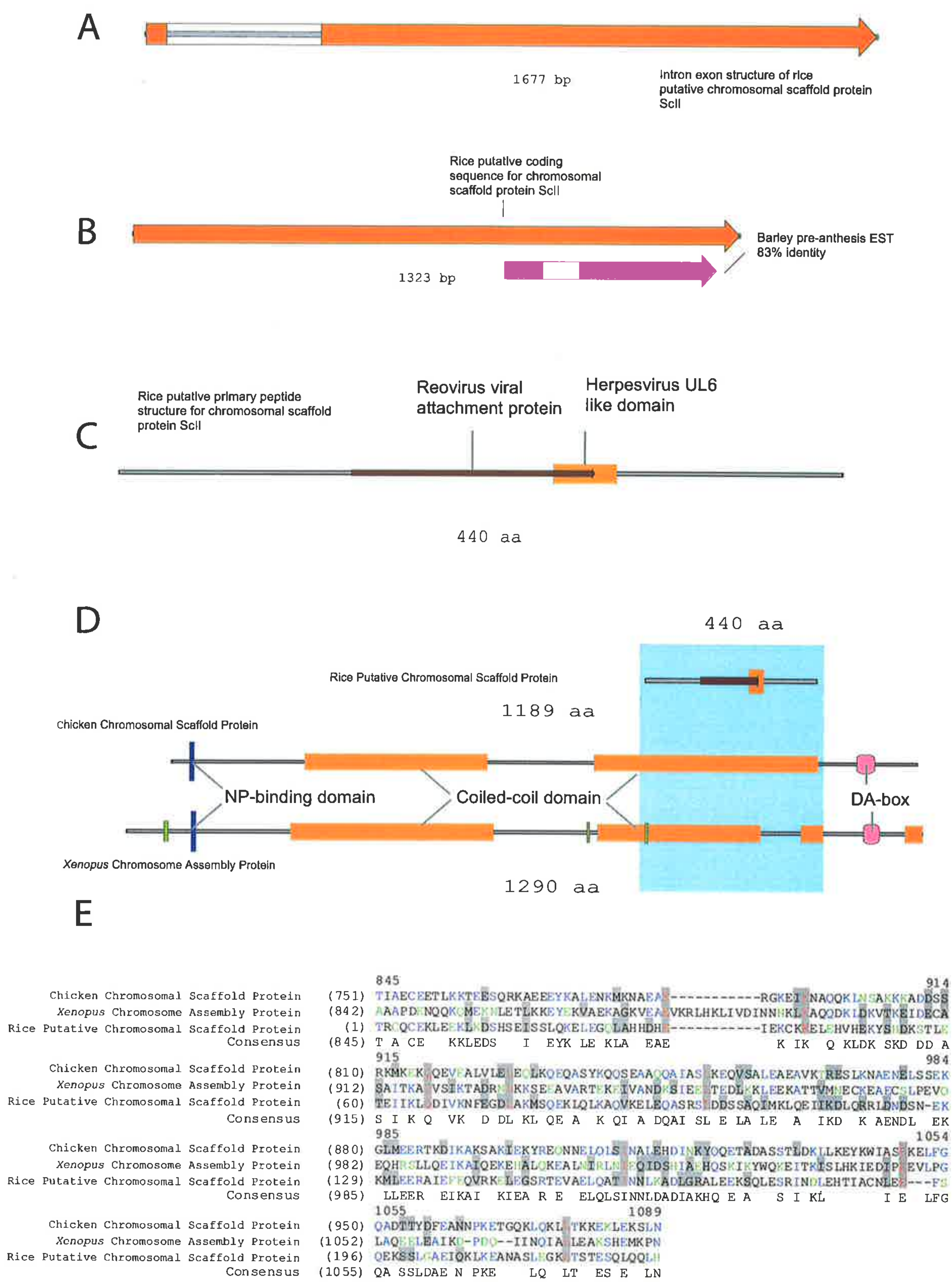

A
1677 bp
Intron exon structure of rice putative chromosomal scaffold protein ScII
B
Rice putative coding sequence for chromosomal scaffold protein ScII
Barley pre-anthesis EST 83% identity
1323 bp
C
Rice putative primary peptide structure for chromosomal scaffold protein ScII
Reovirus viral attachment protein
Herpesvirus UL6 like domain
440 aa
D
440 aa
Rice Putative Chromosomal Scaffold Protein
1189 aa
Chicken Chromosomal Scaffold Protein
NP-binding domain
Coiled-coil domain
DA-box
Xenopus Chromosome Assembly Protein
1290 aa
E
Chicken Chromosomal Scaffold Protein
Xenopus Chromosome Assembly Protein
Rice Putative Chromosomal Scaffold Protein
Consensus

Figure 44: Putative structure of rice damage-specific DNA binding protein. A – Genomic sequence, B – Predicted cDNA highlighting corresponding cereal ESTs, C – Diagramatic representation of the primary peptide sequence depicting putative structural motifs, D – Comparison of peptide sequence diagrams highlighting commonality in areas of conserved structural motifs (boxed), E - Alignment of primary peptide sequence within areas of high conservation as displayed in D.

A

5295 bp

Intron exon structure of rice putative damage specific DNA binding protein

B

Rice putative coding sequence for damage specific DNA binding protein

1479 bp

C

Rice putative primary peptide structure for damage specific DNA binding protein

G-protein beta WD-40 repeats

Regulator of chromosome condensation (RCC1) signature 2

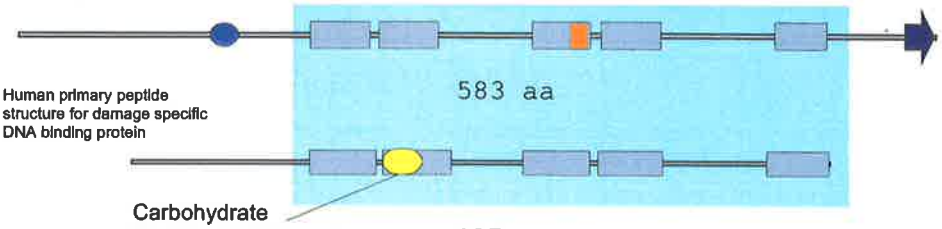
Bipartite nuclear localisation signal

Zn-finger CCHC type

583 aa

D

Rice putative primary peptide structure for damage specific DNA binding protein



Human primary peptide structure for damage specific DNA binding protein

Carbohydrate Kinase, FGGY family

427 aa

E

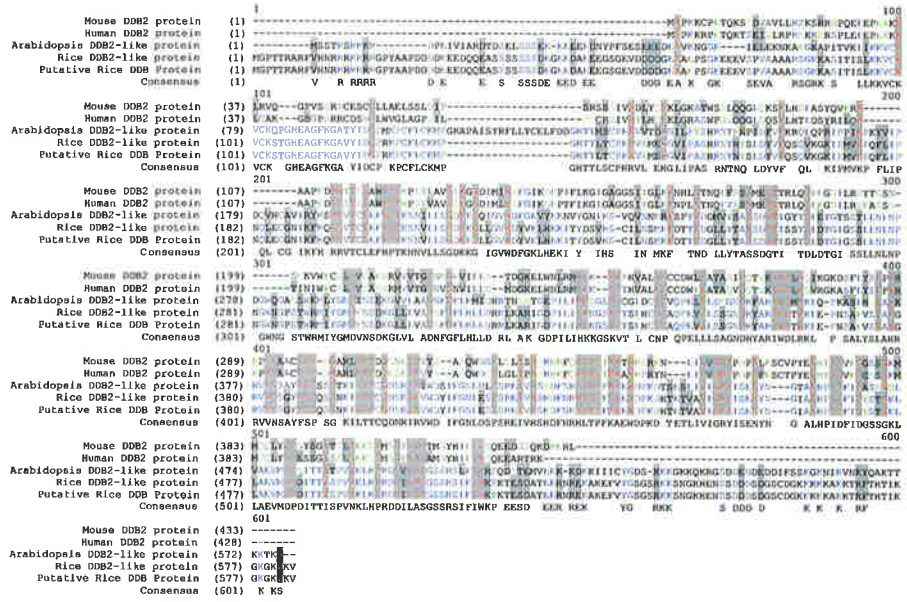


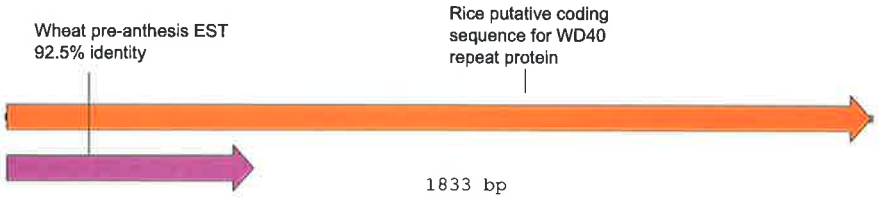
Figure 45: Putative structure of rice WD40 repeat like protein. A – Genomic sequence, B – Predicted cDNA highlighting corresponding cereal ESTs, C – diagrammatic representation of the primary peptide sequence depicting putative structural motifs, D – Comparison of peptide sequence diagrams highlighting commonality in areas of conserved structural motifs (boxed), E - Alignment of primary peptide sequence within areas of high conservation as displayed in D.

