



UNIVERSITY OF
LEICESTER

Using *Brachypodium* as a Model to Study Key
Genes Regulating Reproductive Development in
Temperate Grasses

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

By
Safia Saada

Department of Genetics and Genome Biology
University of Leicester

February 2020

Abstract

MADS-box genes encode a family of transcription factors (TFs) that are central to reproductive development in plants. Within this family, C-class and D-class genes control the development of the third and the fourth whorls in flowers that contain the reproductive organs, stamens and carpel, respectively. Investigating the role of these genes in essential crops such as wheat and barley will provide new insights into the genetic and molecular control of the formation of floral organs and in grain development. This study aimed to investigate the roles of C and D class genes in the monocot grass model system, *Brachypodium* in comparison to wheat and barley. The first objective was to identify all *MADS-box MIKC-type* genes in four species (*Brachypodium distachyon*, *Oryza sativa*, *Triticum aestivum*, *Hordeum vulgare* and *Arabidopsis thaliana*) and investigate their spatio-temporal expression patterns across developmental stages to suggest potential functions. The second objective was to investigate the effect of the loss-of-function of *BdMADS3* from C-lineage, and *BdMADS13* from the D-lineage on *Brachypodium distachyon* and try to select their potential targets.

This study established that there was an expansion in some *MADS-box* genes families in wheat and barley. Also, these genes were expressed during grain development which may suggest roles here. The comparative quantitative expression of C and D-class genes showed a conserved pattern across the three species (*Brachypodium*, wheat and barley) in addition to their high expression during grain filling. This expression was localised in *Brachypodium* grain sections in the degenerating mesocarp prior and after anthesis, then in the cellularising syncytial endosperm at full length stage, and finally in the aleurone layer at maturity. Furthermore, our functional analysis of *BdMADS3* and *BdMADS13* in *Brachypodium* revealed they are essential for plant fertility and proper spikelet formation. Finally, differential expression showed that *BdMADS13* might regulate the reactive oxygen species level at pre-anthesis stage.

Acknowledgements

First and foremost, I would like to thank Almighty God for showering me with mercy and blessings all my life and giving me the strength and guidance to reach this milestone.

I would like to express my deepest gratitude to my supervisor Dr Sinead Drea for her precious advice, direction and supervision throughout my PhD, I really could not have done it without the support of such a great supervisor.

I would like also to thank Dr Richard J Gornall and Dr Julie Morrissey for the guidance and the annual review to improve my work. I want to express my deep appreciation and thanks also to Dr Eamonn Mallon for his suggestions and support. Equally, to Professor David Twell and Dr James Higgins for unlimited help and valuable advices.

I am grateful to each member of the Department Genetics & Genome Biology and I take the chance to thank many people who have made this thesis possible:

Special thanks to Charles Solomon who was an excellent college, thank you for all your help with RNA-seq analyses and your insight and contributions on the PCD paper. Additionally, I would like to thank Dr Philippa Borrill and Dr Abdulkader Alabdullah from John Innes Centre, who helped me with GENIE3 analysis. I would also like to thank Dr Inna Guterman from Dr James Higgins lab, for helping with SDS-PAGE. I would also like to thank Dr Dieter Hackenberg and Dr Lewis Collins Dr Jaypal Darbar from Professor David Twell, for helping me with the transient expression using dual luciferase assay. Special thanks to Neelam Dave support, fast response and excellent problem-solving. I am heartily grateful to Dr Amal Alyamani, for being extremely generous with her knowledge and time. Likewise, I would like to thank Dr Nicolas Sylvius and Caroline Cayrou for helping me with qRT-PCR calculation. I am truly blessed for being surrounded by such amazing and wonderful people.

Safia

Dedication

I dedicate this thesis to my mum Fatima Ait Taleb and my dad Mohammad, who have provided me with their encouragement, love and understanding.

Also, my brothers for their unconditional love, and my sister, who inspired and encouraged me throughout my study.

I would not be where I am today without the support of my husband, who funded my school fees and shifted his working hours to make it possible for me. I really appreciate all the help, support and kindness that you have given me.

I dedicate this work as well to my daughters, RAZAN and MARYAM hoping for them a very bright future Insha Allah.

Contents

Abstract.....	i
Acknowledgements	ii
Dedication	iii
Table of contents.....	iv
List of figures.....	viii
List of tables.....	x
Abbreviation.....	xi
Chapter 1. General introduction and literature review	1
1.1 The grass family (Poaceae).....	1
1.1.1 The <i>Pooideae</i> Subfamily	2
1.1.2 The <i>Triticeae</i> tribe	3
1.1.3 Wheat production, economical value, and its contribution to human diet and health.....	6
1.1.4 Wheat wild ancestors	6
1.1.5 <i>Brachypodium Distachyon</i>	8
1.2 Reproductive Development in Grasses	9
1.2.1 Meristematic determinacy, flower and fruit development in grasses (<i>Poaceae</i>).....	9
1.2.2 Grain Development	11
1.2.3 ABCDE model floral organogenesis	13
1.2.4 The role of transcription factors (TFs) in plants	15
1.2.5 <i>MADS-box</i> transcription factors	15
1.2.6 The diversity of <i>MADS</i> domain transcription factor.....	16
1.2.7 The function of C and D class genes	17
Chapter 2. Phylogeny and expression of <i>MADS-box MIKC-type</i> genes in wild and domesticated grasses.....	21
2.1 Abstract	21
2.2 Introduction	21
2.3 Materials and Methods	23
2.3.1 Plant materials growth conditions	23
2.3.2 Plants sampling	24
2.3.3 Molecular analysis.....	24
2.3.4 Sequence and phylogenetic analysis.....	25
2.3.5 Transcriptome Analysis	26
2.3.6 Genomic distribution of <i>MADS-box</i> genes	27
2.4 Results	27
2.4.1 C and D-class <i>MADS-box</i> Genes ortholog Identification	27
2.4.2 Gene Identification and analyses	28
2.4.3 <i>MADS-box</i> transcription factors protein structure and conserved domain	30
2.4.4 Phylogenetic, Identification and classification analysis of <i>MADS-box</i> genes in some monocots	32
2.4.5 Chromosomal locations of <i>MADS-box</i> gene family members in <i>Triticum aestivum</i> , <i>Oryza sativa</i> and <i>Arabidopsis thaliana</i>	36

2.4.6	Conserved motifs identification	37
2.4.7	C/D-class <i>MADS-box</i> transcription factor expression in wheat.....	43
2.4.8	C/D-class <i>MADS-box</i> transcription factor expression in barley.....	48
2.5	Discussion.....	53
2.6	Conclusion	54
Chapter 3.	Investigating the role of <i>AGAMOUS</i> genes during grain development in <i>Brachypodium distachyon</i>	57
3.1	Abstract	57
3.2	Introduction	58
3.2.1	<i>Brachypodium distachyon</i> as a Reference:.....	59
3.2.2	ABCDE Model of Flower Development	61
3.3	Materials and Methods	65
3.3.1	Plant materials	65
3.3.2	Tissue Collection and storage.....	65
3.3.3	Nucleic acid Analysis	65
3.3.4	mRNA <i>In situ</i> Hybridisation.	70
3.4	Results	74
3.4.1	Expression pattern analyses of C and D lineage genes in <i>Brachypodium</i>	74
3.5	Discussion.....	85
3.6	Conclusion	88
Chapter 4.	Functional analysis of <i>BdMADS3</i> and <i>BdMADS13</i> in <i>Brachypodium</i>	90
4.1	Abstract	90
4.2	Introduction	90
4.3	Material and methods.....	93
4.3.1	Transgenic materials and growth conditions	93
4.3.2	Plant sampling.....	93
4.3.3	Characterisation of the transgenic phenotypes	94
4.3.4	Morphological analysis.....	94
4.3.5	Expression clone and genotyping.....	94
4.3.6	Reverse Transcriptase PCR Analysis	95
4.3.7	Quantitative reverse transcription PCR Analysis.....	95
4.3.8	Scanning electron microscopy (SEM)	96
4.3.9	Light microscopy and histological analysis.....	96
4.4	Results	97
4.4.1	Comparing and contrasting the expression profile between the RNAi lines and the wild-type 98	
4.4.2	Floret and spikelet comparison.....	102
4.4.3	Spikelet and floret histological analysis of the two RNAi lines.....	106
4.4.4	Lodicule morphology.....	112
4.4.5	Anther development and morphology.....	114
4.4.6	Anther morphology and structure analysis using scanning electron microscopy (SEM)	115

4.4.7	Light microscopy analysis.....	116
4.4.8	Pollen grains morphology	119
4.5	Discussion.....	122
4.6	Conclusion	126
Chapter 5.	Potential downstream targets for <i>BdMADS3</i> and <i>BdMADS13</i>.....	128
5.1	Abstract	128
5.2	Introduction	128
5.3	Materials and methods	132
5.3.1	Plant materials	132
5.3.2	Co-expression analysis and potential targets of <i>MADS3</i>	132
5.3.3	cDNA and DNA promoters amplification and manipulation	133
5.3.4	Bacterial transformations.....	136
5.3.5	cDNA and promoters manipulation, cloning, and sequencing.....	138
5.3.6	Dual-luciferase transient expression assay	140
5.3.7	RNA-seq plant material and growth.....	143
5.4	Results	146
5.4.1	Selecting the potential target genes for <i>BdMADS3</i> using co-expression analysis	146
5.4.2	<i>BdMADS3</i> interact with <i>BdMADS7</i> transcription factor to bind and activate metallothioneins (<i>MT</i>) gene promoter	150
5.4.3	Differential gene expression analysis using high-throughput RNA sequencing (RNA-seq).	152
5.5	Discussion.....	167
5.6	Conclusion	169
Chapter 6.	<i>Brachypodium</i> grain development and programmed cell death process.....	171
6.1	Abstract	171
6.2	Introduction	171
6.3	Materials and Methods	176
6.3.1	Plant materials	176
6.3.2	TUNEL assay	177
6.3.3	Phylogenetic analysis of selected protease Family genes.....	178
6.3.4	RNA-seq data processing and expression analysis of protease genes	178
6.3.5	Reverse transcription polymerase chain reaction (RT-PCR).....	179
6.3.6	Quantitative RT-PCR.....	179
6.3.7	mRNA <i>In situ</i> Hybridisation.	180
6.4	Results	181
6.4.1	Phylogeny cysteine protease C1, vacuolar processing enzymes (VPE) C13 and meta-caspase C14 proteases families in wild and domesticated species.....	181
1.1.	Expression profile analysis.....	185
6.4.2	Expression pattern of proteases and RNA-seq validation using RT-qPCR.....	187
6.4.3	Spatial distribution and progress of DNA fragmentation in the developing grain.	188

6.4.4	Expression pattern of <i>BdMADS29</i> in the life cycle of the <i>Brachypodium</i> plant in different tissues samples using RT-PCR	190
6.4.5	The localisation of mRNA of <i>BdMADS29</i> in grain development from Pre-anthesis to mature grain	191
6.5	Discussion.....	194
6.6	Conclusion.....	197
Chapter 7.	Summary and conclusion.....	199
References.....		205
Appendix.....		233