A

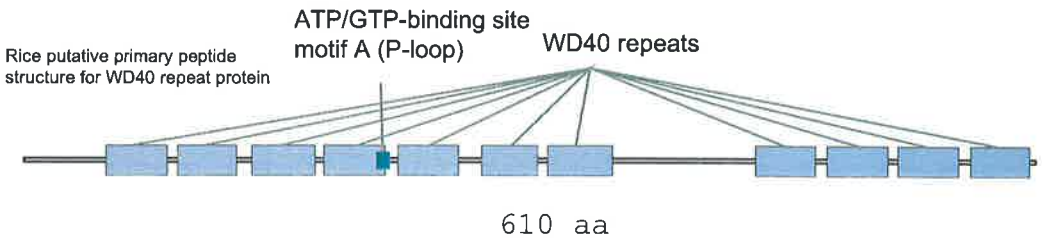
6302 bp

Intron exon structure of rice putative WD40 repeat protein

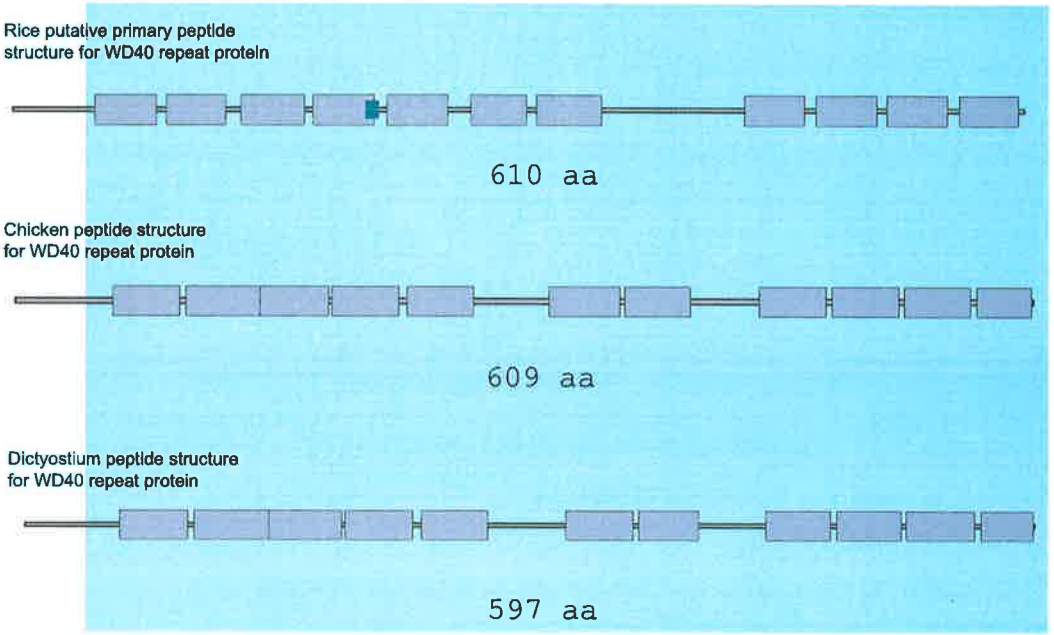
B



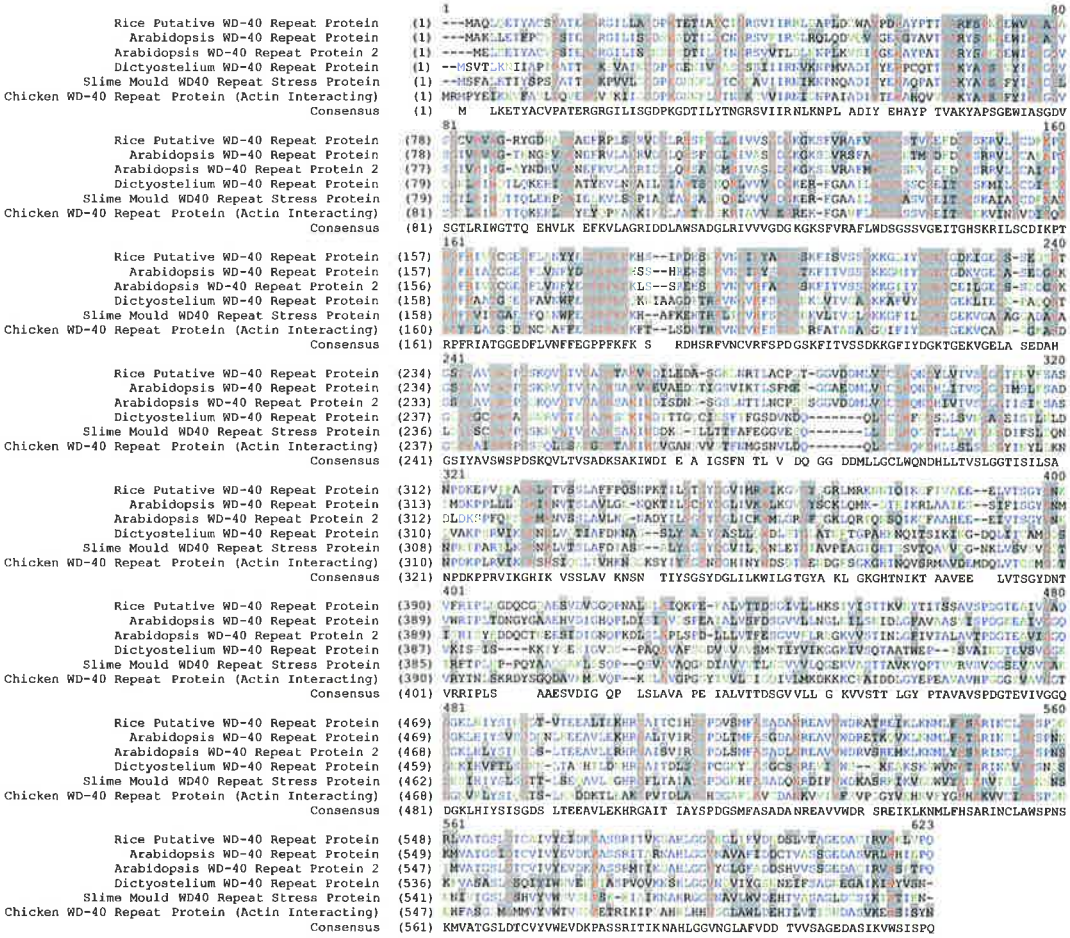
C



D



E


Rice Putative WD-40 Repeat Protein
Arabidopsis WD-40 Repeat Protein
Arabidopsis WD-40 Repeat Protein 2
Dictyostelium WD-40 Repeat Protein
Slime Mould WD40 Repeat Stress Protein
Chicken WD-40 Repeat Protein (Actin Interacting)
Consensus

5.4 Discussion

This study used colinearity of molecular markers in the grasses, in combination with the rice genome sequence, to identify potential *Ph2* gene(s). Comparative mapping between rice, barley and wheat, utilising molecular markers spanning the *Ph2* locus, has identified nine putative meiosis related genes in a 5.5 Mb region of rice.

5.4.1 Physical size of *ph2a* deletion in wheat

Comparison of the wheat and rice linkage maps has highlighted that the region homologous to the wheat deleted *ph2a* region in rice is between 3 and 5.5 Mb. Utilising this estimation of physical size and the difference between the genome size and chromosome complement between hexaploid wheat (15,966 Mb per 1C nucleus, 2n = 42) and rice (431 Mb per 1C nucleus, 2n = 22) (Arumuganathan & Earle, 1991), the estimated size of the *ph2a* deletion is likely to be between 58 and 107 Mb.

5.4.2 Putative gene content of *ph2a* region

Analysis of the RiceGenes and GrainGenes databases highlighted genes corresponding to fertility on the *Triticeae* homoeogroup 3 and the corresponding syntenous region on rice chromosome 1. Specifically a significant QTL (LOD >2.5, p-value = 0.0001) for spikelet fertility (Lin *et al.*, 1996) was found in rice and a male sterility gene termed *msg5*, in barley (Franckowiak, 1997). These results demonstrate that there is a common fertility locus among the *Gramineae*. It is anticipated that similar genes controlling fertility and more specifically meiosis would be present in rice, wheat and barley.

This rice 5.5 Mb region homologous to the *ph2a* deleted region harbours approximately 900 predicted coding sequences of which nine are likely to be specifically involved in meiotic cell development. The predicted gene density within this region of rice is one every 6 Kb whilst it is expected to be one every 64 to 118 Kb in wheat. However, analysis of the *WM1* gene cluster in *T. tauschii* suggests that there is one gene every 19 Kb (refer to 3.3.6). This conflicting result can best be explained by genes being located in "gene islands" flanked by

large stretches of repetitive transposable element-like sequences. From the 900 predicted coding sequences, 280 wheat ESTs were identified. This suggests that at least 31% of the likely wheat orthologues to the predicted rice genes are expressed. It should be noted that the rice gene prediction program is likely to have given an overestimate of gene number.

5.4.2.1 Structural aspects of the Ph2 *region*

The presence of nine potentially meiosis related genes in the region deleted in the homeologous chromosome pairing mutant, *ph2a*, suggests that the *Ph2* locus is most likely to be a complex of linked genes influencing different meiotic processes, rather than a single gene. However, the pairing homeologous wheat mutant *ph2b*, which was generated through ethyl methanesulfonate (EMS) treatment contradicts the idea of a complex of linked meiotic genes. EMS is thought to induce random point mutations (G-C base pair replaced with an A-T pair) and if the *Ph2* locus is a complex of nine linked meiosis genes, the likelihood of having a deleterious EMS induced mutation in each of the linked meiosis genes is low but not inconceivable. It is more likely that only a few of the nine putative meiosis genes are responsible for the *Ph2* phenotype. Further study of each of the potential meiotic genes in the *ph2b* mutant will clarify whether the *Ph2* locus is indeed controlled by a complex of linked genes or a single gene.

5.4.2.2 Candidate Ph2 *genes*

The most likely *Ph2* candidates determined by peptide sequence homology (BLAST 2.0 tblastx) to functionally characterised proteins (**Table 18**) are described below. Possible function of these characterised proteins in terms of meiotic processes and specifically their relevance to the *Ph2* phenotype, are described.

5.4.2.2.1 Asy1-like protein

Homology of the wheat putative *Asy1*-like EST to the *A. thaliana* asynaptic mutant 1 protein is particularly interesting as *Asy1* shows homology with the yeast *HOP1* meiotic gene (**Figure 40**). Mutants of the yeast *HOP1* gene show

reduced meiotic recombination (10-25% of wild-type levels), and extremely low levels of spore viability (<1%) (Hollingsworth *et al.*, 1990). HOP1 is found to be localised within the axial elements of the synaptonemal complex with the protein accumulating at numerous discrete foci on meiotic chromosomes during early prophase. Little protein persists by the time the chromosomes have fully synapsed at pachytene. HOP1 protein accumulation is dependant upon another synaptonemal complex protein, Red1 and their interaction is suggested to be mediated by Mek1, a serine-threonine protein kinase (Hollingsworth & Ponte, 1997).

HOP1 is a 605 amino acid peptide with a non-classical zinc finger about position 371, which has been implicated in DNA binding activity. Asy1 shows homology to HOP1 at the N-terminal HORMA–like domain. This extends to conservation in the HORMA-like domain within the predicted peptide sequence of the wheat and rye putative Asy1-like ESTs (**Figure 40**). HORMA domains have been seen in a variety of proteins that; interact with chromatin, are involved in synaptonemal complex formation, DNA repair or DNA synthesis (Aravind & Koonin, 1998). Homology between Asy1 and HOP1 at the N-terminus of the predicted peptides suggests substantial functional different\ces but it may be the case that although the proteins are functionally homologous, this does not extend to primary amino acid sequence (Caryl *et al.*, 2000). Similar functionality, not extended to the primary peptide sequence, has been reported for other synaptonemal complex proteins including Scp1 from rat (Meuwissen *et al.*, 1992) and yeast Zip1 (Sym *et al.*, 1993). It is yet to be seen whether Asy1 is directly involved in the meiotic recombination process of plants.

The meiotic phenotype of *Asy1 Arabidopsis* T-DNA insertion mutant has been described as partially asynaptic, because a low level of synapsis is indicated by the presence of from one to three bivalents in some metaphase I cells, with an overall average of 1.5 bivalents per cell (Caryl *et al.*, 2000; Ross *et al.*, 1997). Characteristically, mutant embryo and pollen mother cells fail to undergo extensive chromosome synapsis during early prophase I, resulting in frequent univalent formation at diakinesis and metaphase I. This contrasts to wild-type

meiosis where normal chromosome synapsis during zygotene leads to full synpasis at pachytene and the regular presence of five chiasmate bivalents at diakinesis and metaphase I (Caryl *et al.*, 2000).

If the wheat putative *Asy1*-like EST does represent the wheat orthologue to the *Arabidopsis* Asy1 protein then it is a strong candidate for the *Ph2* gene or is functionally important in the *Ph2* phenotype. This proposal is based firstly, on the homology between conserved peptide domains (HORMA domain) of the wheat putative *Asy1*-like EST (**Figure 40**), *Arabidopsis* Asy1 protein and the known meiotic HOP1 protein of yeast. Secondly, the proposed wheat *Asy1* orthologue is likely to be localised within the region deleted in the *ph2a* mutant (**Figure 39**). Thirdly, *Ph2* has also been implicated in the control of homeologous chromosome pairing and recombination as a consequence of asynapsis (Martinez *et al.*, 2001) and any mutation in the orthologous *Ph2* gene in a diploid species like *Arabidopsis* is likely to result in almost complete sterility. T-DNA insertion mutant of *Asy1* is sterilite. *Ph2* mutants of wheat do not show complete sterility. Along with *Ph2* on chromosome 3DS, 3AS possesses another suppressor of homeologous chromosome pairing which may complement the loss of *Ph2*. Essential meiotic genes, such as *Asy1*, may give an insight into the role of ploidy in the control of meiosis. Polyploidy has been shown to modulate gene expression in yeast (Galitski *et al.*, 1999) and in *Arabidopsis* (Mittelsten Scheid *et al.*, 1996). This potential buffering system is likely to be useful in the detailed study of genes that control meiosis.

5.4.2.2.2 TaMSH7 protein

The mismatch repair system plays a critical role in maintaining genetic stability by recognising and processing mismatched nucleotides that may occur during DNA replication, genetic recombination and some types of chemical damage to DNA. Yeast and animal MSH2, MSH3 and MSH6 proteins are major components involved in DNA mismatch repair. A special feature in plants is the presence of two MSH6-like proteins. AtMSH6 and AtMSH7 in *Arabidopsis* have been shown to be homologues of MSH6 (Ade *et al.*, 1999; Culligan & Hays, 2000). Apart from mismatch correction, MSH2, MSH3 and MSH6 also play a

role in preventing recombination between divergent DNA sequences (Chambers *et al.*, 1996; Chen & Jinks-Robertson, 1998; Datta *et al.*, 1996; Hunter *et al.*, 1996; Selva *et al.*, 1997; Selva *et al.*, 1995; Sugawara *et al.*, 1997). In bacteria, the mismatch repair system acts as a barrier to recombination between divergent sequences (Rayssiguier *et al.*, 1989; Vulic *et al.*, 1997). Loss of mismatch repair function results in dramatically elevated homeologous recombination even if the sequence divergence is as high as 20% (Rayssiguier *et al.*, 1989). Further evidence shows that MSH2, MSH3, MSH6 are involved in meiotic recombination. A MSH2-MSH3 complex is required in the processing of non-homologous ends during double-strand break-induced recombination (Sugawara *et al.*, 1997). The MSH2-MSH6 complex has also been shown to bind to Holliday junctions during meiotic recombination (Alani *et al.*, 1994; Marsischky *et al.*, 1999). The suppression of homeologous recombination mediated by the mismatch repair system is thought to be through the recognition of mispairs in the recombination intermediates and the prevention of processing of these recombination intermediates (Baker *et al.*, 1995; de Wind *et al.*, 1995; Edelmann *et al.*, 1996). Given the role of the yeast and animal mismatch repair system in meiosis there is still no evidence in plants that the mismatch repair system is involved in meiotic recombination. Localisation of wheat *TaMSH7* gene (**Figure 39**) within the region deleted in the *ph2a* mutant suggests that a mismatch repair system is involved in chromosome pairing and recombination (Dong *et al.*, 2001b). Specifically *TaMSH7* protein (**Figure 41**) could represent the *Ph2* gene product or be functionally important in the *Ph2* phenotype.

5.4.2.2.3 MFP1-like protein

The interaction of chromatin with the nuclear matrix via matrix attachment regions (MARs) on the DNA is considered to be of fundamental importance for higher-order chromatin organisation and the regulation of gene expression (Harder *et al.*, 2000). The MAR-nuclear matrix interactions are believed to create independent chromatin loop domains, and the location of a gene with respect to this domain structure is thought to influence its expression level (Spiker & William, 1996). The nuclear matrix has been biochemically defined as the insoluble component that remains after treatment of isolated nuclei with

nucleases and extraction of proteins with different methods (Berezney & Coffey, 1974; Mirkovitch *et al.*, 1984). Structurally, the nuclear matrix resembles the dense network of fibers seen in the cytoplasmic skeleton (He *et al.*, 1990). The chromatin loops are presumed to attach to these matrix fibers by protein-DNA interactions with the MARs.

A barley *MFP1*-like EST shows homology with the predicted rice MFP1-like protein (**Figure 42**), which in turn has homology to the plant MAR-binding protein, MFP1, from tomato (Meier *et al.*, 1996). MFP1 has the structural features of a filament-like protein, shows similarity to nuclear and cytoplasmic filament proteins and preferentially binds to MAR sequences from both animals and plants.