List of Figures

Figure 1-1: Phylogenetic tree of the subfamilies of the Poaceae	2
Figure 1-2: (A) Grain anatomy and (B) Phylogenetic tree	5
Figure 1-3: Determination of terminal and lateral spikelet stages in <i>Brachypodium distachyon</i>	10
Figure 1-4: Structure of the floret in wheat barley and <i>Brachypodium</i> . Wheat, barley and <i>Brachypodium</i> have the same floral organs except that <i>Brachypodium</i> has only two stamens and smaller floret. Ca: Carpel, Le: Lemma, Lo: Lodicule Pa: Palea, MG: Mature grain, St: stamen, Stg: Stigma. Adapted from (Heinze, 2017). .	11
Figure 1-5: Schematic representations of the developmental stages of barley caryopsis.	12
Figure 1-6: ABCDE model summarising the homeotic function and the expression profile	14
Figure 2-1: The major clade of <i>MADS-box</i> genes.	22
Figure 2-2: MAFFT alignment of C-and D-class <i>MADS</i> transcription factors	30
Figure 2-3: The consensus sequences of <i>MADS-box</i> domain.	30
Figure 2-4: Maximum likelihood phylogenetic tree	33
Figure 2-5: Chromosome location and distribution of <i>MADS-box</i> transcription factor protein	37
Figure 2-6: Alignment of the C-terminal regions of the A-class amino acid sequences	38
Figure 2-7: Alignment of the C-terminal regions of the B-class amino acid sequences	39
Figure 2-8: C-terminal region of the protein sequence alignment.....	40
Figure 2-9: Alignment of the C-terminal regions of the D-class amino acid sequences.....	41
Figure 2-10: Alignment of the C-terminal regions of the E-class amino acid sequences	42
Figure 2-11: Physiological growth and development in wheat spike and grain.....	43
Figure 2-12: RT-PCR illustrating the expression profile of <i>TaMADS3</i> , 58, 13 and 21 in eight tissues samples	44
Figure 2-13: Quantification of the expression profile	45
Figure 2-14: Expression Analysis of <i>MADS-box</i> MIKC-type genes transcripts	47
Figure 2-15: <i>Hordeum vulgare</i> developmental stages that samples were collected from.	49
Figure 2-16: Expression pattern of <i>HvMADS3</i> (<i>HvAG1</i>), <i>HvMADS58</i> (<i>HvAG2</i>) and <i>HvMADS13</i> in barley	49
Figure 2-17: Quantification of the expression profile	50
Figure 2-18: Expression Analysis of <i>MADS-box</i> MIKC-type genes in <i>Hordeum vulgare</i>	52
Figure 3-1: Schematic phylogenetic relationship of <i>B. distachyon</i> to other <i>Poaceae</i>	59
Figure 3-2: Schematic representation	61
Figure 3-3: Sampling stages of <i>B. distachyon</i>	74
Figure 3-4: Expression patterns of <i>Brachypodium MADS3</i> , 58, 13, and 21 TFs.....	75
Figure 3-5: Schematic summarising <i>Brachypodium distachyon</i> grain development.....	76
Figure 3-6: The mRNA expression pattern of <i>BdMADS3</i>	77
Figure 3-7: <i>BdMADS58</i> mRNA in different stages of <i>Brachypodium distachyon</i>	78
Figure 3-8: Spatial and temporal expression of <i>BdMADS13</i> mRNA	79
Figure 3-9: The mRNA expression pattern of <i>BdMADS21</i> ,	80
Figure 3-10: General illustration of the special localisation of the mRNA signal of (a) C-class genes and (b) D-class genes. DAG: days after germination, DAA: days after anthesis.	81
Figure 3-11: The relative expression of (A) <i>BdMADS3</i> and (B) <i>BdMADS58</i> mRNA of <i>Brachypodium</i>	81
Figure 3-12: The relative expression of <i>Brachypodium MADS13</i> (A) and <i>MADS21</i> (B) mRNA	83
Figure 3-13: Heat map illustrating all <i>MADS</i> MIKC-type in <i>Brachypodium</i> from eight different tissues samples,	84
Figure 4-1: ABCDE floral model suggested for rice and wheat, the model	91
Figure 4-2: The pIPK007 RNAi expression clone generated	95
Figure 4-3: RT-PCR comparing the different expression pattern between <i>Bd21</i> (WT)	98
Figure 4-4: RT-PCR comparing the different expression pattern between <i>Bd21</i>	99
Figure 4-5: mRNA expression of <i>Bdmads3</i> RNAi and <i>Bd21</i> (WT) using qRT-PCR.	100
Figure 4-6: Differential expression of <i>BdMADS13</i> mRNA in the <i>Bdmads13-3</i> RNAi and <i>Bd21</i> (WT) lines.....	101
Figure 4-7: <i>Bdmads3</i> interference line phenotype in comparison with the wild type <i>Bd21</i>	102
Figure 4-8: <i>Brachypodium</i> spikelet developmental stages and phenotypic alteration of <i>Bdmads13-3</i> RNAi spike	104

Figure 4-9: Phenotypic comparison of plants from both lines at two developmental stages.....	105
Figure 4-10: Flowering time period in <i>BdWT</i> , <i>Bdmads3-2</i> and <i>Bdmads13-3</i>	105
Figure 4-11: <i>Brachypodium</i> spikelet at post-anthesis stage.....	106
Figure 4-12: Illustration of <i>Brachypodium distachyon</i> mature grain.....	106
Figure 4-13: Different carpel and grain successive developmental stages	108
Figure 4-14: <i>Brachypodium</i> pre-anthesis floret transversal semi-thin sections.....	109
Figure 4-15: Transverse sections of <i>Brachypodium</i> palea vascular bundle.....	110
Figure 4-16: Transversal sections of <i>Brachypodium</i> ultrastructure of the rachis.....	110
Figure 4-17: SEM of (a-b) the adaxial surface of the lemma and (d-e) awn of <i>Brachypodium</i> floret,.....	111
Figure 4-18: Lodicule morphology comparison between <i>Bdmads3</i> RNAi line and WT line	113
Figure 4-19: <i>Brachypodium</i> lodicule at anthesis stage.	113
Figure 4-20: The anther developmental stages in <i>Brachypodium</i>	114
Figure 4-21: Scanning electron microscope of <i>Bdmads3</i> RNAi line developing anther,	115
Figure 4-22: <i>Brachypodium distachyon</i> mature anther phenotype.	116
Figure 4-23: Light micrographs of longitudinal semi-thin sections of mature anther.....	118
Figure 4-24: comparison of DNA fragmentation detected in <i>Brachypodium</i> anther at maturity stage.....	119
Figure 4-25: Mature pollen of <i>B. distachyon</i> stained with Alexander staining.	120
Figure 4-26 : Mature pollen grain morphology under phase-contrast microscope.	120
Figure 4-27: Fluorescent micrographs of fresh mature pollen grain stained with DAPI.	121
Figure 5-1: Co-expression analysis for <i>MADS3</i> in four species	147
Figure 5-2: Directed edge weight in top 1 million edges <i>TaMADS3</i> co-expressed genes and their using GENI3. Higher numbers indicate a stronger connection.	149
Figure 5-3: Plasmid and genes used to construct vectors.	150
Figure 5-4: Transactivation assays using the luciferase reporter gene	151
Figure 5-5: A workflow diagram of methods for identifying the normalised count.....	152
Figure 5-6: The read counts of the transcripts and their distribution among the eight samples.	153
Figure 5-7: Venn diagram of differentially expressed genes in very young inflorescence (Vy Inf)	153
Figure 5-8: Heat-map of the 39 identified differentially expressed genes.....	154
Figure 5-9: KEGG pathway analysis of the down-regulated genes list	165
Figure 5-10: Quantitative RT-PCR validation of differentially expressed genes that showed a significant expression differences between the wild-type and <i>Bdmads13</i> RNAi line.....	166
Figure 6-1: Phylogenetic Trees of papain C1 family, C13 Vacuolar processing enzyme (VPE)	183
Figure 6-2: Hierarchically Clustered Expression	185
Figure 6-3: Normalised relative expression analysis of selected peptidase candidates.....	187
Figure 6-4: DNA fragmentation detected in <i>Brachypodium</i> grain section in five developmental stages,	189
Figure 6-5: (a) RT-PCR illustrating the expression profile of <i>BdMADS29</i> in different tissues samples.....	190
Figure 6-6: <i>in situ</i> hybridisation of <i>BdMADS29</i> mRNA in different grain developmental stages	192
Figure 6-7: palea vascular bundle, (a-b) TUNEL assay and its negative control,	193
Figure 6-8: Illustration of the expression pattern of C- and D-lineage genes in <i>Brachypodium distachyon</i> .	201
Figure 7-6-9: General illustration of the spatial localisation of the mRNA expression in grain sections in <i>Brachypodium</i> . DAA: days after anthesis.	202

List of Tables

Table 2-1: Identification of <i>Brachypodium MADS</i> Transcription factors and their orthologs in <i>Oryza sativa</i> .	28
Table 2-2: Selection of wheat (<i>Triticum aestivum</i>) C and D class orthologs in <i>Brachypodium</i> and barley.....	29
Table 2-3: Comparison of gene numbers in each <i>MADS-box</i> family in four species	36
Table 3-1: First-strand cDNA synthesis reaction mix.....	67
Table 4-1: Phenotypic alterations in <i>Bdmads3</i> and <i>Bdmads13</i> RNAi lines.....	124
Table 5-1: Gateway® recombination cloning primers adapters sequences used for each vector.	134
Table 5-2: High-Fidelity DNA polymerase Q5 PCR reaction components and amounts required.....	134
Table 5-3: Antibiotics concentration used for selection in <i>E. coli</i> and <i>A. tumefaciens</i>	137
Table 5-4: The BP recombination reaction components and amounts.	138
Table 5-5: the LR recombination reaction components and amounts.	138
Table 5-6: Two-site LR recombination reaction components and amounts.	139
Table 5-7: Optical density measured at a wavelength of 600 nm required	141
Table 5-8: infiltration mixture containing the effector strains and the reporter strains.	141
Table 5-9: Renilla Luciferase and Firefly luciferase buffers components. *FC: final concentration.....	142
Table 5-10: Library preparation, quality and concentration sent for RNA-seq	143
Table 5-11: samples files and their reads quality.	144
Table 5-12: Gene Ontology (GO) of the <i>BdMADS13</i> putative up-regulated 18 genes represented in.....	155
Table 5-13: GO term annotation of the overlap (Vyinf and PrAs samples)	157
Table 5-14:Overlapping gene list (PrAs+VyIn) that are up-regulated with their fold change	159
Table 5-15: Gene Ontology (GO) of the <i>BdMADS13</i> putative down-regulated 21 genes	160
Table 5-16: Go term annotation of the overlap (Vyinf and PrAs samples).....	161
Table 5-17: Overlap gene list (PrAs+VyIn) that are down-regulated fold change	163
Table 6-1: Gene names and IDs for the <i>HvVPE</i> sub-family identified in previous studies and in our work. .	178