MARs are generally AT-rich DNA sequences that range from approximately 300 bp to 2.5 Kb in length and are ubiquitous. Currently the consensus sequence for MARs is "TCTTTAATTTCTAATATATTTAGAA" (Wang *et al.*, 1995) but MARs can also be defined by the ATC rule (Tikhonov, *pers. commun.*), which highlights the significance of structural characteristics such as DNA bending and a narrow minor groove due to oligo (dA) tracts. MARs have been shown to increase transcriptional activity of a linked gene and to confer position-independent, copy-number dependent expression in stably transfected cells (Loc & Stratling, 1988; Mlynarova *et al.*, 1995).

The impact of MFP1 protein on the binding of MARs has direct implications on both local and genome wide chromatin structure and is directly related to the control of transcriptionally active regions. Local chromatin structure and its modification at early meiosis are important in the positioning and frequency of meiotic double-strand breaks in DNA that enable recombination in yeast (Ohta *et al.*, 1994; Wu & Lichten, 1994). Earlier studies (Chandley & McBeath, 1987; Raman & Nanda, 1986) have also described regions of the human genome where the chromatin undergoes conformational changes from mitosis to meiosis could encompass recombinational hot spots. The lack of condensation of early replicating chromosomal segments during pre-meiotic interphase could be a

pre-requisite for crossover at pachytene. This leads to the observation that homologous chromosome pairing occurs when the chromosomes are transcriptionally active, so pairing may be an inevitable consequence of the transcription of partially condensed chromosomes (Cook, 1997). It has been proposed that DNA-protein structures (like that for MFP1 and MARs), inherent to heterochromatin in *Drosophila*, could produce a self-complementary chromosome "landscape" that ensures partner recognition and alignment by "best-fit" mechanisms (Karpen *et al.*, 1996). Specific coiling patterns that could promote strict homologous chromosome pairing, showing apparent denser and weaker zones, presumably reflecting more or less condensed chromatin, were observed at stages before meiotic prophase in the homologous domains of wheat (Schwarzacher, 1997).

This highlights the potential impact of the MFP1-like protein on chromosome pairing and subsequent recombination as correlated with the *Ph2* phenotype.

5.4.2.2.4 ScII protein

A barley putative *ScII*-like EST shows homology primarily to the rice putative ScII-like protein (**Figure 43**) located within the region of rice chromosome 1 that is syntenous to the region deleted in the *ph2a* wheat mutant (**Figure 39**). This rice ScII-like protein in turn shows significant homology to the chicken ScII protein. The chicken ScII protein, otherwise known as a chromosomal scaffold protein, is a member of the large structural maintenance chromosome (SMC) family of proteins. SMCs are considered to be crucial modulators of chromosome architecture and are involved in chromosome condensation and segregation (Hirano & Mitchison, 1994; Saitoh *et al.*, 1994; Saka *et al.*, 1994; Strunnikov *et al.*, 1995), sister chromatid cohesion (Guacci *et al.*, 1997; Michaelis *et al.*, 1997), transcriptional repression (Chuang *et al.*, 1994) and possibly homologous recombination (Jessberger *et al.*, 1996; Lehmann *et al.*, 1995).

SMC proteins are putative ATPases and are characterised by five conserved structural features: an N-terminal globular domain with an NTP-binding motif

(Walker A type), two α-helical regions with the potential to form a coiled-coiled structure separated by a hinge, and a second globular domain in the C-terminal region harbouring the DA-box, a signature motif for the SMC family (Strunnikov *et al.*, 1993), which includes an NTP-binding motif (Walker B) (Walker *et al.*, 1982). Mutations at the DA-box or NTP-binding domains have shown that they are required for SMC function (Chuang *et al.*, 1994; Strunnikov *et al.*, 1993).

In vitro, SMC proteins can perform recombination-related activities such as DNA renaturation (Jessberger *et al.*, 1996; Sutani & Yanagida, 1997). The SMC protein RC-1 is a component of the bovine recombination complex and is involved in *in vitro* recombination, and the repair of gaps and deletions in DNA (Jessberger *et al.*, 1996). SMC proteins, Rad18 (*Saccharomyces pombe*), RHC18 (*S. cerevisiae*), MIM (*A. thaliana*) are involved in the recombinational repair pathway. Although the exact role of SMC–like proteins in DNA repair is general and in recombination in particular, remains to be determined, it is thought that the structural features of chromatin conferred by SMCs are required for recognition of DNA damage and the recruitment of repair complexes. Data on *Arabidopsis* MIM SMC-like protein suggests that chromatin remodelling may be a crucial step in regulating *in vivo* accessibility of chromosomal DNA for recombination in plants (Mengiste *et al.*, 1999).

The wheat orthologue(s) to both the barley *ScII*-like EST and rice putative ScII-like protein could represent the *Ph2* gene or be functionally important in the *Ph2* phenotype. This suggestion is based on homology of both the barley and rice sequences to a representative member of the SMC family (chicken chromosomal scaffold protein). Some members of the SMC family have been implicated in homologous recombination and chromosome structure. Secondly, the wheat orthologue(s) to the barley and rice *ScII*-like sequences are likely to be located on chromosome 3DS within the region deleted in the *ph2a* mutant (**Figure 39**). Proposed function of SMC proteins would complement the idea that *Ph2* influences homeologous chromosome pairing and recombination.

5.4.2.2.5 Damage-specific DNA binding protein

The rare human hereditary disease, *Xeroderma pigmentosum* (XP), is characterised biochemically by defective nucleotide excision repair, which manifests clinically as sensitivity to ultraviolet light and a high incidence of skin cancer. A putative rice "damage-specific DNA binding protein" shows homology to the human DDB2 (*p48*) protein (**Figure 44**), which is part of a complex that binds specifically to damaged DNA (Nichols *et al.*, 2000). In addition to its high affinity for UV-damaged DNA, the DDB complex also recognises a wide spectrum of DNA lesions (Carew & Feldberg, 1985; Chu & Chang, 1988; Feldberg, 1980; Hirschfeld *et al.*, 1990; Payne & Chu, 1994; Protic *et al.*, 1989). It has been proposed that the DDB complex has a DNA repair function, possibly acting in damage recognition prior to nucleotide excision repair (Liu *et al.*, 2000). Since the "damage specific DNA binding protein" may have a recognition role in DNA repair, it could be involved in detecting and resolving mis-pairing in heteroduplex DNA during the process of chromosome pairing and/or recombination. Effective disruption of such a protein could result in pairing of divergent DNA sequences (homeologous chromosome pairing) or allowing mispairs in the recombination intermediates (homeologous recombination). Consequently, the wheat orthologue to the rice, "damage specific DNA binding protein" could represent the *Ph2* gene(s).

5.4.2.2.6 WD40 repeat like protein

The large family of proteins that contain the WD repeat are found in all eukaryotes and are implicated in a wide variety of crucial functions. One group of WD repeat proteins are involved in RNA-processing complexes with some members forming part of snRNP particles (Bjorn *et al.*, 1989). Another group comprises of transcriptional regulators, including the TFIID subunit of the TATA-box-binding complex (Hoey *et al.*, 1993; Keleher *et al.*, 1992; Williams *et al.*, 1991). Others play roles in the regulation of vesicle formation and vesicular trafficking (Pryer *et al.*, 1993), control various aspects of cell division or regulate sulfur metabolism in fungi. More importantly, some WD repeat proteins have been found to be involved in cyto-skeleton assembly and mitotic-spindle formation (de Hostos *et al.*, 1991; Vaisman *et al.*, 1995). The WD-repeat

'propeller-like structure' is a conserved motif with 4 to 16 repeats in any single protein. If there is a common functional theme, it appears to be that the WD-repeat propeller structure creates a stable platform that can form complexes reversibly with several proteins, thus coordinating sequential and/or simultaneous interactions involving several sets of proteins (Smith *et al.*, 1999). Although the WD repeat proteins are a large family, there is the possibility that the wheat orthologue to the putative rice WD-40 protein (**Figure 45**) represents the *Ph2* gene(s).

5.4.2.2.7 Other predicted genes

Previously described as *Wm5.12-tap*, the *WM3* messenger transcript is specifically and highly expressed in wheat floral tissue (carpels and anthers) and preferentially in the tapetum of meiotic anthers. The complete sequence of the cDNA revealed that it was similar to non-specific plant lipid transfer proteins (LTPs) (Letarte, 1996). This group of extracellularly secreted proteins has been found in several aerial organs of diverse plant species and more specifically in the tapetum of both *Nicotiana* sp. (Crossland and Tuttle *pers. commun.*) and *Brassica* sp. (Foster *et al.*, 1992). Although unlikely to represent the *Ph2* gene or be functionally important in the *Ph2* phenotype, the WM3 protein may be important in pollen mother cell development through meiosis.

5.4.2.3 Possible role for multiple Ph2 *candidate genes in maintaing strict diploid-like behaviour*

Interestingly, of the nine potential meiosis genes, four are likely to be involved in recombination and two of which may be specifically involved in heteroduplex DNA pairing and recombination (Asy1 and TaMSH7 proteins). Of the remaining two proteins that are likely to be involved in the recombination process, the SMC-like protein, ScII, may induce an altered chromatin structure, required for the recognition by damage-specific DNA binding proteins of mis-pairing in heteroduplex DNA. Chromatin structural change and recognition of mispairing is a likely precursor for recruitment of repair complexes such as TaMSH7 of the mismatch repair system. Local and genome wide structure of the chromatin (possible role of MFP1 protein) is also likely to be crucial for the pairing process.

Each of these proteins highlights a potential structural role of the *Ph2* gene product(s) in maintaining a strict diploid-like behaviour in allohexaploid wheat. However the WD40 repeat-like protein may temporally control meiotic development. Interestingly, functionality is not likely to be exclusive for all described proteins. For example, the Asy1-like protein is likely to have a role in DNA repair/synthesis but may also form an important structural component of the synaptonemal complex.

The clustering of these potential meiotic genes in both wheat and rice is intriguing as it suggests that structural co-localisation of meiosis genes may be important for their function. Apparent clustering of functionally related genes within the *Ph2* region shows similarity to the *Dopa decarboxylase* (*Ddc*) gene cluster of *Drosophila*, which is responsible for cuticle morphology (Wright, 1996). The function of the clustered meiosis genes within the *ph2a* deletion region may be dependent on a common coordinate regulatory mechanism. This has been suggested for the cluster of neurally expressing genes in the 14D region of the *Drosophila* genome. These genes have been implicated in the common physiological process of neuron Na+ transport (Hong & Ganetzky, 1996). The structural importance of this region would agree with the idea that as chromatin continually condenses from interphase to metaphase I, it would be expected that less of the chromatin is likely to be transcriptionally active due to the inaccessibility of the transcription enzymes to the euchromatic regions of DNA. Those genes that are essential for the processes of meiosis have to maintain transcriptional activity and therefore they are likely to be clustered in regions of the genome that are structurally different for this reason. This study tentatively supports the idea that the region deleted in the *ph2a* wheat mutant is one such structurally important region.

5.5 Conclusions

The suggestion that *Ph1* and *Ph2* bring about diploidisation of allohexaploid wheat by different mechanisms has important implications on interpretation of the comparative mapping of the *Ph2* locus. It was specifically suggested *Ph2*

affects synaptic progression, probably in a similar way to a diploid species, and that *Ph2* itself may not represent a pairing homeologous (*Ph*) locus but a synaptic (*Syn*) locus (Martinez *et al.*, 2001).

The identification of a rice *Asy1* orthologue within the rice chromosome 1 region homologous to the region deleted in the *ph2a* mutant supports the findings of Martinez *et al.* (2001). This tentatively suggests that the wheat orthologue to *Asy1* plays an important role in synaptic progression contributing to the wheat *Ph2* phenotype. However, the identification of *TaMSH7* within the region deleted in the wheat *ph2a* mutant can account for the homoeologous chromosome recombination aspect of the *Ph2* phenotype. This disagrees with the suggestion by Martinez *et al.* (2001) that *Ph2* represents a synaptic locus as opposed to a pairing homoeologous (*Ph*) locus.

Further study through transformation and protein localisation of both the wheat Asy1 homologue and TaMSH7 will help elucidate the role each of these genes plays in the *Ph2* phenotype.

6 Chapter

GENERAL DISCUSSION

Although studies in budding yeast *S. cereviseae* have lead to the identification and cloning of a large number of meiosis-specific genes, little is known about their counterparts in higher eukaryotes. This is particularly true in plants, where in many cases there is the added complexity of polyploidy. The ability of allohexaploid *T. aestivum* to act as a diploid at meiosis poses special questions related to the mechanisms that ensure correct homologous chromosome pairing at early prophase of meiosis 1.

This study illustrates a systematic approach to the analysis of *T. aestivum* genes likely to be involved in meiosis, in particular, the genes that could represent the *Ph2* gene(s). Based on the results of this study and the aim to isolate the *Ph2* gene(s), a series of new research opportunities can be considered.

6.1 Future direction

The *WM1* gene family can be further investigated from two perspectives; firstly in relation to its possible role in floral development leading to meiosis and secondly in the potential role of these genes in the perception of pathogen attack. Protein immunolocalisation of individual WM1 gene products within floral tissues may clarify the role of these genes. GFP reporter constructs using *WM1* gene promoter regions could help identify regulatory control of these genes whilst GFP protein fusions will validate predictions of the subcellular localisation of the WM1 proteins and may confirm the receptor-like function of these type Ia plasma membrane anchored LRR proteins.