Abbreviations

AG	<i>AGAMOUS</i> gene
AT	<i>Arabidopsis thaliana</i>
Bd	<i>Brachypodium distachyon</i>
Ca	carpel
CDS	coding sequence
DAA	days after anthesis
DPA	days post anthesis
EmSPK	emerging spikelet
FAA	formaldehyde, acetic acid and ethanol
FLG	full-length grain
FLuc	firefly luciferase activity
FLuc/RLuc	normalised dual luciferase activity
Fmol	fentamole
Hv	<i>Hordeum vulgare</i>
IWGSC	International Wheat Genome Sequencing Consortium
Lo	lodicules
MADS	MCM1, <i>AGAMOUS</i> , DEFICIENS, SRF gene
MEF2	myocyte enhancer factor-2
MG	mature grain
MLG	mid-length grain
MT	metallothionein
Os	<i>Oryza sativa</i>
Ov	ovule
PAGE	polyacrylamide-based discontinuous gel
PCD	programmed cell death
PPFD	photosynthetic photon flux density
PrAs	Pre-anthesis
Rluc	Renilla luciferase activity
ROS	reactive oxygen species
SDS	sodium dodecyl sulfate
SPK	spikelet
SRF	Serum response factor
St	stamen
Ta	<i>Triticum aestivum</i>
TF	Transcription factor
TUNEL	Terminal deoxynucleotidyl transferase dUTP nick end labelling
VyInf	very young inflorescence
YG	young grain

Chapter 1

General introduction and literature review

Chapter 1. General introduction and literature review

1.1 The grass family (Poaceae)

The grass family contains about 12,000 species, which makes it the fifth largest flowering plant family. It is distributed world-wide and is arguably our planet's most important family in economic terms. Together the domesticated species can be said to feed almost the entire human population, by providing just over half of all human nutrition and also are poised to become major sources of renewable energy. The family is divided into 13 subfamilies (Fig. 1-1). The Pooideae is the largest grass subfamily with 177 genera and 3850 species (Kellogg, 2015; Vogel et al., 2010), Inhabiting temperate and cold regions, mainly in the northern hemisphere (Barberá et al., 2017). The second largest subfamily in Poaceae is the Panicoideae, with 212 genera and approximately 3316 species (Burke et al., 2016). It is mainly distributed in tropical and subtropical environments of Asia, Australia, Africa, and South America (Craine et al., 2013; Vicentini et al., 2008). The southern hemisphere Panicoideae contains millets, sorghum, maize and sugarcane, and the northern Pooideae contains the cereals wheat, barley, rye, oats and an array of forage grasses; rice occupies a relatively small subfamily, Oryzoideae (Fig. 1-1). The present work focuses on the Pooideae, and especially on the genus *Brachypodium* as a model taxon to help understand more about the biology of fruit production in our cereal crops.

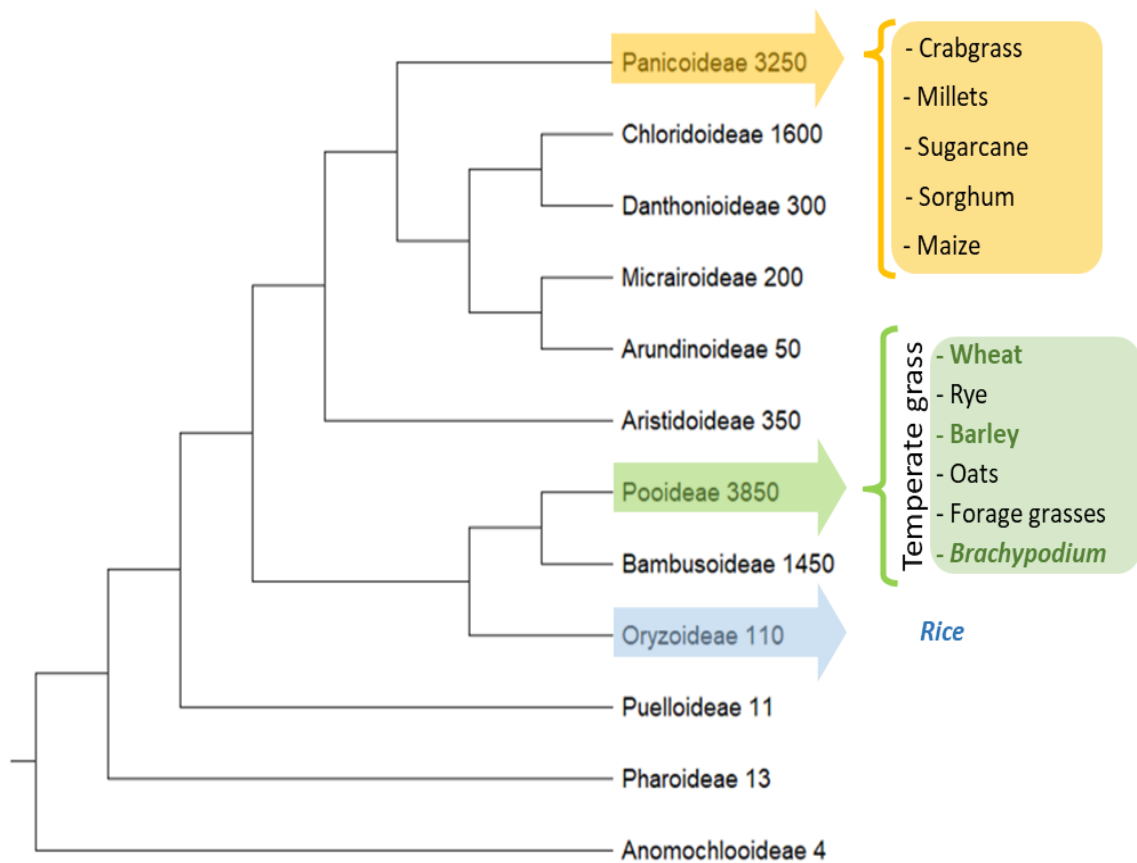


Figure 1-1: Phylogenetic tree of the subfamilies of the Poaceae (after Stevens, P. F. (2001>). Numbers are representing species. Angiosperm Phylogeny Website (mobot.org/MOBOT/research/APweb). Version 14, July 2017.

1.1.1 The *Pooideae* Subfamily

The *Pooideae* subfamily is a large sub-clade that contains almost all the world's major temperate cereals such as wheat, barley, and important forage crop species including *Lolium* and *Phleum* (figure1-1, green highlight).

The various members of the *Pooideae* show significant diversification which was driven by the climate cooling force. It is believed that the *Pooideae* ancestor had adaptations to the temperate climate and cold stress responses and these characteristics are implemented in modern *Pooideae* grasses (Schubert *et al.*, 2019) that shows several variations of physiological, morphological, and genetic diversity. This includes the core *Pooids* that form a prominent and important sub-group of the *Pooideae* containing the *Poeae*, *Triticeae*, and *Bromeae* clades (Catalán *et al.*, 1997).

1.1.2 The *Triticeae* tribe

The cereal grain *Triticeae* tribes contain important crops including wheat, rye and barley. These crop plants are essential in terms of global food security as they provide half of the food eaten by humankind together with rice and maize.

1.1.2.1 Barley (*Hordeum vulgare*. L)

Barley (*Hordeum vulgare*. L), belongs to the genus *Hordeum*, it is an annual, self-pollinating, diploid temperate grass, with $2n = 14$ chromosomes and a genome size of 5.1 gigabases. Its production worldwide constitutes 142.37 million metric tons in the 2017/2018 crop year (statista.com). Approximately 75% of total production is used for animal feed, 20% is malted for a range of beverages, and 5% as breads, stews, and health products (Ullrich, 2010).

Barley is successfully grown at higher latitudes and altitudes with wide adaptation to diverse environmental conditions more than other cereal crops including its close relative wheat (Nevo *et al.*, 2012). Barley pericarp-seed coat and the aleurone layer are particularly high in soluble dietary fiber so it is becoming evident that the consumption of the whole grain has considerable potential to improve human health and to substantially reduces the risk of serious human diet-related diseases such as type II diabetes, colorectal cancers and cardiovascular disease that causes high mortality rate worldwide (Collins *et al.*, 2010).

After anthesis, many important changes occur in the grain structure as maturation proceeds. The crease that is located along the ventral side of the grain is adjusted to an elongated chalazal tract through which nutrients and reserve materials pass from the vascular bundle in the ovary wall to the developing cells of the endosperm (1-2).

At maturity stage, the seed is protected by a relatively hard shield called the pericarp layer and lying immediately after it is a seed coat which contains the integumentary system. These semi-permeable oil-rich cuticular membranes control the penetration of water and solutes into the growing endosperm during grain filling and also into the embryo during germination (Collins, 1918). Adjacent to that, comes a well-defined nucellar epidermis that is characterised with distinctive rectangular cells that form after successive breakdown processes secure nutrients to the developing ovule (Norstog, 1974). In the central region of the kernel, and considered the larger, is the endosperm enclosed by the aleurone layer (Figure 1-2). Barley endosperm cells have a significantly high starch content (up to 64% of

the dry weight) which is considered the main calorie contributor to the human diet, and its cell walls, beta-glucans that account for three-quarter of the total cell wall polysaccharides (Sullivan *et al.*, 2013).

1.1.2.2 Bread wheat (*Triticum aestivum*)

Bread wheat (*Triticum aestivum*), evolved through two polyploidisation events and some progenitors gene flow; first between *Triticum urartu* (2n=14, AA genome) and the goat grass *Aegilops Speltoides* (BB, diploid genome) 500,000 years ago, creating *Triticum diccocooides*. This species was cultivated and called *Emmer* with a tetraploid genome of 4n=28, AABB, which went through a subsequent hybridisation with the wild goat grass (*Aegilops tauschii*) (2n= DD genome) 10 000 years ago, resulting in the modern hexaploid bread wheat (*Triticum aestivum*) (6n=42, AABBDD) (Feldman *et al.*, 1995; Huang *et al.*, 2002; El Baidouri *et al.*, 2017). A recent study in Nature Genetics showed that during the early phase of domestication and cultivation of the wild *Emmer* wheat, at least two distinct lineages of domesticated tetraploids rose; *Triticum dicoccum* and *Triticum durum*, and that the modern *Triticum aestivum* is most likely to be derived from an ancestral hybridization event between the *Triticum durum* (AABB) lineage and wild *Aegilops tauschii* (DD) (Pont *et al.*, 2019).

Evidence shows that the current varieties of the allohexaploid wheat were originated from the Fertile Crescent, a region extending from Palestine, Jordan, Lebanon, Syria to South-East Turkey, to northern Iraq and western Iran. Then went under successive human selection during domestication and cultivation for at least 10,000 years resulting in a wide range of varieties that have an adaptation to several environments around the globe (Pont *et al.*, 2019).

The grasses phenotypic plasticity and adaptability were some of the characteristics that made them colonising species and dominant amongst the major crop species. This flexibility makes them perfect candidates for selection and domestication where they went through dramatic genetic and phenotypic changes. Domestication was a gradual process involving a stage of cultivation of wild plants that came before morphological domestication.

A generally accepted definition for the term domestication is the process whereby wild plants have been evolved into crop plants through early hybridisation event followed by

selective breeding. In this process, the plant accumulated sufficient mutations and adaptations to be unable to compete in its cycle and rely upon human interference for its propagation (Charmet, 2011).

Archaeological data, as well as population genetic simulations, indicate that the first domesticated wheat and barley established a tough rachis as a key trait for domesticated cereals, and that it was potentially evolved around 1500 and 2000 yr. Then came the tenacious glume and non-free thresh-ability in *durum* and common wheat (Dubcovsky & Dvorak, 2007), even with unconscious selection: only seeds falling to the ground contribute to future generations (Hillman & Davies, 1990; Tanno & Willcox, 2006; Fuller, 2007). Then harvesters start selecting seeds from the plant that present beneficial traits such as the size of the grain and spikes to used it for the next generations (Glémin & Bataillon, 2009). Cycles of harvest and sowing of collected seeds maintained a seed retentive phenotype that became dominant, as they contributed most highly to future generations. As domesticated plants were moved with the migration of people into different regions. The cultivation selected not only for desirable traits as a crop but also for adaptations favourable to the local climate and conditions, as only the individuals that carry the adaptation genes survive.

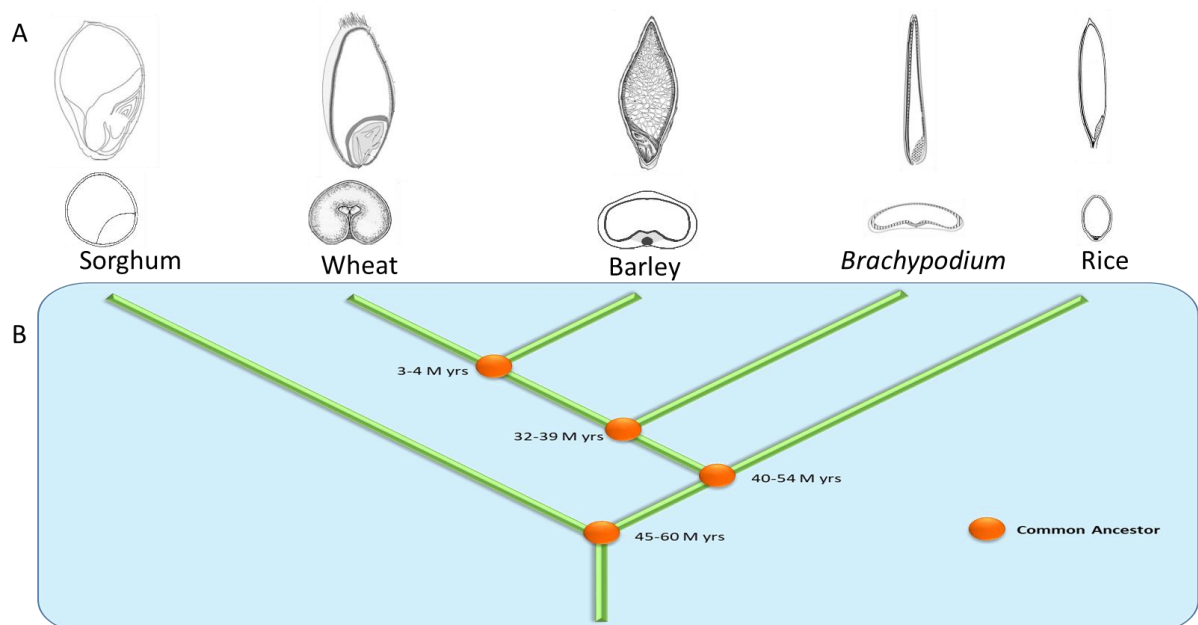


Figure 1-2: (A) Grain anatomy and (B) Phylogenetic tree showing the evolutionary relationship between some wild and domesticated grasses. M yrs: Million years. Adapted from (Hands & Drea, 2012) (WheatBP).

1.1.3 Wheat production, economical value, and its contribution to human diet and health

The domestication of cereals and many other staple plants has been linked to the evolution of human civilisation as a source of food for at least the last 3000 years. Currently, the economic importance of wheat and its impact on the diets of humans and livestock cannot be disputed. It is estimated that wheat is the third largest staple food crop and a key ingredient in cereal production around the world constituting around 15% of the world's calorie intake, after maize and rice in 2019 (<http://www.worldatlas.com>).

There is an increasing global demand for wheat in countries beyond its region of climatic adaptation. This is because of the ability to make exceptional food products due to the gluten protein fraction that facilitate the processing of wheat to produce bread, pasta and noodles, and a range of functional ingredients. These products may be more suitable to produce or consume than traditional foods (Shewry & Hey, 2015b).

The amylose and amylopectin composition, as well as protein amount, are key determinants of quality in wheat. At maturity, the wheat grain consists of 80% (w/w) of starchy endosperm (ZOOK *et al.*, 1970). The protein content of some wheat lines has been reported to range from seven to twenty-two per cent of the dry weight (Vogel *et al.*, 1976). Therefore, Wheat is considered primarily as a source of carbohydrate energy and represents a healthy source of multiple nutrients, dietary fiber and bioactive compounds, including minerals, vitamins, lipids and phytochemicals, especially if consumed as a whole-grain (Andersson *et al.*, 2013).

A meta-analysis confirmed that regular consumption of whole-grain (rich in dietary fiber) has strongly associated with reducing the risk of cardiovascular disease, caused by high cholesterol, triglycerides, blood glucose, blood pressure and also decrease the possibility of having type two diabetes significantly, and also certain forms of cancer (Kelly *et al.*, 2007; Aune *et al.*, 2016).

1.1.4 Wheat wild ancestors

It is believed that wild wheat varieties such as *Einkorn*, *Emmer*, *Khorasan* and *Spelt* have remained unchanged over the last hundred years. However, the modern species of wheat were subjected to extensive cross-breeding. The major outcome of the breeding

programmes in the fifties in many countries was the semi-dwarfing genes that were bred during the green revolution to development current varieties characterised by higher yield, an increased diseases and pest resistance, and improved tolerance to environmental stresses, a consistent maturation, short maturation cycle and a higher gluten content (Salvi *et al.*, 2013). However, these intensive breeding programs decrease genetic variability, minimise the nutritional and nutraceutical properties of the wheat, this mainly caused by the complete replacement of ancient local breeds with current varieties.

The increasing interest in ancient wheat cultivars comes from the need for nutritionally healthy food products and crop diversification. A comparison of ancient and modern wheat species underlined that *Emmer*, *Einkorn* and *Spelt* wheat have higher concentrations of magnesium, phosphorus, selenium and zinc compared to *durum* and bread wheat (Zhao *et al.*, 2009; Piergiovanni *et al.*, 1997). *Einkorn*, *Khorasan*, and *Emmer* wheat cultivars showed the highest content of total carotenoids and lutein, which are essential for growth, immune system function and eye health (Lachman *et al.*, 2013; Abdel-Aal *et al.*, 2007). By evaluating the average protein content, *Einkorn* protein values were around sixty per cent higher than bread and *durum* wheat, in addition, to contain the highest concentration of phytochemicals; phytosterols and tocols (Shewry & Hey, 2015a), which contribute to the reduction of cholesterol levels and decreasing risk of coronary heart failure and many diseases (Ogbe *et al.*, 2015).

It appears that the more we learn about the past, the better prepared we are for the future (Roosevelt, 2013). Therefore, in our study, we are using *Brachypodium distachyon* (wild grass), which shared a common ancestor with *Triticeae* around 32 million years (figure 1-2), as a model to study key genes regulating reproductive development in temperate grasses.

1.1.5 *Brachypodium Distachyon*

Studying the biology and the genetic makeup of cereals such as wheat has proved challenging due to their height, long lifecycle, susceptibility to plant microbes, limited transgenics and large genomes. Using model organisms such as *Brachypodium distachyon* with a small genomes, simple growth requirements, the ability to inbreed, short generation time and simple biological systems, enable researchers to make enormous discoveries that benefit humanity and improve their life (Vogel & Bragg, 2009).

The perennial grass *Brachypodium distachyon*, is commonly named purple false brome. It is classified in its own tribe *Brachypodieae* close to the valuable crops such as wheat and barley (Figure 1-1) and a rather a small, invasive species, currently distributer worldwide (Scholthof *et al.*, 2018), however, it was native to the Middle East and the Mediterranean basin, an area mostly covering the Fertile Crescent where the grain cereals ancestors first found (Hands & Drea, 2012; Opanowicz *et al.*, 2008).

The genus *Brachypodium* comprises around 17 perennial species distributed worldwide (Schippmann, 1991; Catalan *et al.*, 2015), however, only three independent species were intensively investigated; *B. stacei* ($2n=2x=20$, $x=10$), *B. hybridum* ($2n=4x=30$, $x=10+5$) and *B. distachyon* ($2n=2x=10$, $x=5$) (Catalán *et al.*, 2012).

Some of the pioneering research that brought *Brachypoideae* under light was by Professor Clive Stace working at Leicester University, who initiated the phylogenetic analysis in the 1980s. His study on five perennial European species of genus *Brachypodium* revealed their breeding taxonomic relationships based on morphological, anatomical, biochemical and molecular characteristics (Khan & Stace, 1999).

Subsequent studies position *B. distachyon* as an intermediate between rice and the *Triticeae* (Kellogg, 2001) (Figure 3-1). This made it a candidate for consideration as a sister group to this important section of the core *pooideae* (*Bromeae*, *Aveneae*, *Triticeae*, and *Poeae*). Therefore, it has been used as a relevant reference system for temperate cereals (Draper *et al.*, 2001) with the aim of investigating and studying cereal quality and development (Paterson *et al.*, 2009; Schnable *et al.*, 2009; Vogel *et al.*, 2010). In addition, *Brachypodium* also provides insights into the effects of cultivation and domestication as this is the first wild grass to be sequenced.

Brachypodium distachyon was first proposed as model species (Draper *et al.*, 2001) for *pooid* grass and selected for sequencing (International *Brachypodium* Initiative, 2010) because it reflects the number of advantages; such as a short life cycle, small stature, simple growth requirement and relatively small genome size (5 chromosomes= 355 Mbp). All this and more, make it ideal for laboratory research. Although *Arabidopsis* research has brought numerous scientific breakthroughs (Flavell, 2009), this dicot model system is limited when trying to fully understand monocot-specific processes.

It was estimated that *Oryza sativa* first diverged around 40-54 million years ago (Mys) then the divergence of *Brachypodium* occurred much later (around 32-39 million years ago) positioning *Brachypoideae* closer to the *Triticeae* (Figure 1-2) (Vogel *et al.*, 2006; Mur *et al.*, 2011). This enables researchers, for instance, to understand more about chromosome numbers diversification in the major grass subfamilies by nested insertions of whole chromosomes into centromeres to generate more than an ancestral set of five chromosomes (International *Brachypodium* Initiative, 2010). This ancestor is considered as protected wild grass that was unaltered or shaped by domestication pressures and yet offers a similar grain structure and displays many biological features related to agricultural traits that are of great importance in the temperate cereals, such as cold tolerance, resistance to some pathogens, meristem dormancy mechanisms and injury tolerance (Ozdemir *et al.*, 2008; Hands & Drea, 2012). *Brachypodium* offers a valuable comparative model and a comprehensive understanding of the impact of the domestication process in temperate crops (Shewry, 2009; Opanowicz *et al.*, 2011).