WM5 transgenic lines also need further study. Lines expressing both reporter and introduced *WM5* genes require more detailed phenotypic and reporter gene analysis. Suspected phenotypic changes in apical meristems and meiotic

development need to be assessed closely. Further reporter gene transgenics will help define the structure of the full-length promoter segment of the *WM5* gene.

Comparative mapping among the grasses identified two strong candidates for the *Ph2* gene(s). However, wheat ESTs identified from the homologous rice sequence to the region deleted in the *ph2a* mutant need to be mapped back to the wild-type Chinese Spring to confirm localition within the region deleted in the *ph2a* mutant. Each wheat EST also needs to be assessed for temporal expression pattern during meiosis and differential expression between wild-type and mutant lines. Particular attention should be paid to the candidate *Ph2* gene(s) *Asy1* and mismatch repair protein homologues. An *Arabidopsis* mutant is available for *Asy1* and complementation with the wheat homologue should be a prelude to complementation of wheat mutants, *ph2a* and *ph2b*. Similarly, complementation should be attempted with *TaMSH7*.

6.2 Wheat genes have multiple levels of complexity

Analysis of the *WM1* and *WM5* gene families, as well as comparative mapping among the grasses have confirmed that wheat genes are more complex than expected in both their structural organisation and transcriptional control relative to rice.

6.2.1 Structural organisation of the *Ph2* region

The availablility of mutants in this study proved important. The deletion mutant *ph2a* aided in comparative mapping across the grasses and led to the identification of candidate *Ph2* gene(s). This study identified what appears to be a meiotic gene cluster. Such a complex cluster of structurally different genes, but likely to influence the same developmental process, has not previously been described in plants and further research may reveal whether the observed clustering is coincidental or functionally important.

Additionally, structural analysis of the complex *WM1* gene family revealed that gene content across homeologous genomes differs greatly. Although this result

has only been shown with the *WM1* gene family, it may apply to other genes within the region of *Ph2*. It is clear that the three homeologous genomes have diverged in gene content but whether this occurred before or after the polyploidisation events is unclear. Nevertheless analysis of the 220 Kb of *T. tauschii* DNA provided clues of the likely evolutionary mechanisms causing divergence in WM1 gene content. Studying this 220 Kb region also revealed several novel transposable elements, highlighting the complex organisation of the wheat genome.

The complexity of the *Ph2* region and the *WM1* gene cluster raises many questions with respect to polyploidisation and meiosis:

What is the significance of the apparent clustering of meiotic genes in grasses (barley, wheat and rice) and what was the impact of polyploidisation on these genes? The observed diversity in gene content between homeologues in allohexaploid *T. aestivum* raises the question; did this variation arise before (as appears to have been the case for the WM1 family) or after allopolyploidisation? Is this variation observed only in the physical presence/absence of genes across the homeologues, as shown with the *WM1* genes, or is variation observed in the functionality of the genes across the homeologues? If this diversity was present before polyploidisation, did this pre-dispose diploid progenitors to allopolyploidisation? What was the selective advantage derived from the apparent clustering of meiotic genes and how did it occur?

These are just some of the questions that may be answered through further understanding of genome changes during polyploidisation.

6.2.2 Transcriptional control of meiotic wheat genes

In addition to the genome complexity revealed through the study of the *WM1* gene family, transcriptional complexity was found for the *WM5* genes. *WM5* shows complex temporal control in transcription. Strict temporal control of transcription at meiosis for specific genes is expected to be critical for the progression of meiocytes through to haploid, pollen cells. Although *WM5* may not represent the *Ph2* gene(s) its preferential expression at meiosis suggests

that a search for the *Ph2* gene(s) through meiosis specificity alone may not yield the *Ph2* gene(s). The criteria for identifying *Ph2* candidates needs to be re-assessed and it will be important to analyse both the temporal transcription and the protein localisation for genes of interest.

Through the structural, functional and comparative analysis of genes within the grasses, this study has successfully increased awareness of the complexity associated with genome structure and gene expression during meiosis.

6.3 Cereal genome and wheat breeding

A detailed understanding of genome changes, transcription and post-translational modifications of gene products involved in chromosome pairing and recombination will help enhance cereal breeding programs. This knowledge will specifically help plant breeders develop meiotically stable synthetic polyploids and wheat lines that can be used for alien gene introgression. Through the use of such synthetic polyploids and wheat lines, cereal-breeding programs will benefit from the greater number of landraces and wild relatives available for alien gene introgression into commercial wheat cultivars. It is also expected that the time required for these breeding strageties will be greatly reduced, allowing breeders to react more quickly to consumer demand in the global market place.

BIBLIOGRAPHY

Abe M., Takahashi T., and Komeda Y. (1999). Cloning and characterization of an L1 layer-specific gene in *Arabidopsis thaliana*. *Plant Cell Physiology* **40:** 571-80.

Abe M., Takahashi T., and Komeda Y. (2001). Identification of a *cis*-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant Journal* **26:** 487-94.

Abramic M., Levine A. S., and Protic M. (1991). Purification of an ultraviolet-inducible, damage-specific DNA-binding protein from primate cells. *Journal of Biological Chemistry* **266:** 22493-500.

Ade J., Belzile F., Philippe H., and Doutriaux M. P. (1999). Four mismatch repair paralogues coexist in *Arabidopsis thaliana*: *AtMSH2*, *AtMSH3*, *AtMSH6-1* and *AtMSH6-2*. *Molecular & General Genetics* **262:** 239-49.

Ahn S., Anderson J. A., Sorrells M. E., and Tanksley S. D. (1993). Homoeologous relationships of rice, wheat and maize chromosomes. *Molecular & General Genetics* **241:** 483-90.

Alani E., Reenan R. A., and Kolodner R. D. (1994). Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. *Genetics* **137:** 19-39.

Albini S. M., and Jones G. H. (1987). Synaptonemal complex spreading in *Allium cepa* and *A. fistulosum*. I. The initiation and sequence of pairing. *Chromosoma* **95:** 324-338.

Altschul S. F., Madden T. L., Schaffer A. A., Zhang J. H., Zhang Z., Miller W., and Lipman D. J. (1997). Gapped blast and psi-blast - a new generation of protein database search programs. *Nucleic Acids Research* **25:** 3389-3402.

Ansley H. R. (1958). Histones of mitosis and meiosis in *Loxa flavicolis* (hemipteran). *Journal of Biophysical & Biochemical Cytology* **4:** 59-62.

Apweiler R., Attwood T. K., Bairoch A., Bateman A., Birney E., Biswas M., Bucher P., Cerutti L., Corpet F., Croning M. D., Durbin R., Falquet L., Fleischmann W., Gouzy J., Hermjakob H., Hulo N., Jonassen I., Kahn D., Kanapin A., Karavidopoulou Y., Lopez R., Marx B., Mulder N. J., Oinn T. M., Pagni M., and Servant F. (2001). The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Research* **29:** 37-40.

Aragon-Alcaide L., Reader S., Miller T., and Moore G. (1997). Centromeric behaviour in wheat with high and low homeologous chromosomal pairing. *Chromosoma* **106:** 327-333.

Aravind L., and Koonin E. V. (1998). The HORMA domain: a common structural denominator in mitotic checkpoints, chromosome synapsis and DNA repair. *Trends in Biochemical Sciences* **23:** 284-6.

Arumuganathan K., and Earle E. D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9:** 208-219.

Avivi L., and Feldman M. (1973a). Mechanism of non-random chromosome placement in common wheat. *In* "Proceedings of the 4th International Wheat Genetics Symposium", pp. 629-633, Columbia, Missouri, USA.

Avivi L., and Feldman M. (1973b). The mechanism of somatic association in common wheat, *Triticum aestivum* L. IV. Further evidence for modification of spindle tubulin through the somatic association genes as measured by vinblastine binding. *Genetics* **73:** 379-385.

Baker S. M., Bronner C. E., Zhang L., Plug A. W., Robatzek M., Warren G., Elliott E. A., Yu J., Ashley T., Arnheim N., Flavell R. A., and Liskay R. M.

(1995). Male mice defective in the DNA mismatch repair gene *PMS2* exhibit abnormal chromosome synapsis in meiosis. *Cell* **82:** 309-19.

Baulcombe D. C., Chapman S., and Cruz S. S. (1995). Jellyfish green fluorescent protein as a reporter for virus infections. *Plant Journal* **7:** 1045-1053.

Becker D. (1999). **(personal communication)**.

Becker D., Brettschneider R., and Lorz H. (1994). Fertile transgenic wheat from microprojectile bombardment of scutellar tissue. *Plant Journal* **5:** 299-307.

Becraft P. W. (1998). Receptor kinases in plant development. *Trends in Plant Science* **3:** 384-388.

Benavente E., and Orellana J. (1991). Chromosome differentiation and pairing behavior of polyploids: an assessment on preferential metaphase I associations in colchicine- induced autotetraploid hybrids within the genus *Secale*. *Genetics* **128:** 433-42.

Benavente E., Orellana J., and Fernandezcalvin B. (1998). Comparative analysis of the meiotic effects of wheat *ph1b* and *ph2b* mutations in wheat X rye hybrids. *Theoretical & Applied Genetics* **96:** 1200-1204.

Benghezal M., Wasteneys G . O ., a nd J ones D . A . (2000). T he C -terminal dilysine motif confers endoplasmic reticulum localization to type I membrane proteins in plants. *Plant Cell* **12:** 1179-201.

Bennett M. D. (1984). Premeiotic events and meiotic chromosome pairing. *Symposia of the Society for Experimental Biology* **38:** 87-121.

Bennett M. D., Finch R. A., Smith J. B., and Rao M. K. (1973). The time and duration of female meiosis in wheat, rye and barley. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **183:** 301-319.

Bennett M. D., and Smith J. B. (1979). The effect of colchicine on fibrillar material in wheat meiocytes. *Journal of Cell Science* **38:** 33-47.

Bennett M. D., and Smith J. B. (1991). Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **334:** 309-345.

Berezney R., and Coffey D. S. (1974). Identification of a nuclear protein matrix. *Biochemical & Biophysical Research Communications* **60:** 1410-7.

Birch R. G., and Franks T. (1991). Development and optimisation of microprojectile systems for plant genetic transformation. *Australian Journal of Plant Physiology* **18:** 453-469.

Bjorn S. P., Soltyk A., Beggs J. D., and Friesen J. D. (1989). *PRP4* (RNA4) from *Saccharomyces cerevisiae*: its gene product is associated with the U4/U6 small nuclear ribonucleoprotein particle. *Molecular & Cellular Biology* **9:** 3698-709.

Blake N. K., Lehfeldt B. R., Hemphill A., Shan X., and Talbert L. E. (1998). DNA sequence analysis suggests a monophyletic origin of the wheat B genome. *In* "Proceedings of the 9th International Wheat Genetics Symposium" (A. E. Slinkard, Ed.), pp. 14-16, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Bonierbale M., Plaisted R. L., and Tanksley S. D. (1988). RFLP maps of potato and tomato based on a common set of clones reveal modes of chromosomal evolution. *Genetics* **120:** 1095-1103.

Bower R., Elliott A. R., Potier B. A. M., and Birch R. G. (1996). High-efficiency, microprojectile-mediated cotransformation of sugarcane, using visible or selectable markers. *Molecular Breeding* **2:** 239-249.

Braun D. M., and Walker J. C. (1996). Plant transmembrane receptors - New pieces in the signaling puzzle. *Trends in Biochemical Sciences* **21:** 70-73.

Burge C., and Karlin S. (1997). Prediction of complete gene structures in human genomic DNA. *Journal of Molecular Biology* **268:** 78-94.

Burset M., and Guigo R. (1996). Evaluation of gene structure prediction programs. *Genomics* **34:** 353-67.

Cameriniotero R. D., and Hsieh P. (1995). Homologous recombination proteins in prokaryotes and eukaryotes. *Annual Review of Genetics* **29:** 509-552.

Carew J. A., and Feldberg R. S. (1985). Recognition of a cytosine base lesion by a human damage-specific DNA binding protein. *Nucleic Acids Research* **13:** 303-15.

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

Casper S. J., and Holt C. A. (1996). Expression of the green fluorescent protein-encoding gene from a tobacco mosaic virus-based vector. *Gene* **173:** 69-73.

Ceoloni C., and Donini P. (1993). Combining mutations for the two homoeologous pairing suppressor genes *Ph1* and *Ph2* in common wheat and in hybrids with alien *Triticeae*. *Genome* **36:** 377-386.

Chaboute M. E., Chaubet N., Gigot C., and Philipps G. (1993). Histones and histone genes in higher plants: structure and genomic organization. *Biochimie* **75:** 523-31.

Chambers S. R., Hunter N., Louis E. J., and Borts R. H. (1996). The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. *Molecular & Cellular Biology* **16:** 6110-20.

Chandley A . C. (1986). A model for effective pairing and recombination at meiosis of early replicating sites (R-bands) along chromosomes. *Human Genetics* **72:** 50-57.

Chandley A. C., and McBeath S. (1987). DNase I hypersensitive sites along the XY bivalent at meiosis in man include the XpYp pairing region. *Cytogenetics & Cell Genetics* **44:** 22-31.

Chen W., and Jinks-Robertson S. (1998). Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. *Molecular & Cellular Biology* **18:** 6525-37.

Cheng M., Fry J. E., Pang S. Z., Zhou H. P., Hironaka C. M., Duncan D. R., Conner T. W., and Wan Y. C. (1997). Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. *Plant Physiology* **115:** 971-980.

Cheung W. Y., and Gale M. D. (1990). The isolation of high molecular weight DNA from wheat, barley and rye for analysis by pulse-field gel electrophoresis. *Plant Molecular Biology* **14:** 881-8.

Chiu W., Niwa Y., Zeng W., Hirano T., Kobayashi H., and Sheen J. (1996). Engineered GFP as a vital reporter in plants. *Current Biology* **6:** 325-30.

Christou P., Ford T. L., and Kofron M. (1991). Production of transgenic rice (*Oryza sativa* L.) plants from agronomically *indica* and *japonica* varieties via electrical discharge particle acceleration of exogenous DNA i nto i mmature zygotic embryos. *Biotechnology* **9:** 957-962.

Chu G., and Chang E. (1988). *Xeroderma pigmentosum* group E cells lack a nuclear factor that binds to damaged DNA. *Science* **242:** 564-7.

Chuang P. T., Albertson D. G., and Meyer B. J. (1994). DPY-27:a chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through association with the X chromosome. *Cell* **79:** 459-74.

Clark S. E., Running M. P., and Meyerwitz E. M. (1995). *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. *Development* **121:** 2057-2067.

Clark S. E., Williams R. W., and Meyerowitz E. M. (1997). The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* **89:** 575-585.

Cook P. R. (1997). The transcriptional basis of chromosome pairing. *Journal of Cell Science* **110:** 1033-40.

Cuadrado C., Romero C., and Lacadena J. R. (1991). Meiotic pairing control in wheat-rye hybrids. I. Effects of different wheat chromosome arms of homoeologous groups 3 and 5. *Genome* **34:** 72-75.

Culligan K. M., and Hays J. B. (2000). *Arabidopsis MutS* homologs-*AtMSH2*, *AtMSH3*, *AtMSH6*, and a novel *AtMSH7*-form three distinct protein heterodimers with different specificities for mismatched DNA. *Plant Cell* **12:** 991-1002.

Cunado N., Callejas S., Garcia M. J., Fernandez A., and Santos J. L. (1996). Chromosome pairing in the allotetraploid *Aegilops b iuncialis* a nd a t riploid intergeneric hybrid. *Genome* **39:** 664-670.

Dangl J. L. (1995). Piece de Resistance: novel classes of plant disease resistance genes. *Cell* **80:** 363-6.

Datta A., Adjiri A., New L., Crouse G. F., and Jinks Robertson S. (1996). Mitotic crossovers between diverged sequences are regulated by mismatch repair proteins in *Saccaromyces cerevisiae*. *Molecular & Cellular Biology* **16:** 1085-93.

Davis L., and Smith G. R. (2001). Meiotic recombination and chromosome segregation in *Schizosaccharomyces pombe*. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 8395-402.

de Barros E. G., Tingey S., and Rafalski A. (1999). Sequence characterisation of hypervariable region in the soybean genome: leucine-rich repeats and SSRs.

de Hostos E. L., Bradtke B., Lottspeich F., Guggenheim R., and Gerisch G. (1991). Coronin, an actin binding protein of *Dictyostelium discoideum* localized to cell surface projections, has sequence similarities to G protein beta subunits. *EMBO Journal* **10:** 4097-104.

de Wind N., Dekker M., Berns A., Radman M., and te Riele H. (1995). Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* **82:** 321-30.

de Wit P. J. G. M., and Joosten M. H. A. J. (1999). Avirulence and resistance genes in the *Cladosporium fulvum*-tomato interaction. *Current Opinion in Microbiology* **2:** 368-373.

Del Fosse F. E., and Church K. (1981). Presynaptic chromosome behaviour in *Lilium*. I. Centromere orientation and movement during premeiotic interphase in *Lilium speciosum* cv. Rosemede. *Chromosoma* **81:** 701-716.

Dixon M. S., Jones D. A., Keddie J. S., Thomas C. M., Harrison K., and Jones J. D. G. (1996). The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* **84:** 451-459.

DNASTAR I. (1997). DNASTAR, DNASTAR Inc., Madison, WI 53715 USA.

Dong C., Thomas S. W., Guo R., and Langridge P. (2001a). Isolation and characterisation of *WM5*, a gene preferentially expressed during meiosis in wheat. *Plant Journal* **(submitted)**.

Dong C., Whitford R., and Langridge P. (2001b). A mismatch repair gene linked to the *Ph2* locus in wheat. *Genome* **(in press)**.

Driscoll C. J. (1972). Genetic suppression of homoeologous chromosome pairing in hexaploid wheat. *Canadian Journal of Genetics & Cytology* **14:** 39-42.

Driscoll C. J. (1973). Minor genes affecting homoeologous pairing in hybrids between wheat and related genera. *Genetics* **74s:** 566.

Du H., and Klessig D. F. (1997). Identification of a soluble, high-affinity salicylic acid binding protein in tobacco. *Plant Physiology* **113:** 1319-1327.

Edelmann W., Cohen P. E., Kane M., Lau K., Morrow B., Bennett S., Umar A., Kunkel T ., C attoretti G ., C haganti R ., P ollard J . W ., K olodner R . D ., a nd Kucherlapati R. (1996). Meiotic pachytene arrest in *MLH1*-deficient mice. *Cell* **85:** 1125-34.

Ellis J., and Jones D. (1998). Structure and function of proteins controlling strain-specific pathogen resistance in plants. *Current Opinion in Plant Biology* **1:** 288-293.

Feinberg A. P., and Vogelstein B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132:** 6-13.

Feldberg R. S. (1980). On the substrate specificity of a damage-specific DNA binding protein from human cells. *Nucleic Acids Research* **8:** 1133-43.

Feldman M. (1966). The effect of chromosomes 5B, 5D and 5A on chromosomal pairing in *Triticum aestivum. Proceedings of the National Academy of Sciences of the United States of America* **55:** 1447-1453.

Feldman M. (1968). Regulation of somatic association and meiotic pairing in common wheat. *In* "Proceedings of the Third International Wheat Genetics Symposium" (K. W. Finlay, and K. W. Shepherd, Eds.), pp. 31-40, Australian Academy of Science, Canberra.

Feldman M . (1993). C ytogenetic a ctivity and mode of action of the pairing homoeologous (*Ph1*) gene of wheat. *Crop Science* **33:** 894-897.

Feldman M., and Avivi L. (1973). The pattern of chromosomal arrangement in nuclei of common wheat and its genetic control. *In* "Proceedings of the 4th International Wheat Genetics Symposium", pp. 675-684, Columbia, Missouri, USA.

Feldman M., and Avivi L. (1988). Genetic control of bivalent pairing in common wheat: the mode of *Ph1* action. *In* "Kew Chromosome Conference. 3rd Proceedings Chromosome Conference." (P. E. Brandham, Ed.), pp. 269-279, HMSO, Jodress Lab. Royal Botanical Gardens, Kew, England.

Feuillet C., Penger A., Gellner K., Mast A., and Keller B. (2001). Molecular evolution of receptor-like kinase genes in hexaploid wheat. Independent evolution of orthologs after polyploidization and mechanisms of local rearrangements at paralogous loci. *Plant Physiology* **125:** 1304-13.

Flavell R. B., Bennett M. D., Seal A. G., and and Hutchinson J. (1987). Chromosome structure and o rganisation. *I n* " Wheat B reeding: i ts s cientific basis" (F. G. H. Lupton, Ed.), pp. 211-268, Chapman and Hall Ltd, London.

Foster G. D., Robinson S. W., Blundell R. P., Robert M. R., Hodge R., Draper J., and Scott R. J. (1992). A *Brassica napus* mRNA encoding a protein homologous to phospholipid transfer proteins, is expressed specifically in the tapetum and developing microspores. *Plant Science* **84:** 187-192.

Franckowiak J . (1997). Revised linkage maps for morphological markers in barley, *Hordeum vulgare*. *Barley Genetics Newsletter* **26:** 9-21.

Franklin A. E., McElver J., Sunjevaric I., Rothstein R., Bowen B., and Cande W. Z. (1999). Three-dimensional microscopy of the *rad51* recombination protein during early meiotic prophase. *Plant Cell* **11:** 809-824.

Fromm M. E., Morrish F., Armstrong C., Williams R., Thomas J., and Klein T. M. (1990). Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Biotechnology* **8:** 833-9.

Gale M. D., and Miller T. E. (1987). The introduction of alien genetic variation into wheat. *In* "Wheat Breeding: Its Scientific Basis" (F. G. H. Lupton, Ed.), pp. 173-210, Chapman and Hall Ltd., London.

Galitski T., Saldanha A. J., Styles C. A., Lander E. S., and Fink G. R. (1999). Ploidy regulation of gene expression. *Science* **285:** 251-254.

Game J. C., Sitney K. C., Cook V. E., and Mortimer R. K. (1989). Use of a ring chromosome and pulsed field gels to study interhomologue recombination, double-strand DNA breaks and sister chromatid exchange in yeast. *Genetics* **123:** 695-713.

Gay N. J., Packman L. C., Weldon M. A., and Barna J. C. J. (1991). A leucine-rich repeat peptide derived from the *Drosophila* toll receptor forms extended filaments with β-sheet structure. *FEBS Letters* **291:** 87-91.

Gordon D., Abajian C., and Green P. (1998). CONSED - a graphical tool for sequence finishing. *PCR Methods & Applications* **8:** 195-202.

Gordon-Kamm W. J., Spencer T. M., Mangano M. L., Adams T. R., Daines R. J., Start W. G., O'Brian J. V., Chambers S. A., Adams J. W. R., Willets N. G., Rice T. B., Mackey C. J., Krueger W., Kausch A. P., and Lemaux P. G. (1990). Transformation of maize cells and regeneration of fertile ransgenic plants. *Plant Cell* **2:** 603-618.

Grant M. R., Godiard L., Straube E., Ashfield T., Lewald J., Sattler A., Innes R. W., and Dangl J. L. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* **269:** 843-6.

Grimes B., Hallick R., Williams K., Wells M., Lapointe M., Ryan C., Wagenheim M., and Hersoff R. (2001). The Biology Project, University of Arizona.

Guacci V., Koshland D., and Strunnikov A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of *MCD1* in *S. cerevisiae*. *Cell* **91:** 47-57.

Guidet F., Rogowsky P., Taylor C., Song W., and Langridge P. (1991). Cloning and characterisation of a new rye-specific repeated sequence. *Genome* **34:** 81-87.

Hajdukiewicz P., Svab Z., and Maliga P. (1994). The small, versatile Ppzp family of *Agrobacterium* binary vectors for plant transformation. *Plant Molecular Biology* **25:** 989-994.

Harder P. A., Silverstein R. A., and Meier I. (2000). Conservation of matrix attachment region-binding filament-like protein 1 among higher plants. *Plant Physiology* **122:** 225-34.

Harushima Y., Yano M., Shomura P., Sato M., Shimano T., Kuboki Y., Yamamoto T., Lin S. Y., Antonio B. A., Parco A., Kajiya H., Huang N., Yamamoto K., Nagamura Y., Kurata N., Khush G. S., and Sasaki T. (1998). A high-density rice genetic linkage map with 2275 markers using a single F-2 population. *Genetics* **148:** 479-494.

Haseloff J., Siemering K. R., Prasher D. C., and Hodge S. (1997). Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 2122-2127.

Hasenkampf C. A. (1996). The synaptonemal complex - The chaperone of crossing over. *Chromosome Research* **4:** 133-140.

He D. C., Nickerson J. A., and Penman S. (1990). Core filaments of the nuclear matrix. *Journal of Cell Biology* **110:** 569-80.

He D. G., Mouradov A., Yang Y. M., Mouradova E., and Scott K. J. (1994). Transformation of wheat (*Triticum aestivum* L) through electroporation of protoplasts. *Plant Cell Reports* **14:** 192-196.

Henikoff J. G., Greene E. A., Pietrokovski S., and Henikoff S. (2000). Increased coverage of protein families with the blocks database servers. *Nucleic Acids Research* **28:** 228-30.

Henikoff S., Henikoff J. G., and Pietrokovski S. (1999). Blocks+: a non-redundant database of protein alignment blocks derived from multiple compilations. *Bioinformatics* **15:** 471-9.

Higo K., U gawa Y ., I wamoto M ., a nd K orenaga T . (1999). P lant c is-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27:** 297-300.

Hirano T., and Mitchison T. J. (1994). A heterodimeric coiled-coil protein required for mitotic chromosome condensation *in vitro*. *Cell* **79:** 449-458.

Hirschfeld S ., L evine A . S ., O zato K ., a nd Protic M. (1990). A constitutive damage-specific DNA-binding protein is synthesized at higher levels in UV-irradiated primate cells. *Molecular & Cellular Biology* **10:** 2041-8.

Hobolth P. (1981). Chromosome pairing in allohexaploid wheat var. Chinese Spring. Transformation of multivalents into bivalents, a mechanism for exclusive bivalent formation. *Carlsberg Research Communication* **46:** 129-173.

Hoey T., Weinzierl R. O., Gill G., Chen J. L., Dynlacht B. D., and Tjian R. (1993). Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* **72:** 247-60.

Hofmann K., Bucher P., Falquet L., and Bairoch A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Research* **27:** 215-9.

Hofmann K., and Stoffel W. (1993). TMbase - A database of membrane spanning proteins segments. *Biological Chemistry* **374:** 166.

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

Hollingsworth N. M., and Ponte L. (1997). Genetic interactions between *HOP1*, *RED1* and *MEK1* suggest that *MEK1* regulates assembly of axial element components during meiosis in the yeast *Saccharomyces cerevisiae*. *Genetics* **147:** 33-42.