1.2 Reproductive Development in Grasses

1.2.1 Meristematic determinacy, flower and fruit development in grasses (*Poaceae*)

For more than a decade, there were attempts to identify genes that control floral organogenesis and fruit development, as they are considered some of the most important stages in the plant development.

Grasses species develop complex inflorescences and distinctive flowers, primary inflorescence branches are produced distichously rather than the phyllotaxis arrangement in most angiosperms. In the grass subfamily *Pooideae*, the shoot apical meristem differentiates into the primary branches of the inflorescence, controlling inflorescence phenotype and regulating grain structure (Kellogg *et al.*, 2013).

In *Brachypodium*, the inflorescence produced in a distichous pattern (Figure 1-4 (F)). At flowering time, the vegetative shoot apical meristem differentiates into three to four primary branches on the main stem (Figure 1-4 (A)). Subsequently, these branches initiate lateral organs and axillary spikelet meristems containing a group of undifferentiated cells that produce the inflorescence meristem for a complete flower (Figure 1-4 (E))(Kellogg *et al.*, 2013). This then transforms into the four whorls of the floral organs; lemma-palea in the first whorl, lodicules in the second whorl, stamens, in the third and carpels in the fourth whorl, defining the basic flower structure (Figure 3-2).

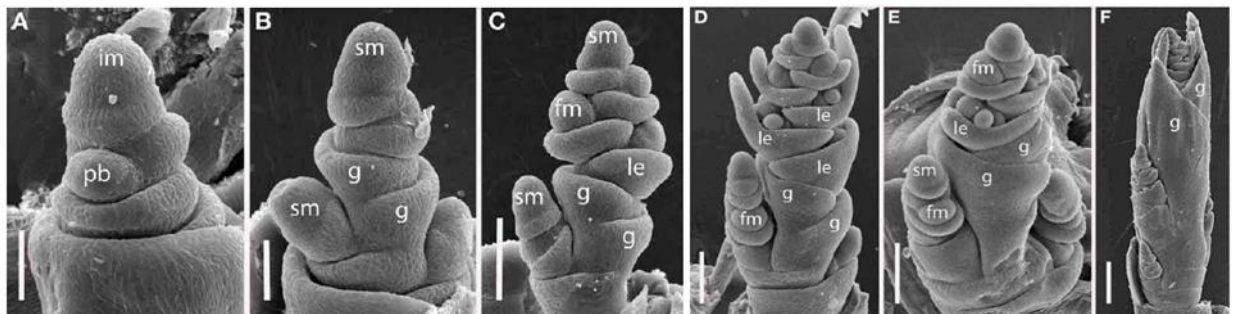


Figure 1-3: Determination of terminal and lateral spikelet stages in *Brachypodium distachyon*. (A) Distichous primary branch formation; (B) inflorescence meristem and uppermost branch meristem converted to spikelet meristems; (C–F), successive stages of development, showing differentiation of the terminal spikelet well ahead of the lateral spikelets. im, inflorescence meristem; pb, primary branch; sm, spikelet meristem, g, glume, le, lemma, fm, floral meristem. Scale bars: (A, B), 50 μm ; (C–E), 100 μm ; (F), 200 μm (Kellogg *et al.*, 2013)

The unique floral morphology of grasses also includes the distinctive fruit form of the grain. The grass caryopsis normally consists of three genetically distinguishable sections; the pericarp, the embryo and the endosperm (Hands & Drea, 2012). The maternal tissues represented by the two layers of the integument, the nucellus and the pericarp, occupy the bulk of the carpel before pollination and form a watertight layer with a mainly protective function (Percival, 1921). The endocarp is rich in chloroplasts that produce oxygen and photosynthetic materials (glucose) (Xiong *et al.*, 2013). As the grain develops the starch storage accumulated in the pericarp layer, is used for amylose synthesis in the growing endosperm and the maternal layers diminish at maturity. After double fertilisation, the diploid embryo emerges from the egg cell migrate to proximal region of the grain while the triploid endosperm derived from central cell, develop into persistent endosperm that accumulates carbohydrates, proteins and lipid, and forms the largest compartment in the mature grain (Nawaschin, 1898; Raghavan, 2006; Hands & Drea, 2012). The regulation of these developmental stages is accomplished through a very complex gene network that part of it is controlled by major transcription factor families.

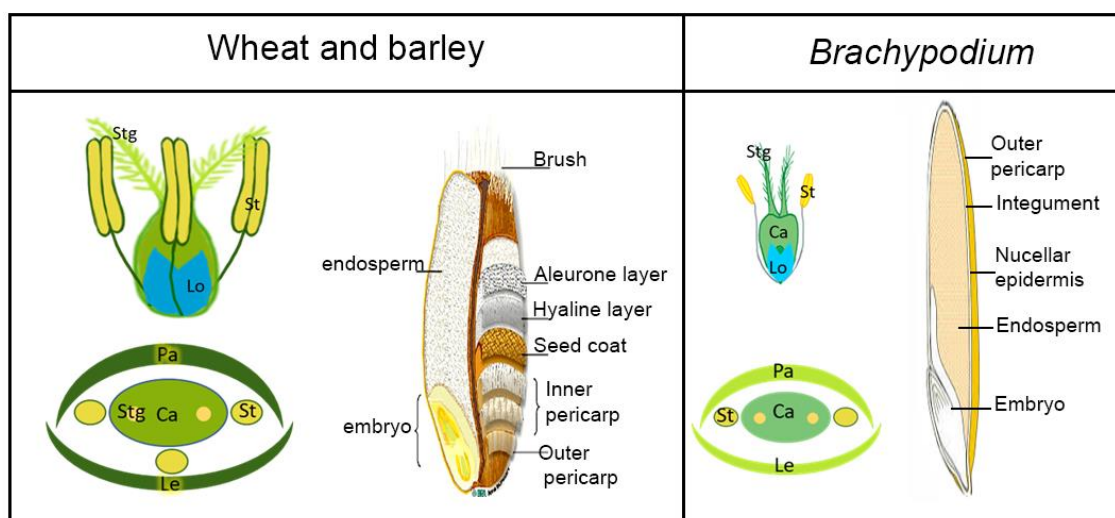


Figure 1-4: Structure of the floret and mature grain in wheat barley and *Brachypodium*. Wheat, barley and *Brachypodium* have the same floral organs except that *Brachypodium* has only two stamens and smaller floret. Ca: Carpel, Le: Lemma, Lo: Lodicule Pa: Palea, MG: Mature grain, St: stamen, Stg: Stigma. Adapted from (Heinze, 2017).

1.2.2 Grain Development

In most flowering plant species, double fertilisation is initiated by the delivery of the sperm cells by the pollen tube that ruptures in the degenerating synergid and releases two sperm cells. One sperm cell fuses with the egg cell, whereas the other one fuses with the

homodiploid secondary nucleus in the central cell, launching embryo and endosperm development, respectively (Olsen, 2004; Hands *et al.*, 2016).

Nuclear fusion in the central cell induces parental gene expression, and that results in an increase in auxin biosynthesis and influx (Forestan *et al.*, 2010). Subsequently, this activates nuclear mitosis and cell proliferation in the central cell (Hands *et al.*, 2016).

Endosperm initiation and early development

The triploid endosperm nucleus which is located in the central cell in the basal cytoplasm undergoes a synchronous division without establishing the microtubules of the phragmoplast, and no anticlinal cell wall is formed between dividing daughter nuclei (Figure 1-5). After rapid division then nuclear migration and cellularisation, the endosperm coenocyte forms the first layer of cells containing evenly spaced nuclei in the entire peripheral cytoplasm (Brown & Lemmon, 2001) (Figure 1-5). The syncytial endosperm cellularises from the outer side in radial form to fill the central vacuole and enters a maturation process storing energy-rich molecules including starches, proteins, and lipids to provide nutrient to the embryo development and growth (Olsen, 2004). This was observed in *Arabidopsis*, rice, barley and wheat (figure 1-5).

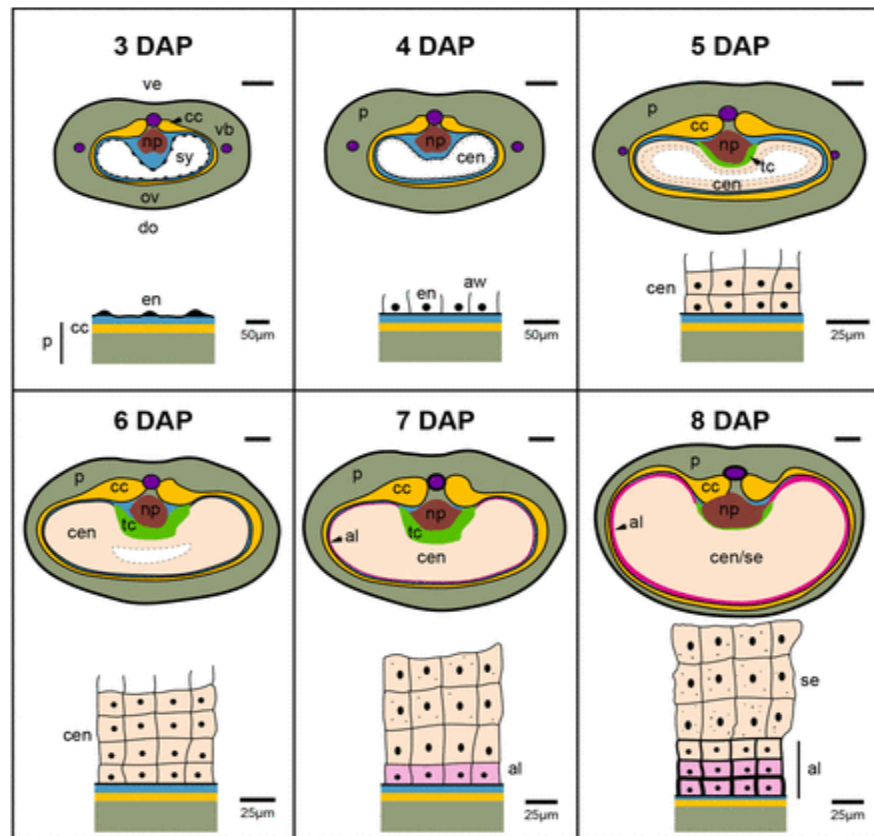


Figure 1-5: Schematic representations of the developmental stages of barley caryopsis.

Transverse sections showing the early events of endosperm development. (al, Aleurone; aw, anticlinal walls; cc, chlorenchyma layer; cen, cellular endosperm; do, dorsal; en, endosperm nuclei; np, nucellar projection; ov, ovary; P, pericarp; se, starchy endosperm; sy, syncytium; tc, transfer cells; vb, vascular bundle; ve, ventral from (Zhang *et al.*, 2016).

In major core *poooids* including wheat and barley, the endosperm remains as a persistent structure at maturity stage reducing the space for the rest of the grain tissues. It consists of four main cell types; transfer cells, aleurone cells, starchy endosperm cells, and embryo-surrounding region cells. The endosperm is enclosed within layers of nourishing cells in the nucellar epidermis and nucellar projection, which compress and degenerate with the protective shield of the pericarp and integument (seed coat) during maturation (Figure 1-5) (Hands *et al.*, 2012).

The maternal tissues control nutrient supply routes to the developing endosperm. In the wheat grain, it is believed that generally the nucellar projection and modified aleurone (transfer cells) provide a channel for this transport from the vascular bundle and nucellus (Wang *et al.*, 1994). Although the organisation of the aleurone layers varies among *Pooideae*, it conserves its properties of storing large quantities of oils and neutral lipids in the oleosomes (membrane-bound organelles) and also containing some hydrolytic enzymes. These lipoxygenase and glyoxylate cyclase are involved in lipid breakdown and gluconeogenesis, in addition to the enzymes involved in the degradation of the starch and proteins stored in the starchy endosperm cells during germination (Doig *et al.*, 1975; Holtman *et al.*, 1994; Jones, 1972; Jones, 1969).

1.2.3 ABCDE model floral organogenesis

Intensive studies in *Arabidopsis* settled on the ABCDE floral model that states the identity of floral organs is determined by the homeotic genes belonging to A, B, C, D and E classes (Theissen *et al.*, 2000; Rijpkema *et al.*, 2007; Rijpkema *et al.*, 2010)(Figure1-6).

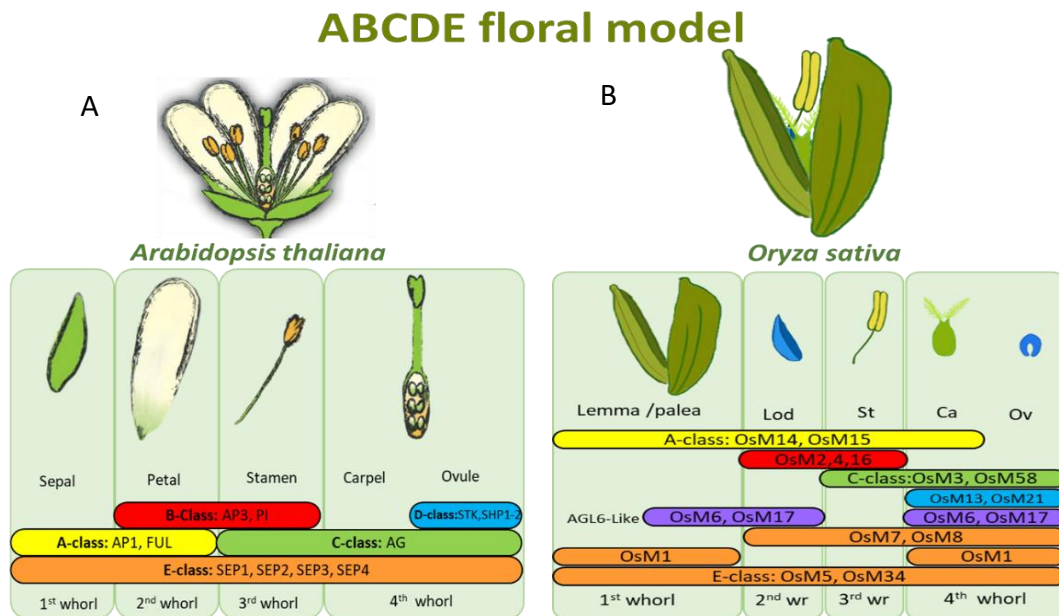


Figure 1-6: ABCDE model summarising the homeotic function and the expression profile of *MADS-box* genes, (A) in *Arabidopsis* based on (Theißen *et al.*, 2016) and in *Oryza sativa* from (Wu *et al.*, 2018) and (Yoshida & Nagato, 2011). wr: whorl, OsM: OsMADS, Lod: lodicule, St: stamen, Ca: carpel, Ov: ovule.

This model describes how five classes of homeotic genes identify floral organs in specific whorls of the flower (Rijkema *et al.*, 2010). According to the floral quartet model of floral organ specification, proteins from different classes interact and bind to structure multimeric complexes in order to regulate its gene targets (Smaczniak *et al.*, 2012). In *Arabidopsis thaliana*, the A- and E-class protein complex are specifically involved in the development of the sepals in the first floral whorl, the A-, B- and E-class protein complex identify petals in the second whorl, the B-, C- and E-class protein complex are required for stamens development in the third whorl, and the C- and E-class protein complex specify gynoecium in the fourth whorl. The D-class proteins interact in a larger complex with the E-class proteins to specify ovule identity (Figure 1-6 (A)).

A similar model was established for rice and suggests that the ABCDE model might equally apply to monocots species such as *Brachypodium*, wheat and barley. However, recent studies on rice (Wu *et al.*, 2018) showed a different expression pattern of these genes in the floral whorls. This suggests that the dicot model *Arabidopsis* does not reflect fully what occurs in monocot plant, as *Oryza sativa* presented a large number of genes acting in the same class (Figure 1-6 (B)). Ultimately, tropical grasses could give only a limited understanding on temperate grasses such as wheat and barley that phylogenetically far from rice.

1.2.4 The role of transcription factors (TFs) in plants

Eukaryotic genomes contain 3-10 % of genes encoding transcription factors that play a major role in the development of living organisms (Levine & Tjian, 2003). Generally, transcription factors are proteins that regulate the expression of genes by annealing with cis-regulatory elements in the promoter of their target genes. They act as repressors or activators of genes, either alone or with other proteins, forming multimeric protein complexes (Chen *et al.*, 1997). This function can also be achieved through recruiting and interaction with chromatin-modifying enzymes which facilitate and modulate the accessibility to a specific region of the genome (Gonzalez, 2015). Consequently, transcription factors are classified and named in families based on the DNA-binding domain specificity (Stegmaier *et al.*, 2004; Vaquerizas *et al.*, 2009; Charoensawan *et al.*, 2010). The activation of some genes can be achieved also through establishing protein-protein interaction with other elements of the transcription machinery or participation of chromatin changing enzymes (Gonzalez, 2015). There are four major super-classes of TFs; (1) containing basic domain: bZip and bHLH, (2) Zinc-coordinating DNA binding domains, such as Zinc finger and Cys6 cysteine-zinc cluster, (3) Helix-turn-helix: TEA domain, Homeodomain and Paired box, (4) beta-scaffold factor with minor groove contacts: Rel homology region (RHR), p53, STAT and *MADS-box* (Gonzalez, 2015).

1.2.5 *MADS-box* transcription factors

In plants, *MADS-box* transcription factors are probably the most studied family of transcriptional regulators due to their essential roles in various developmental processes such as floral organogenesis and fruit formation, in addition to their involvement in response to natural changing environments and stress conditions (Yamaguchi *et al.*, 2006; Lopez-Dee *et al.*, 1999; T. Zhao *et al.*, 2006; Castelán-Muñoz *et al.*, 2019; Hands & Drea, 2012). In other organisms, this family contributes to a wide range of biological processes, fluctuating from muscle development, cell proliferation and differentiation in animals to pheromone responses in fungi (Messenguy & Dubois, 2003).

The *MADS-box* transcription factor family was identified based on the presence of a conserved motif of 60-amino-acid long found within the DNA-binding domain at the N-terminal of the sequence. This domain recognises a CA-rich-G-box site or the so-called “Serum Response Element (SRE)-type CArG box” (Verelst, Saedler, & Munster, 2007). The first *MADS-box* gene isolated was *ARG80* from *Saccharomyces cerevisiae* (budding yeast) (Dubois et al., 1987; Bercy et al., 1987), however, the term was created later after four genes were subsequently characterised: *MCM1* (minichromosome maintenance protein 1 from budding yeast), *AGAMOUS* from the thale cress *Arabidopsis thaliana* (Yanofsky et al., 1990), *DEFICIENS* from snapdragon *Antirrhinum majus* (Schwarz-Sommer et al., 1990) and SRF (Serum response factor) from human *Homo sapiens* (Riechmann & Meyerowitz, 1997).

1.2.6 The diversity of *MADS* domain transcription factor

These *MADS-box* transcription factors were found in numerous eukaryotic organisms, including Protista, animal, fungi and plants. A detailed understanding of *MADS-box* gene phylogeny may consequently highlight the diversity and the phenotypic development of plants on land.

An ancestral gene duplication that took place around 2 billion years ago (Figure 2-1) led to limited extension in nearly all major eukaryotic groups but a large diversity of the *MADS-box* family in land plants. The subsequent divergence that occurred 1 BYA, split the family into two types termed type I and type II in land plants, their orthologs in animal are called SRF and MEF2, while in fungi, ARG and RLM1 (Gramzow *et al.*, 2010; Becker & Theißen, 2003). These groups are distinguished by their number of exon-intron, domain construction and their degree of functional redundancy. For instance, MADS-SRF and MADS-MEF2 domains vary in sequence specificity in DNA binding and the amount of DNA bending that they induce in the promoter (Huang *et al.*, 2000). In land plants, type I is characterised by one or two exons that encode for MADS-SRF like domain and further subdivided into three groups named, M α , M β and M γ , that lack the keratin-like (K) domain (Figure 2-1). While, type II genes (commonly referred to as *MIKC-type*) have an average of seven exons that translate into MADS-MEF2 domain (MYOCYTE ENHANCER FACTOR 2), intervening domain (I), keratin-like domain (K) and carboxy-terminal (C) domains (De Bodt *et al.*, 2003). Subsequently, type II MADS-domain proteins are subdivided into the groups MIKC^C and MIKC* based on the number of exons that encode the intervening domain (I)

and structural differences in the keratin-like (K) domain (Becker & Theißen, 2003; Henschel *et al.*, 2002; Parenicová *et al.*, 2003).

The MIKC^C (MIKC-type) transcription factors contain the well-conserved MADS-domain in addition to three additional functional domains; the Intervening (I) domain that influences the specificity of DNA-binding, protein-protein dimers and predicted to form an α -helix, this is followed by Keratin-like (K) domain that is involved in protein-protein interaction - after the MADS domain, it is the most highly conserved domain of MIKC-type proteins, and it has been subdivided into three subdomains K1, K2, and K3 (Kaufmann *et al.*, 2005). Finally, the C-terminal domain that is involved in transcriptional activation and forming higher-order multimeric MADS protein complexes (Gramzow & Theissen, 2010). The I domain and the C-terminal domain are less conserved than the MADS and K domains.

Functional studies link variations in these sub-clades to sub-functionalisation. For instance, studies on *Arabidopsis* revealed that type I *MADS-box* genes are important for female gametophyte, embryo sac and seed development (Bemer *et al.*, 2009; Yoo *et al.*, 2006; Colombo *et al.*, 2008; Kang *et al.*, 2008; Bemer *et al.*, 2008). While, type II *MADS-box* genes containing MIKC* and MIKC^C group have been suggested to have in general critical roles in gametophyte development in all land plants, such as SHP1 and SHP2 from C-class that involved in seed dehiscence and also members of class A, AP1, CAL and FUL that play a crucial role in flowering time (Pařenicová *et al.*, 2003; Rijpkema *et al.*, 2007). In animal, *MADS-MEF2* box found to regulate muscle differentiation (Olson *et al.*, 1995).

In land plant, early studies have identified *MADS-box* as homeotic genes due to the spectacular phenotypes they presented in mutation lines. However, recently it is becoming clear that these genes are involved in several aspects of plant growth and development, including male and female gametophyte development, embryo and seed development, as well as root, flower and fruit development.