Holm P. B. (1986). Chromosome pairing and chiasmata formation in allohexaploid wheat, *Triticum aestivum* analysed by spreading of meiotic nuclei. *Carlsberg Research Communication* **51:** 239-294.

Holm P. B. (1988a). Chromosome pairing and synaptonemal complex formation in allohexaploid wheat, monosomic for chromosome 5B. *Carlsberg Research Communication* **53:** 57-89.

Holm P. B. (1988b). Chromosome pairing and synaptonemal complex formation in allohexaploid wheat, nullisomic for chromosome 5B. *Carlsberg Research Communication* **53:** 91-110.

Holm P. B., and Wang X. (1988). The effect of chromosome 5B on synapsis and chiasma formation in wheat *Triticum aestivum* cv. Chinese Spring. *Carlsberg Research Communication* **53:** 191-208.

Hong C. S., and Ganetzky B. (1996). Molecular characterization of neurally expressing genes in the para sodium channel gene cluster of *Drosophila*. *Genetics* **142:** 879-892.

Hu W., and Cheng C. L. (1995). Expression of *Aequorea* green fluorescent protein in plant cells. *FEBS Letters* **369:** 331-334.

Hulbert S. H. (1997). Structure and evolution of the *rp1* complex conferring rust resistance in maize. *Annual Review of Phytopatholgy* **35:** 293-310.

Hunt M. D., and Ryals J. A. (1996). Systemic acquired resistance signal transduction. *Critical Reviews in Plant Sciences* **15:** 583-606.

Hunter N., Chambers S. R., Louis E. J., and Borts R. H. (1996). The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO Journal* **15:** 1726-33.

Hyatt D., Snoddy J., Schmoyer D., Chen G., Fischer K., Parang M., Vokler I., Petrov S., Locascio P., Olman V., Miriam Land, Shah M., and Uberbacher E. (2000a). GRAIL-EXP and the genome analysis toolkit. *The 13th Annual Cold Spring Harbor Meeting on Genome Sequencing & Biology*.

Hyatt D., Snoddy J., Schmoyer D., Chen G., Fischer K., Parang M., Vokler I., Petrov S., Locascio P., Olman V., Miriam Land, Shah M., and Uberbacher E. (2000b). Improved analysis and annotation tools for whole-genome computational annotation and analysis: GRAIL-EXP genome analysis toolkit and related analysis tools. *Genome Sequencing & Biology Meeting*.

Islam A. K. M. R., and Shepherd K. W. (1988). Induced pairing between wheat and barley chromosomes. *In* "Proceedings of the 7th International Wheat Genetics Symposium", pp. 309-314, Cambridge.

Ito S., Sakai A., Nomura T., Miki Y., Ouchida M., Sasaki J., and Shimizu K. (2001). A novel *WD40* repeat protein, *WDC146*, highly expressed during spermatogenesis in a stage-specific manner. *Biochemical & Biophysical Research Communications* **280:** 656-663.

Jenkins G., and Jimenez G. (1995). Genetic control of synapsis and recombination in *Lolium* amphidiploids. *Chromosoma* **104:** 164-168.

Jenkins G., White J., and Parker J. S. (1988). Elimination of multialents during meiotic prophase in *Scilla autumnalis*. II. Tetraploids. *Genome* **30:** 940-946.

Jessberger R ., R iwar B ., B aechtold H ., a nd A khmedov A . T. (1996). SMC proteins constitute two subunits of the mammalian recombination complex RC-1. *EMBO Journal* **15:** 4061-8.

Ji L. H. (1992). A study of meiosis in allohexaploid wheat: The molecular aspects. *In* "Department of Plant Science", Adelaide University, Adelaide: Australia.

Ji L. H., and Langridge P. (1994). An early meiosis cDNA clone from wheat. *Molecular & General Genetics* **243:** 17-23.

Jones D. A., Thomas C. M., Hammond-Kosack K. E., Balintkurti P. J., and Jones J. D. G. (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266:** 789-793.

Karpen G. H., Le M. H., and Le H. (1996). Centric heterochromatin and the efficiency of achiasmate disjunction in *Drosophila* female meiosis. *Science* **273:** 118-22.

Kawchuk L. M., Hachey J., Lynch D. R., Kulcsar F., van Rooijen G., Waterer D. R., Robertson A., Kokko E., Byers R., Howard R. J., Fischer R., and Prufer D. (2001). Tomato *Ve* disease resistance genes encode cell surface-like receptors. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 6511-5.

Keeney S., Giroux C. N., and Kleckner N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by *Spo11*, a member of a widely conserved protein family. *Cell* **88:** 375-384.

Keleher C. A., Redd M. J., Schultz J., Carlson M., and Johnson A. D. (1992). *Ssn6-Tup1* is a general repressor of transcription in yeast. *Cell* **68:** 709-19.

Kempanna C., and Riley R. (1962). Relationships between the genetic effects of deficiencies for chromosomes III and V on meiotic pairing in *Triticum aestivum*. *Nature* **195:** 1270-1273.

Kerber E. R. (1988). Telocentric mapping in wheat of the *Lr32* for resistance to leaf rust. *Crop Science* **28:** 178-179.

Kessman H., and Nordmeyer D. (1996). Plant activation - a new technology for disease control. *Agrow Supplement Autumn*: 23-27.

Kleckner N. (1996). Meiosis - how could it work. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 8167-8174.

Klein S. (1994). Choose your partner: chromosome pairing in yeast meiosis. *BioEssays* **16:** 869-871.

Kobe B., and Deisenhofer J. (1993). Crystal structure of *Porcine* ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **366:** 751-756.

Kobe B., and Deisenhofer J. (1994). The leucine-rich repeat - A versatile binding motif. *Trends in Biochemical Sciences* **19:** 415-421.

Koebner R. M. D., and Shepherd K. W. (1985). Induction of recombination between rye chromosome 1RL and wheat chromosomes. *Theoretical & Applied Genetics* **71:** 208-215.

Koziel G. M., Beland G. L., and Bowman C. (1993). Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Biotechnology* **1:** 194-200.

Lam E., Hanley-Bowdoin L., and Chua N. H. (1988). Characterization of a chloroplast sequence-specific DNA binding factor. *Journal of Biological Chemistry* **263:** 8288-93.

Lander E. S., Green P., Abrahamson J., Barlow A., Daly M. J., Lincoln S. E., and Newburg L. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1:** 174-81.

Langridge P., Karakousis A., Collins N., Kretschmer J., and Manning S. (1995). A consensus linkage map of barley. *Molecular Breeding* **1:** 389-395.

Laroche A., Frick M. M., Huel R., Nykiforuk C., Conner B., and Kuzyk A. (2000). Molecular identification of the wheat stripe rust resistance gene *Yr10*, the first full-length leucine zipper-nucleotide binding site-leucine-rich-repeat resistance gene in cereals, GenBank Accession Number AF149112 (Crop Sciences, Agriculture and Agri-Food Canada, 5403 1st Avenue S., PO Box 3000, Lethbridge, AB T1J 4B1, Canada).

Lehmann A. R., Walicka M., Griffiths D. J., Murray J. M., Watts F. Z., McCready S., and Carr A. M. (1995). The *rad18* gene of *Schizosaccharomyces pombe* defines a new subgroup of the SMC superfamily involved in DNA r epair. *Molecular & Cellular Biology* **15:** 7067-80.

Letarte J. (1996). Identification and characterisation of early meiotic genes in wheat. *In* "Department of Plant Science", Adelaide University, Adelaide.

Letourneur F., and Klausner R. D. (1992). A novel di-leucine motif and a tyrosine-based motif independently mediate lysosomal targeting and endocytosis of CD3 chains. *Cell* **69:** 1143-57.

Liharska T., Koornneef M., van Wordragen M., van Kammen A. B., and Zabel P. (1996). Tomato chromosome 6: effect of alien chromosomal segments on recombinant frequencies. *Genome* **39:** 485-491.

Lin H. X., Qian H. R., Zhuang J. Y., Lu J., Min S. K., Xiong Z. M., Huang N., and Zheng K. L. (1996). RFLP mapping of QTLs for yield and related characters in rice (*Oryza sativa* L). *Theoretical & Applied Genetics* **92:** 920-927.

Liu B., Vega J. M., and Feldman M. (1998). Differentiation of homoeologous chromosomes in polyploid wheat. *In* "9th International Wheat Genetics Symposium" (A. E. Slinkard, Ed.), pp. 70-71, University Extension Press, Saskatoon, Saskatchewan, Canada.

Liu W., Nichols A. F., Graham J. A., Dualan R., Abbas A., and Linn S. (2000). Nuclear transport of human DDB protein induced by ultraviolet light. *The Journal of Biological Chemistry* **275:** 21429-21434.

Loc P. V., and Stratling W. H. (1988). The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. *EMBO Journal* **7:** 655-64.

Loidl J. (1990). The initiation of meiotic chromosome pairing: The cytological view. *Genome* **33:** 759-778.

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

Lörz H., Becker D., and Lutticke S. (1998). Molecular wheat breeding by direct gene transfer. *Euphytica* **100:** 219-223.

Luo M. C., Dubcovsky J., and Dvorak J. (1996). Recognition of homeology by the wheat *Ph1* locus. *Genetics* **144:** 1195-1203.

Maguire M. P. (1977). Homologous chromosome pairing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **277:** 245-258.

Maguire M. P. (1983). Chromosome behaviour at premeiotic mitosis in maize. *The Journal of Heredity* **74:** 93-96.

Manly K. F. (1993). A Macintosh program for storage and analysis of experimental genetic mapping data. *Mammalian Genome* **4:** 303-13.

Marino C. L., Nelson J. C., Lu Y. H., Sorrells M. E., Leroy P., Tuleen N. A., Lopes C. R., and Hart G. E. (1996). Molecular genetic maps of the group 6 chromosomes of hexaploid wheat (*Triticum aestivum* L *Em Thell*). *Genome* **39:** 359-366.

Marks M. S., Woodruff L., Ohno H., and Bonifacino J. S. (1996). Protein targeting by tyrosine- and di-leucine-based signals: evidence for distinct saturable components. *Journal of Cell Biology* **135:** 341-54.

Marsischky G. T., Lee S., Griffith J., and Kolodner R. D. (1999). *Saccharomyces cerevisiae MSH2/6* complex interacts with Holliday junctions and facilitates their cleavage by phage resolution enzymes. *Journal of Biological Chemistry* **274:** 7200-6.

Martinez M., Cunado N., Carcelen N., and Romero C. (2001). The *Ph1* and *Ph2* loci play different roles in synaptic behaviour of hexaploid wheat *Triticum aestivum*. *Theoretical & Applied Genetics* **103:** 398-405.

Martinez-Perez E., Shaw P., Reader S., Aragon-Alcaide L., Miller T., and Moore G. (1999). Homologous chromosome pairing in wheat. *Journal of Cell Science* **112:** 1761-1769.

McElroy D., Blowers A. D., Jenes B., and Wu R. (1991). Construction of expression vectors based on rice actin 1 (Act1) 5' region for use in monocot transformation. *Molecular & General Genetics* **231:** 150-160.

McLeish J., and Snoad B. (1958). "Looking at Chromosomes," St Martin's, Macmillon.

Meier I., Phelan T., Gruissem W., Spiker S., and Schneider D. (1996). *MFP1*, a novel plant filament-like protein with affinity for matrix attachment region DNA. *Plant Cell* **8:** 2105-15.

Mello-Sampayo T. (1971). Genetic regulation of meiotic chromosome pairing by chromosome 3D of *Triticum aestivum*. *Nature New Biology* **230:** 23-24.

Mello-Sampayo T. (1972). Compensated monosomic 5B-trisomic 5A plants in tetraploid wheat. *Canadian Journal of Genetics & Cytology* **14:** 463-475.

Mello-Sampayo T., and Canas P. (1973). Suppressors of meiotic chromosome pairing in common wheat. *In* "Proceedings of the Fourth International Wheat Genetics Symposium", pp. 709-713, Columbia, Missouri.

Mello-Sampayo T., and Lorente R. (1968). The role of chromosome 3D in the regulation of meiotic pairing in hexaploid wheat. *EWAC Newsletter* **2:** 16-24.

Mengiste T., Revenkova E., Bechtold N., and Paszkowski J. (1999). An SMC-like protein is required for efficient homologous recombination in *Arabidopsis*. *EMBO Journal* **18:** 4505-12.

Meuwissen R. L., Offenberg H. H., Dietrich A. J., Riesewijk A., van Iersel M., and Heyting C. (1992). A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. *EMBO Journal* **11:** 5091-100.

Michaelis C., Ciosk R., and Nasmyth K. (1997). Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91:** 35-45.

Mikhailova E. I., Naranjo T., Shepherd K., Eden J. W., Heyting C., and de Jong J. H. (1998). The effect of the wheat *Ph1* locus on chromatin organisation and meiotic chromosome pairing analysed by genome painting. *Chromosoma* **107:** 339-350.

Milanesi L., D'Angelo D., and Rogozin I. B. (1999). GeneBuilder: interactive *in silico* prediction of gene structure. *Bioinformatics* **15:** 612-21.

Milanesi L., Muselli M., and Arrigo P. (1996). Hamming-Clustering method for signals p rediction i n 5 ' a nd 3 ' regions of eukaryotic genes. *Computational Applied Bioscience* **12:** 399-404.