1.2.7 The function of C and D class genes

The C and D class genes are the most important genes because they are responsible for the morphogenesis of the male and the female sex organs, respectively. The C class genes

regulate mainly the stamens that are in the third whorl while the D class are dominant in the carpel and ovule in the fourth whorl.

In rice, two members were identified from C-class called *OsMADS3* and *OsMADS58*. These two paralogs were found to be expressed in the floral primordia organs at the very early stages increasing and becoming more restricted to the third whorl (stamens; male gametophyte) for *OsMADS3*, while, *OsMADS58* expression was distributed uniformly in the third and fourth whorl (female gametophyte) at the pre-anthesis stage. Their mutations suggest that both genes specify floral meristem determinacy; however, *OsMADS3* plays a dominant role in specifying the third whorl, while, *OsMADS58* is focused more on regulating carpel morphogenesis (Yamaguchi *et al.*, 2006). With regard to the D-class, two genes *OsMADS13* and *OsMADS21* were identified in rice and found to be expressed in the developing ovules (Lopez-Dee *et al.*, 1999). *OsMADS13* function in meristem determinacy and ovule identity was established. However, although there are paralogs in D-class, *Osmads13* mutant ovules were converted into carpelloid structures, and *Osmads21* mutant shows that it has no function in ovule identity determination (Dreni *et al.*, 2007), suggesting that there was loss-of-function in *OsMADS21* following the duplication events. In *Arabidopsis*, two genes from C-class *SHP1* and *SHP2* and one gene from D-class *STK* were found to control ovule identity (Pinyopich *et al.*, 2003) and these finding led to believe that there might be some functional conservation between the two species. However, there was no similar indeterminacy control observed in *Arabidopsis* and *Petunia* mutants. The rice (monocot plant) flower anatomy and its developmental program clearly vary from the one of *Arabidopsis* (dicotyledon plant), and therefore floral developmental processes are more likely to be different. From here comes the important to study a model system such as *Brachypodium* that is close to the important crops such as wheat and barley but does not present the challenge of big genomes and long-life cycle.

Our important crops are facing many challenges in the future including a rapidly growing population, climate changes and the rising level of pests that evolved resistance to pesticides. Therefore, the need to understand genes involved in the essential developmental stages is important to prepare for future challenges such as increasing grain yield and improving grain quality for various end uses.

Thesis aims and objectives:

A central aim of the thesis is to investigate the roles of C and D class *MADS-box* genes in grain development using *Brachypodium* as a reference species in comparison to related temperate crops barley and wheat. The objectives were as follows:

- Perform a genome-wide identification of all *MADS-box* genes in wheat, barley, rice and *Brachypodium*, and identify how they are phylogenetically related. Then using the well-characterised genes from rice and *Arabidopsis* to cluster and name genes from wheat, barley and *Brachypodium*.
- Dissect the role of C and D-class genes in the grass model system *Brachypodium distachyon* by investigating their temporal and spatial expression pattern through plant developmental stages.
- Explore their function using available transgenic RNAi lines.
- Identify potential targets using transcriptome analysis.

Chapter 2

Phylogeny and expression of *MADS-box MIKC-type* genes
in wild and domesticated grasses.

Chapter 2. Phylogeny and expression of *MADS-box MIKC-type* genes in wild and domesticated grasses

2.1 Abstract

MADS-box genes are important transcription factors in controlling many developmental processes of angiosperms. First was identified as floral homeotic and organ identity genes then quickly became apparent that they have roles through the whole lifecycle of the plant and play roles within the gene networks that govern vegetative and reproductive organs. Therefore, studying these important genes in high economic value species such as rice (*Oryza sativa*), wheat (*Triticum aestivum*), and barley (*Hordeum vulgare*) is valuable especially in the context of increasing demand and the impact of changing environments on reproduction and yield. We report here a comprehensive phylogenetic and structural analysis of *MADS-box* MIKC-type transcription factors in five species including *Arabidopsis thaliana*, *Oryza sativa Japonica*, *Triticum aestivum*, *Hordeum vulgare* and *Brachypodium distachyon*, representing the dicot plant model system, wild and the domesticated monocot grasses. We illustrate their expression pattern in several stages in the life cycle using RNA-seq data and elaborate on four candidate genes that have been characterised in *Arabidopsis* and rice (*MADS3/58/13* and *21*).

2.2 Introduction

The *MADS-box* transcription factor family was identified based on the presence of a conserved motif of 60-amino-acids found within the DNA-binding domain at the N-terminal of the sequence. The first *MADS-box* gene isolated was ARG80 from *Saccharomyces cerevisiae* (budding yeast) (Dubois *et al.*, 1987; Bercy *et al.*, 1987) and the term was generated after four genes were subsequently characterised, *MCM1* (minichromosome maintenance protein 1 from budding yeast), *AGAMOUS* from the thale cress *Arabidopsis thaliana* (Yanofsky *et al.*, 1990), *DEFICIENS* from snapdragon *Antirrhinum majus* (Schwarz-Sommer *et al.*, 1990) and *SRF* (Serum response factor) from human *Homo sapiens* (Riechmann & Meyerowitz, 1997). Some functional studies first identified *MADS-box* transcription factor as homeotic genes due to the spectacular phenotypes they presented in mutation lines. However, recently it is becoming clear that these genes are involved in several aspects of plant growth and development, including male and female gametophyte

development, embryo and seed development, as well as root, flower and fruit development.

The diversity of *MADS* domain transcription factor

These *MADS-box* transcription factors were found in numerous eukaryotic organisms, including Protista, animal, fungi and plants. A detailed understanding of *MADS-box* gene phylogeny may consequently highlight the diversity and the phenotypic development of plants on land.

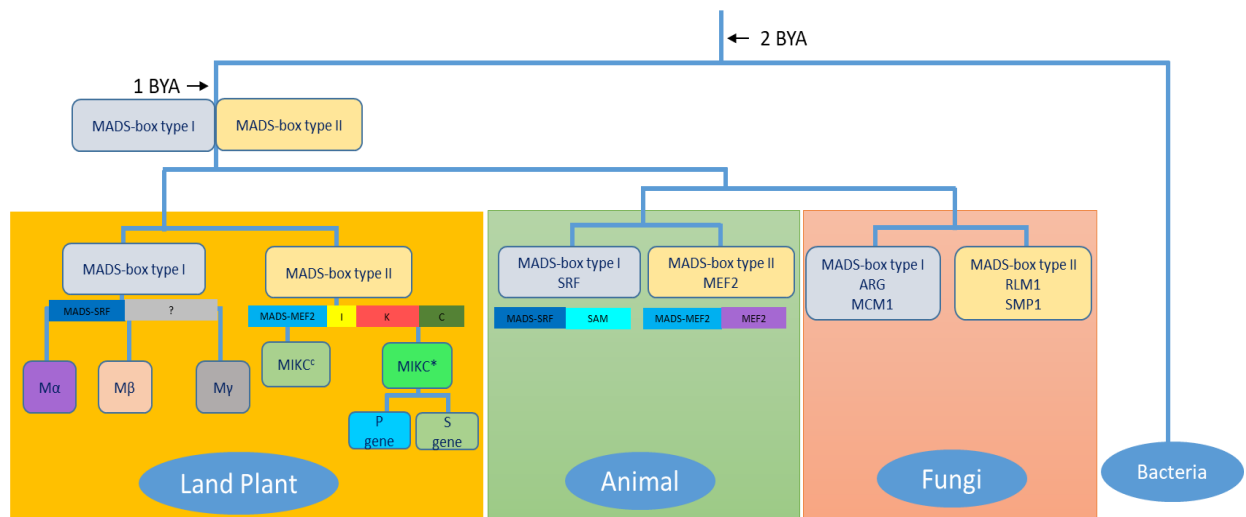


Figure 2-1: The major clade of *MADS-box* genes. The figure was adapted from (Becker & Theißen, 2003) and adjusted to illustrate the protein domains. BYA: billion years ago.

An ancestral gene duplication that took place around 2 billion years ago (Figure 2-1) led to the diversification of The *MADS-box* family in different ways in the various eukaryotic lineages during evolution as a result there was divergence in three kingdom fungi with ARG MCM1 and RLM1 SM1, animal with their distinctive SAM and MEF2 domain and land plant. Subsequent gene duplication events led to the further diversification of *MADS-box* genes in the latter. In land plants two fundamentally different types of plant *MADS-box* genes have been identified, *MADS-box* type I which contain *MADS-SRF* like domain and further subdivided into three groups named, Mα, Mβ and Mγ (Figure 2-1) these groups are distinguished by exon-intron and domain structure. The second is *MADS-box* type II that have four domains, *MADS-MEF2* DNA binding (MYOCYTE ENHANCER FACTOR 2), Interference domain, Keratin-like domain and C terminal. This sub-clade is further subdivided into two groups, MIKCC and MIKC* based on the size of the K domain, which is small in MIKCC. The *MADS-SRF* and *MADS-MEF2* domains vary in sequence specificity in DNA binding and the amount of DNA bending that they induce.

Functional studies link these variations to the sub-functionalisation resulted in the ancient duplication for the different classes of genes. For instance, studies on *Arabidopsis* revealed that type I *MADS-box* genes are important for female gametophyte, embryo sac and seed development (Bemer *et al.*, 2009; Yoo *et al.*, 2006; Colombo *et al.*, 2008; Kang *et al.*, 2008; Bemer *et al.*, 2008). While, type II *MADS-box* genes containing MIKC* and MIKC^C group have been suggested to have in general critical roles in gametophyte development in all land plants, such as SHP1 and SHP2 from C-class that involved in seed dehiscence and also members of class A, AP1, CAL and FUL that play a crucial role in flowering time (Pařenicová *et al.*, 2003; Rijpkema *et al.*, 2007). In animal, *MADS-MEF2* box found to regulate muscle differentiation (Olson *et al.*, 1995).

Since the dicotyledon plant, *Arabidopsis* and monocotyledonous *Brachypodium*, barley and wheat are distantly related plant species; it is interesting to investigate whether the floral organ identity factors have been conserved in their functions, and understand the variation that happened during the species divergences.

Here we first identified the *MADS-box MIKC-type* in five species and classified the one in barley and wheat, based on the one characterised in rice and *Arabidopsis*, then investigate the expression pattern of candidate gene from the *AGAMOUS* family in wheat and barley. These results provide a comprehensive analysis of the sub-family *MADS-MEF2-like*, taking into consideration the number of paralogues in each species and orthologs in-between species, highlighting the members that were present or absent.

2.3 Materials and Methods

2.3.1 Plant materials growth conditions

Bread wheat seeds (*Triticum aestivum*) (wild-type) was kindly provided by Dr James Higgins lab. Barley (*Hordeum vulgare*) seeds were obtained from Dr Sinead's lab. The seeds were de-husked and grown as described (Drea *et al.*, 2005). All the seeds were sterilised for 10 minutes in 70% ethanol with 0.05% Triton® then washed with 100% ethanol for 2 minutes then air-dried on sterile Whatman® filter paper, and vernalised for 48 hours at 4 °C to accelerate flowering stage. The seeds were sown in the vermiculite (Dupré Minerals Limited) and soil (Levington® seed and modular compost) mixture (1/3) to retain moisture and improve aeration in pot size of 360 cm³ for germination, and kept in a phytotron under

the following growth conditions: 51% humidity, 23 °C for 20 hours photoperiod, 4 hours darkness. Plants were watered by flooding under tray every other two days.