Mirkovitch J., Mirault M. E., and Laemmli U. K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* **39:** 223-32.

Mittelsten Scheid O., Jakovleva L., Afsar K., Maluszynska J., and Paszkowski J. (1996). A change of ploidy can modify epigenetic silencing. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 7114-9.

Mlynarova L., Jansen R. C., Conner A. J., Stiekema W. J., and Nap J. P. (1995). The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants. *Plant Cell* **7:** 599-609.

Moens P. B. (1994). Molecular perspectives of chromosome pairing at meiosis. *BioEssays* **16:** 101-106.

Moore G. (1998). To pair or not to pair - Chromosome pairing and evolution. *Current Opinion in Plant Biology* **1:** 116-122.

Moore G., Devos K. M., Wang Z., and Gale M. D. (1995). Cereal genome evolution - Grasses, line up and form a circle. *Current Biology* **5:** 737-739.

Morgante M. (2000). **(personal communication)**.

Moullet O., Zhang H. B., and Lagudah E. S. (1999). Construction and characterisation of a large DNA insert library from the D genome of wheat. *Theoretical & Applied Genetics* **99:** 305-313.

Murashige T., and Skoog F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* **15:** 473-497.

Nag D. K., Scherthan H., Rockmill B., Bhargava J., and Roeder G. S. (1995). Heteroduplex DNA formation and homolog pairing in yeast meiotic mutants. *Genetics* **141:** 75-86.

Nakai K., and Kanehisa M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14:** 897-911.

Nehra N. S., Chibbar R. N., Leung N., Caswell K., Mallard C., Steinhauer L., Baga M., and Kartha K. K. (1994). Self-fertile transgenic wheat plants

regenerated from isolated scutellar tissues following microprojectile bombardment with two distinct gene constructs. *Plant Journal* **5:** 285-297.

Nelson J. C., Sorrells M. E., Vandeynze A. E., Lu Y. H., Atkinson M., Bernard M., Leroy P., Faris J. D., and Anderson J. A. (1995a). Molecular mapping of wheat - Major genes and rearrangements in homoeologous groups 4, 5, and 7. *Genetics* **141:** 721-731.

Nelson J. C., Vandeynze A. E., Autrique E., Sorrells M. E., Lu Y. H., Merlino M., Atkinson M., and Leroy P. (1995b). Molecular mapping of wheat - Homoeologous group 2. *Genome* **38:** 516-524.

Nelson J. C., Vandeynze A. E., Autrique E., Sorrells M. E., Lu Y. H., Negre S., Bernard M., and Leroy P. (1995c). Molecular mapping of wheat - Homoeologous group 3. *Genome* **38:** 525-533.

Newman S. M., Eannetta N. T., Yu H., Prince J. P., de Vicente M. C., Tanksley S. D., and Steffens J. C. (1993). Organisation of the tomato polyphenol oxidase gene family. *Plant Molecular Biology* **21:** 1035-51.

Nichols A. F., Itoh T., Graham J. A., Liu W., Yamaizumi M., and Linn S. (2000). Human damage-specific DNA binding protein *p48*. *The Journal of Biological Chemistry* **275:** 21422-21428.

Niedz R. P., Sussman M. R., and Satterlee J. S. (1995). Green fluorescent protein - an *in vivo* reporter of plant gene expression. *Plant Cell Reports* **14:** 403-406.

Nielsen H., Engelbrecht J., Brunak S., and von Heijne G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10:** 1-6.

Ohta K., Shibata T., and Nicolas A. (1994). Changes in chromatin structure at recombination initiation sites during yeast meiosis. *EMBO Journal* **13:** 5754-63.

Okamoto M. (1957). Asynaptic effect of chromosome V. *In* "Wheat Information Server" (R. J. Callanan, and K. Scogna, Eds.), pp. 6, Mosby-Year Book, Inc, St. Louis, Missouri.

Oparka K. J., Roberts A. G., Prior D. A. M., Chapman S., Baulcombe D., and Santacruz S. (1995). Imaging the green fluorescent protein in plants - Viruses carry the torch. *Protoplasma* **189:** 133-141.

Padmore R., Cao L., and Kleckner N. (1991). Temporal comparison of recombination and synaptonemal complex formation during meiosis in *S. cerevisiae*. *Cell* **66:** 1239-1256.

Pang Q., Hays J. B., and Rajagopal I. (1992). A plant cDNA that partially complements *Escherichia coli recA* mutations predicts a polypeptide not strongly homologous to *RecA* proteins. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 8073-7.

Pang S. Z., Deboer D. L., Wan Y., Ye G. B., Layton J. G., Neher M. K., Armstrong C. L., Fry J. E., Hinchee M. A. W., and Fromm M. E. (1996). An improved green fluorescent protein gene as a vital marker in plants. *Plant Physiology* **112:** 893-900.

Parniske M., Hammondkosack K. E., Golstein C., Thomas C. M., Jones D. A., Harrison K., Wulff B. B. H., and Jones J. D. G. (1997). Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* **91:** 821-832.

Parniske M ., a nd J ones J . D . G . (1999). R ecombination between diverged clusters of the tomato *Cf-9* plant disease resistance gene family. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 5850-5855.

Parniske M., Wulff B. B. H., Bonnema G., Thomas C. M., Jones D. A., and Jones J. D. G. (1999). Homologues of the *Cf-9* disease resistance gene

(*Hcr9s*) are present at multiple loci on the short arm of tomato chromosome 1. *Molecular Plant-Microbe Interactions* **12:** 93-102.

Payne A., and Chu G. (1994). *Xeroderma pigmentosum* group E binding factor recognizes a broad spectrum of DNA damage. *Mutation Research* **310:** 89-102.

Peterson D. G., Price H. J., Johnston J. S., and Stack S. M. (1996). DNA content of heterochromatin and euchromatin in tomato (*Lycopersicum esculentum*) pachytene chromosomes. *Genome* **39:** 77-82.

Piedras P., Rivas S., Droge S., Hillmer S., and Jones J. D. G. (2000). Functional, C-Myc-tagged *Cf-9* resistance gene products are plasma-membrane localized and glycosylated. *Plant Journal* **21:** 529-536.

Prestridge D. S. (1991). SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Computational Applied Bioscience* **7:** 203-6.

Protic M., Hirschfeld S., Tsang A. P., Wagner M., Dixon K., and Levine A. S. (1989). Induction of a novel damage-specific DNA binding protein correlates with enhanced DNA repair in primate cells. *Molecular Toxicology* **2:** 255-70.

Pryer N. K., Salama N. R., Schekman R., and Kaiser C. A. (1993). Cytosolic *Sec13p* complex is required for vesicle formation from the endoplasmic reticulum *in vitro*. *Journal of Cell Biology* **120:** 865-75.

Qiagen (1998). The Qiagen guide to template purification and DNA sequencing, Qiagen Pty Ltd.

Raman R., and Nanda I. (1986). Mammalian sex chromosomes. I. Cytological changes in the chiasmatic sex chromosomes of the male musk shrew, *Suncus murinus*. *Chromosoma* **93:** 367-74.

Raven P. H., and Johnson G. B. (1992). "Biology," Mosby-Year Book, Inc, St. Louis, Missouri.

Rayssiguier C., Thaler D. S., and Radman M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342:** 396-401.

Rechsteiner M., and Rogers S. W. (1996). PEST sequences and regulation by proteolysis. *Trends in Biochemical Sciences* **21:** 267-271.

Reinke R., Krantz D. E., Yen D., and Zipursky S. L. (1988). Chaoptin, a cell surface glycoprotein required for *Drosophila* photoreceptor cell morphogenesis, contains a repeat motif found in yeast and human. *Cell* **52:** 291-301.

Riley R. (1960). The diploidisation of polyploid wheat. *Heredity* **15:** 407-429.

Riley R., and Chapman V. (1958). Genetic control of the cytologically diploid behavior of hexaploid wheat. *Nature* **182:** 713-715.

Riley R., Chapman V., Young R. M., and Belfield A. M. (1966). Control of meiotic chromosome pairing by the chromosomes of homoeologous group 5 of *Triticum aestivum*. *Nature* **212:** 1475-1477.

Roberts M. A., Reader S. M., Dalgliesh C., Miller T. E., Foote T. N., Fish L. J., Snape J. W., and Moore G. (1999). Induction and characterization of *Ph1* wheat mutants. *Genetics* **153:** 1909-1918.

Rogers S. W., Wells R., and Rechsteiner M. (1986). Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. *Science* **234:** 364-368.

Ross K. J., Fransz P., Armstrong S. J., Vizir I., Mulligan B., Franklin F. C. H., and Jones G. H. (1997). Cytological characterization of four meiotic mutants of *Arabidopsis* isolated from T-DNA-transformed lines. *Chromosome Research* **5:** 551-559.

Ryals J. A., Neuenschwander U. H., Willits M. G., Molina A., Steiner H.-Y., and Hunt M. D. (1996). Systemic aquired resistance. *Plant Cell* **8:** 1809-1819.

Saitoh N., Goldberg I., and Earnshaw W. C. (1995). The SMC proteins and the coming of age of the chromosome scaffold hypothesis. *BioEssays* **17:** 759-766.

Saitoh N., Goldberg I. G., Wood E. R., and Earnshaw W. C. (1994). ScII - an abundant chromosome scaffold protein is a member of a family of putative ATPases with an unusual predicted tertiary structure. *Journal of Cell Biology* **127:** 303-318.

Saji S., Umehara Y., Antonio B., Yamane H., Tanoue H., Baba T., Aoki H., Ishige N., Wu J. Z., Koike K., Matsumoto T., and Sasaki T. (2001). A physical map with yeast artificial chromosome (YAC) clones covering 63% of the 12 rice chromosomes. *Genome* **44:** 32-37.

Saka Y., Sutani T., Yamashita Y., Saitoh S., Takeuchi M., Nakaseko Y., and Yanagida M. (1994). Fission yeast *Cut3* and *Cut14*, members of a ubiquitous protein family, are required for chromosome condensation and segregation in mitosis. *EMBO Journal* **13:** 4938-4952.

Sakata K., Nagasaki H., Idonuma A., Waki K., Kise M., and Sasaki T. (1999). A computer program for prediction of gene domain on rice genome sequence. *The 2nd Georgia Technical International Conference on Bioinformatics* **Abstracts:** 78.

Sambrook F., Fritsch E. F., and Maniatis F. (1989). "Molecular cloning: A laboratory manual," Cold Spring Harbour Laboratory, New York.

Scherthan H., Bahler J., and Kohli J. (1994). Dynamics of chromosome organisation and pairing during meiotic prophase in fission yeast. *Journal of Cell Biology* **127:** 273-85.

Schmekel K., Wahrman J., Skoglund U., and Daneholt B. (1993). The central region of the synaptonemal complex in *Blaps cribrosa* studied by electron microscope tomography. *Chromosoma* **102:** 669-681.

Schwacha A., and Kleckner N. (1994). Identification of joint molecules that form frequently between homologs but rarely between sister chromatids during yeast meiosis. *Cell* **76:** 51-63.

Schwarzacher T. (1997). Three stages of meiotic homologous chromosome pairing in wheat - Cognition, alignment and synapsis. *Sexual Plant Reproduction* **10:** 324-331.

Sears E. R. (1954). The aneuploids of common wheat. *Missouri Agricultural Experimental Station Research Bulletin* **572:** 58.

Sears E . R . (1972). T he n ature o f mutation in hexaploid wheat. *Symposia Biologica Hungarica* **12:** 73-82.

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 & Cytology* **19:** 585-593.

Sears E. R. (1982). A wheat mutant conditioning an intermediate level of homoeologous chromosome pairing. *Canadian Journal of Genetics & Cytology* **24:** 715-719.

Sears E. R., and Okamoto M. (1958). Intergenomic chromosome relationships in hexaploid wheat. *In* "Proceedings of the Tenth International Congress of Genetics", pp. 258-259.

Selva E. M., Maderazo A. B., and Lahue R. S. (1997). Differential effects of the mismatch repair genes *MSH2* and *MSH3* on homeologous recombination in *Saccharomyces cerevisiae*. *Molecular & General Genetics* **257:** 71-82.

Selva E. M., New L., Crouse G. F., and Lahue R. S. (1995). Mismatch correction acts as a barrier to homeologous recombination in *Saccharomyces cerevisiae*. *Genetics* **139:** 1175-88.

Shaw P., and Moore G. (1998). Meiosis: *vive la* difference! *Current Opinion in Plant Biology* **1:** 458-62.

Sheen J., Hwang S. B., Niwa Y., Kobayashi H., and Galbraith D. W. (1995). Green-fluorescent protein as a new vital marker in plant cells. *Plant Journal* **8:** 777-784.

Sherman J. D., Stack S. M., and Anderson L. K. (1989). Two-dimensional spreads of synaptonemal complexes from solanaceous plants IV. Synaptic irregularities. *Genome* **32:** 743-753.

Smilde W. D., Haluskova J., Sasaki T., and Graner A. (2001). New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from r ice expressed sequence tags. *Genome* **44:** 361-7.

Smith T. F., Gaitatzes C., Saxena K., and Neer E. J. (1999). The WD repeat: a common architecture for diverse functions. *Trends in Biochemical Sciences* **24:** 181-5.

Song W. Y., Wang G. L., Chen L. L., Kim H. S., Pi L. Y., Holsten T., Gardner J., Wang B., Zhai W. X., Zhu L. H., Fauquet C., and Ronald P. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* **270:** 1804-1806.

Southern E. M. (1975). Detection of s pecific D NA s equences a mong D NA fragments seperated by gel electrophoresis. *Journal of Molecular Biology* **98:** 503-517.

Spiker S., and William F. (1996). Nuclear matrix attachment regions and transgene expression in plants. *Plant Physiology* **110:** 15-21.

Stein N., Feuillet C., Wicker T., Schlagenhauf E., and Keller B. (2000). Subgenome chromosome walking in wheat: A 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proceedings of the National Academy of Sciences of the United States of America* **97:** 13436-41.

Strunnikov A. V., Hogan E., and Koshland D. (1995). *SMC2*, a *Saccharomyces cerevisiae* gene essential for chromosome segregation and condensation, defines a subgroup within the SMC family. *Genes & Development* **9:** 587-599.

Strunnikov A. V., Larionov V. L., and Koshland D. (1993). *SMC1*: an essential yeast gene encoding a putative head-rod-tail protein is required for nuclear division a nd d efines a new ubiquitous protein family. *Journal of Biological Chemistry* **123:** 1635-48.

Sugawara N ., P aques F ., Colaiacovo M., and Haber J. E. (1997). Role of *Saccharomyces cerevisiae MSH2* and *MSH3* repair proteins in double-strand break-induced recombination. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 9214-9.

Sun H., Treco D., Schultes N. P., and Szostak J. W. (1989). Double strand breaks at an initiation site for meiotic gene conversion. *Nature* **338:** 87-90.

Sutani T., and Yanagida M. (1997). DNA renaturation activity of the SMC complex implicated in chromosome condensation. *Nature* **388:** 798-801.

Suzuki N., Choe H.-R., Nishida Y., Yamawaki-Kataoka Y., Ohnishi S., Tamaoki T., and Kataoka T. (1990). Leucine-rich repeats and carboxyl terminus are required for interaction of yeast adenylate cyclase with RAS proteins. *Proceedings of the National Academy of Sciences of the United States of America* **87:** 8711-8715.

Sym M., Engebrecht J. A., and Roeder G. S. (1993). *Zip1* is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell* **72:** 365-78.

Thomas C. M., Dixon M. S., Parniske M., Golstein C., and Jones J. D. (1998). Genetic and molecular analysis of tomato *Cf* genes for resistance to *Cladosporium fulvum*. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **353:** 1413-24.

Thomas C. M., Jones D. A., Parniske M., Harrison K., Balint-Kurti P., Hatzixanthis K., and Jones J. D. G. (1997). Characterisation of the *Cf-4* gene for r esistance t o *C ladosporium f ulvum* i dentifies s equences t hat determine recognitional specificity in *Cf-4* and *Cf-9*. *The Plant Cell* **9:** 2209-2224.

Thomas H. M. (1990). Analysis of synaptonemal complexes in the amphidiploid of *Lolium multiflorum* x *Festuca drymeja*. *Genome* **33:** 903-907.

Thomas S. W. (1997). Molecular studies of homologous chromosome pairing in *Triticum aestivum*. *In* "Department of Plant Science", pp. 173, Adelaide University, Adelaide.

Tokunaga Y., Keon J. P. R., and Hargreaves J. A. (1999). Isolation of a barley homolog of *Cf2/Cf5* disease resistance genes by differential display of benzothiodiazole-treated plants, GenBank Accession Number AF166121(Cell Biology, IACR-Long Ashton, Long Ashton, Bristol BS41 9AF, UK).

Townsley F. M., and Ruderman J. V. (1998). Proteolytic ratchets that control progression through mitosis. *Trends in Cell Biology* **8:** 238-244.

Twyman R. M., Ed. (1998). "Advanced molecular biology: A concise reference," BIOS Scientific Publishers Ltd, Oxford.

Uphadya M. D., and Swaminathan M. S. (1967). Mechanisms regulating chromosome pairing in *Triticum*. *Biologisches Zentralblatt* **87s:** 239-255.

Vaisman N., Tsouladze A., Robzyk K., Ben-Yehuda S., Kupiec M., and Kassir Y. (1995). The role of *Saccharomyces cerevisiae Cdc40p* in DNA replication and mitotic spindle formation and/or maintenance. *Molecular & General Genetics* **247:** 123-36.

Vandeynze A. E., Nelson J. C., Odonoughue L. S., Ahn S. N., Siripoonwiwat W., Harrington S. E., Yglesias E. S., Braga D. P., McCouch S. R., and Sorrells M. E. (1995). Comparative mapping in grasses - Oat relationships. *Molecular & General Genetics* **249:** 349-356.

Vasil V., Castillo A. M., Fromm M. E., and Vasil I. K. (1992). Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Biotechnology* **10:** 667-674.

Vega J. M., and Feldman M. (1998a). Effect of the pairing gene *Ph1* and premeiotic colchicine treatment on intra- and interchromosome pairing of isochromosomes in common wheat. *Genetics* **150:** 1199-1208.

Vega J. M., and Feldman M. (1998b). Effect of the pairing gene *Ph1* on centromere misdivision in common wheat. *Genetics* **148:** 1285-1294.

Vinatzer B. A., Patocchi A., Gianfranceschi L., Tartarini S., Zhang H. B., Gessler C., and Sansavini S. (2001). Apple contains receptor-like genes homologous to the *Cladosporium fulvum* resistance gene family of tomato with a cluster of genes cosegregating with *Vf* apple scab resistance. *Molecular Plant-Microbe Interactions* **14:** 508-15.

Vincent J. E., and Jones G. H. (1993). Meiosis in autopolyploid *Crepis capillaris*. I. Triploids and trisomics; implications for models of chromosome pairing. *Chromosoma* **102:** 195-206.

von Heijne G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research* **14:** 4683-90.

von Wettstein D., Rasmussen S. W., and Holm P. B. (1984). The synaptonemal complex in genetic segregation. *Annual Review of Genetics* **18:** 331-413.

Vulic M., Dionisio F., Taddei F., and Radman M. (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in

enterobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 9763-7.

Walker J. E., Saraste M., Runswick M. J., and Gay N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO Journal* **1:** 945-51.

Wall A. M., Riley R., and Chapman V. (1971). Wheat mutants permitting homoeologous meiotic chromosomes pairing. *Genetical Research* **18:** 311-328.

Walters M. S. (1970). Evidence of the time of chromosome pairing from the preleptotene spiral stage in *Lolium longiflorum* 'Croft". *Chromosoma* **29:** 375-418.

Wang B. C., Dickinson L. A., Koivunen E., Ruoslahti E., and Kohwishigematsu T. (1995). A novel matrix attachment region DNA binding motif identified using a random phage peptide library. *Journal of Biological Chemistry* **270:** 23239-23242.

Weiner B. M., and Kleckner N. (1994). Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast. *Cell* **77:** 977-991.

Wicker T. (2001). **(personal communication)**.

Wicker T., Stein N., Albar L., Feuillet C., Schlagenhauf E., and Keller B. (2001). Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. *Plant Journal* **26:** 307-16.

Williams F. E., Varanasi U., and Trumbly R. J. (1991). The *CYC8* and *TUP1* proteins involved in glucose repression in *Saccharomyces cerevisiae* are associated in a protein complex. *Molecular & Cellular Biology* **11:** 3307-16.

Wolters P. (2000). **(personal communication).**

Wright T. R. F. (1996). Phenotypic analysis of the *DOPA* decarboxylase gene cluster mutants in *Drosophila melanogaster*. *Journal of Heredity* **87:** 175-190.

Wu T. C., and Lichten M. (1994). Meiosis-induced double-strand break sites determined by yeast chromatin structure. *Science* **263:** 515-8.

Yacobi Y. Z., Mello-Sampayo T., and Feldman M. (1982). Genetic induction of bivalent interlocking in common wheat. *Chromosoma* **87:** 165-175.

Yanagisawa S. (1997). *Dof* DNA-binding domains of plant transcription factors contribute to multiple protein-protein interactions. *European Journal of Biochemistry* **250:** 403-10.

Yanagisawa S. (1998). *Dof* proteins: involvement of transcription factors with a novel DNA- binding domain in tissue-specific and signal-responsive gene expression [Japanese]. *Seikagaku* **70:** 280-5.

Yanagisawa S. (2000). *Dof1* and *Dof2* transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. *Plant Journal* **21:** 281-8.

Yanagisawa S., and Schmidt R. J. (1999). Diversity and similarity among recognition sequences of *Dof* transcription factors. *Plant Journal* **17:** 209-14.

Yanagisawa S., and Sheen J. (1998). Involvement of maize *Dof* zinc finger proteins in tissue-specific and light-regulated gene expression. *Plant Cell* **10:** 75-89.

Yang D., Parco A., Nandi S., Subudhi P., Zhu Y., Wang G., and Huang N. (1997). Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. *Theoretical & Applied Genetics* **95:** 1147-1154.

Zickler D., and Kleckner N. (1998). The leptotene-zygotene transition of meiosis. *Annual Review of Genetics* **32:** 619-+.

Zickler D., and Kleckner N. (1999). Meiotic chromosomes: Integrating structure and function. *Annual Review of Genetics* **33:** 603-754.

Appendix

Table 19: Transgenic plants produced from individual microprojectile bombardment experiments. Each experiment is determined by date of microprojectile bombardment and the cultivar plus number of isolated scutella. The co-bombardment class, number of plantlet regenerants into glasshouse and the identified number of transformation events for each experiment are given.

Date	Cultivar	Number of Scutella bombarded	Co-bombardment class (refer to Table 11)	Plantlet regenerants	Transformation Events
1/3	Veery	90 (C)	X^2	0	0
		20	Control		
5/3	Veery	45 (A + C)	X	1	0
		23 (B)	Y^1	0	0
		24 (D)	Z^3	0	0
		10	Control		
	Pavon	64 (A + C)	X	2	0
		32 (B)	Y^1	0	0
		28 (D)	Z^3	0	0
		10	Control		
	Combi	28 (A)	X^1	3	0
		27 (B)	Y^1	0	0
		10	Control		
8/3	Veery	20 (C)	X^2	0	0
		60 (B)	Y^1	4	0
		60 (D')	Z^1	1	0
		10	Control		
13/3	Veery	54 (C')	X^2	0	0
		54 (B')	Y^1	0	0
		54 (D')	Z^1	1	0
		10	Control		
	Pavon	36 (C')	X^2	0	0
		54 (B')	Y^1	1	0
		36 (D')	Z^1	2	1
		10	Control		
	Combi	15 (D')	Z^1	2	0
18/3	Veery	72 (A)	X^1	0	0
		48 (B)	Y^1	2	1
		48 (E)	Z^3	1	0
		16	Control		
	Pavon	44 (A)	X^1	0	0
		44 (B)	Y^1	0	0
		44 (E)	Z^3	1	1
	Combi	18 (A)	X^1	3	1
		17 (B)	Y^1	0	0
22/3	Veery	140	X	1	0
		96	Y	0	0
		112	Z	0	0
	Pavon	66	X	0	0
		44	Y	2	1
		66	Z	3	1
26/3	Veery	72	X	4	0
		78	Y	10	0
		78	Z	16	1
		14	Control		
	Pavon	24	X	1	0
		24	Y	1	0

		24	Z	0	0
		10	Control		
31/3	Veery	114	X	12	0
		86	Y	24	0
		114	Z	8	0
		12	Control		
	Pavon	24	X	2	0
		24	Y	2	0
		24	Z	0	0
4/4	Veery	82	X	26	1
		56	Y	0	0
		56	Z	0	0
		12	Control		
5/4	Veery	84	X	28	0
		56	Y	47	3
		84	Z	19	0
		20	Control		
9/4	Veery	72	X	28	0
		48	Y	7	1
		72	Z	1	0
		5	Control		
13/4	Veery	54	X	0	0
		54	Y	0	0
		27	Z	1	0
		6	Control		
	Combi	24	X	14	
		24	Y	12	1
		24	Z	7	0
		12	Control		
	Florida	26	X	1	0
		26	Z	1	0
		6	Control		
16/4	Veery	60	X	28	1
		60	Y	25	2
		60	Z	18	2
	Combi	25	X	7	0
	Florida	60	X	1	0
		60	Y	4	0
		90	Z	4	0
		10	Control		
19/4	Veery	24	X	1	0
		24	Y	0	0
		24	Z	0	0
	Florida	60	X	2	0
		30	Y	0	0
		30	Z	0	0
		6	Control		
22/4	Veery	30	X	3	0
		30	Y	0	0
		30	Z	1	0
		12	Control		
	Combi	26	X	11	2
		26	Y	17	1
		26	Z	23	0
		8	Control		
	Florida	62	X	0	0
		62	Y	1	0

		62 8	Z Control	3	0
26/4	Veery	30 30 30 12	X Y Z Control	15 11 8	1 1 0
	Florida	60 60 60 12	X Y Z Control	28 4 0	1 0 0
29/4	Veery	140 120 140 12	X Y Z Control	3 0 25	0 0 1
3/5	Veery	90 90 90 10	X Y Z Control	6 4 4	0 0 1
	Combi	21 21 21	X Y Z	7 1 0	0 0 0
6/5	Veery	120 120 120 10	X Y Z Control	8 30 19	0 0 0
14/5	Veery	120 120 12	X Z Control	9 1	0 0
	Combi	25 25	X Z	0 3	0 0
Total		5832		637	26

[1]except *WM5* GUS and *WM5* GFP reporter gene constructs
[2]except Antisense *WM5* and *WM5* GFP constructs
[3]except *WM5* GUS reporter gene construct
[4]except *WM5* GFP reporter gene construct