2.3.2 Plants sampling

Plant samples were manually collected from different stages of the flowering phase starting from very young inflorescence at the formation of the meristem floral organs primordia inside the flag leaf, until mature grain stage for the two species (*Triticum aestivum* and *Hordeum vulgare*). All samples were flash-frozen in liquid nitrogen straight after collection then stored at -80°C for RNA extraction.

2.3.3 Molecular analysis

The total RNA was extracted from frozen ground tissues using TRIsure reagent (BIOLINE) for RT-PCR, and the ISOLATE II RNA extraction kit (BIOLINE) for the quantitative RT-PCR due to its excellent purification of the RNA from any inhibitors, following the manufacturer instructions. The samples were then quantified using spectrophotometer NanoDrop™ 2000 UV-Vis (Thermo Scientific), and based on that concentration, a 700 ng of the Total RNA was used to synthesize cDNA using Tetro Reverse Transcriptase enzyme (200u/ µl) (BIOLINE). RT-PCR was performed on 1:5 dilution from the cDNA with the corresponding primers and Dream Taq Green PCR Master Mix (2X) in the total volume of 10µl, then amplified for ~30 cycles. After that it was loaded in 1% w/v agarose gel with 0.5 nM ethidium bromide then visualised in a transilluminator. Quantitative RT-PCR reactions were carried out using SYBR Green (SensiMix SYBR Low-ROX Kit) (BIOLINE) on Applied Biosystems™ 7500 Real-Time PCR Systems with 7500 Software v2.0.6, on 96well MicroAmp® Fast Optical 96-Well Reaction Plate.

Real-time quantitative RT-PCR

Primers were designed using Integrated DNA Technologies (IDT) with an amplicon size of around 100bp to anneal efficiently in 65°C and contained 50% GC. For wheat, mostly the common 5'UTR/3'UTR regions between A, B and D was targeted with no specification of alleles. Primer sequences are provided in Table S5. Detailed protocol is described in chapter 3.

2.3.4 Sequence and phylogenetic analysis

2.3.4.1 Orthologs identification

Four rice *MADS-box* *OsMADS3/58/13/21* genes ID from a previous study (Dreni *et al.*, 2007) were used as a starting point using the transcript sequence from *Oryza sativa Japonica* Group (IRGSP-1.0). BLAST searches were carried out against *Brachypodium distachyon* (Bd_v3.0) using Ensembl Plants (<http://plants.ensembl.org/index.html>), which resulted in finding four genes that present a high similarity to the rice with an E-Value of $\approx 6.7E-66$ and 94.6% identity. Another BLAST search was performed on *Triticum aestivum* (RefSeqV1.0) with a similar result with three copies for each gene represented in the A, B and D homoeologues. The orthologs in *Brachypodium* and wheat were given the name similar to the characterised one in rice.

To identify the putative *MADS-box* MIKC-type proteins in *Brachypodium distachyon*, *Oryza sativa*, *Triticum aestivum*, *Hordeum vulgare* and *Arabidopsis thaliana*, used to construct the phylogenetic tree (Figure 2-4). A BioMart search was performed using the Interpro ID (IPR033896) scanning for the MADS MIKC type domain.

2.3.4.2 Full identification of all *MADS-box* MIKC-type transcription factor in five species.

InterPro ID (IPR033896) was used to scan for *MADS-box* MIKC-type domain which binds to DNA in the CArG-box motif. The search was performed using BioMart software (Smedley *et al.*, 2015) from EnsemblPlants (<http://plants.ensembl.org/index.html>), in TAIR10 for *Arabidopsis thaliana*, IRGSP-1.0 for *Oryza sativa Japonica*, IWGSC for *Triticum aestivum*, IBSC_v2 for *Hordeum vulgare* and Bd_v3.0 for *Brachypodium distachyon*. The full-length protein sequences result were cleaned from duplicated sequences and filtered for the longest protein sequence using Geneious R10 (<https://www.geneious.com>). The total number was 326 sequences including; *Arabidopsis thaliana* 47 TFs, *Oryza sativa Japonica* 43 TFs, *Triticum aestivum* 155 TFs, *Hordeum vulgare* 41 TFs and *Brachypodium distachyon* 40 TFs.

2.3.4.3 Protein sequences alignment

Using MEGA10, full-length protein sequences from the selected five species were submitted to select, substitution model for *MADS* transcription factor subfamily. WAG+G+F was used based on the lowest BIC scores (Bayesian information criterion), AICc (Akaike

information criterion, corrected) and highest lnL (log likelihood value). The alignment was performed in Geneious R10 (<https://www.geneious.com>) using MAFFT v7.308 (Katoh & Standley, 2013) under the following parameters: protein model: WAG+G+F and JTT+G+F, BLOSUM62 as scoring matrix, Gap open penalty of 1.53 and an offset value of 0.123, scoring matrix: BLOSUM62.

2.3.4.4 Conserved domain and motif analysis and their distribution

The consensus for *MADS-box* domain was identified in (Theißen *et al.*, 1996) in some eukaryotes plant species of about 180 base pairs in length. In our study, we used the same consensus with a wide number of mismatches to generate a known consensus specific to our targeted species from wild and domesticated grasses. The same procedure was followed for the I domain from (Wei *et al.*, 2014) and K domain from (Kaufmann *et al.*, 2005) using SMARTBLAST (<http://blast.ncbi.nlm.nih.gov/smartblast/>) from NCBI (Coordinators, 2017). For C-class TFs motifs that were identified in (Kramer *et al.*, 2004) study was used in the alignment. The consensus of AG motif I sequence is YDSRNFLQXNIM, whereas, the AG motif II is QQPQXYSQ for the five species selected and also conserved in many angiosperm AG-homologs (Kramer *et al.*, 2004).

2.3.4.5 Phylogenetic tree construction

The uploaded sequences were filtered and checked for duplications then aligned in Geneious R10 software. The alignment was performed by MAFFT (v7.308) (Katoh & Standley, 2013; Katoh *et al.*, 2002), using automatic algorithm and BLOSUM62 as a scoring matrix, the Gap open penalty was 1.53 with an offset value of 0.123. The tree was constructed using maximum likelihood phylogenetic method with Jukes-Cantor as genetic distance Model, and Bootstrap as resampling with 100 replicates, the threshold set was 50%.

2.3.5 Transcriptome Analysis

2.3.5.1 Wheat (*Triticum aestivum*) RNA-Seq expression

After obtaining the sequences of *MADS-box* MIKC-type family using Interpro ID scanning for the *MADS* domain. Only the IDs were selected and uploaded in the (<http://wheat-expression.com/>) database which contains Transcriptomic analyses from 28 studies for *Triticum aestivum* in various tissues. A heat-map was constructed using the Log2 TPM values for all 155 genes of wheat *MADS-box* MIKC-type family, then filtered for the genes

that their total expression were zero in order to have only the genes that expressed in the available tissues samples. Some duplicated samples were removed as well. The heat-map was built in Rstudio in R (RCore, 2016) with pheatmap package.

2.3.5.2 Barley (*Hordeum vulgare*) RNA-Seq expression

Forty-one genes were obtained from Interpro Scan in *Hordeum vulgare* genome, and these genes IDs were submitted in the publicly available database Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) and extract the expression values from eight barley tissues from different developmental stages (International Barley Genome Sequencing Consortium, 2012). In this study, only 26 genes expression were found out of the 41 genes identified. The heat-map was constructed using (<https://biit.cs.ut.ee/clustvis/>).

In order to generate a heat-map for all *MADS-box* genes in *Triticum aestivum*, (Jian Ma *et al.*, 2017) study where they used Genome-wide identification and analysis of the *MADS-box* gene family in bread wheat was examined. 182 gene IDs were generated from this study however it was with the old IWGSC assembly, 2014, therefore, we converted each gene ID manually to the new assembly TGACv1, 2017 using Ensembl Plants database and we generated a list of 153 unique IDs. Their expression pattern was examined in three different studies. In *Brachypodium* only one study (Bo Wei *et al.*, 2014) tried to identify *MADS-box* TFs and selected only 32 MIKC^c sequences with old accession number which make it difficult to find in the latest version v3 of *Brachypodium* assembly - our study identified 40. The heat-map was generated using RStudio using source "https://bioconductor.org/biocLite.R" and the ComplexHeatmap library. These results were dropped from the chapter and only listed in the appendix (Figure S12) because this list in this study was not completed due to the old assembly used.

2.3.6 Genomic distribution of *MADS-box* genes

The locations of the identified *MADS-box* MIKC-type for *Arabidopsis thaliana*, *Triticum aestivum*, and *Oryza sativa* were illustrated using (<https://knetminer.rothamsted.ac.uk/KnetMiner/>).

2.4 Results

2.4.1 C and D-class *MADS-box* Genes ortholog Identification





MADS-box genes are essential transcription factors for establishing floral organ identities (Alvarez-Buylla *et al.*, 2000). According to their roles in flower development and depending

on the organ they specify, *MADS-box* genes are classified as having A, B, C, D and E functions (Pelaz *et al.*, 2000). In this study, we focused on the C and D class genes involved in fruit and seed development and identified four genes (Table 2-1) that we predict will be involved in grain development. In grasses, *MADS-box* genes have been especially well studied in *Oryza sativa*. Therefore, rice sequence was primarily used in BLASTp searches and compared to wheat and other temperate cereal species for corroboration. We successfully identified sequences listed in table 2-1 and 2-2 as putative *Brachypodium* orthologs for these genes.

2.4.2 Gene Identification and analyses

Due to the importance of *MADS-box* genes in the fruit organ identification and grain development, we have focused on two candidate genes from C and two from D families. The transcription factors protein sequences with known functions in rice were also indicated in table 2-1 below.

Table 2-1: Identification of *Brachypodium* *MADS* Transcription factors and their orthologs in *Oryza sativa*.

Class (MIKC)	<i>Oryza sativa</i> sequences used in BLAST homology identification	Identity % to rice	BLAST E-Value	<i>Brachypodium</i> Gene name and ID	Chromosome and location in Bd
C-class	<i>OsMADS3</i> Os01g0201700	89.8%	8.7E-68	<i>BdMADS3</i> Bradi2g06330	Ch 2 4,754,368-4,765,920 Mb 
	<i>OsMADS58</i> Os05g0203800	91.2 %	5.4E-54	<i>BdMADS58</i> Bradi2g32910	Ch 2 32,917,648-32,928,228 Mb 
D-class	<i>OsMADS13</i> Os12g0207000	88.8%	7.9E-47	<i>BdMADS13</i> Bradi4g40350	Ch 4 44,746,080-44,751,256 
	<i>OsMADS21</i> Os01g0886200	90.3%	5.6E-49	<i>BdMADS21</i> Bradi2g25090	Ch 2 22,912,795-22,918,817 

These genes play a crucial role in floral development, grouped according to their gene function classes. They are used as a reference in the BLAST analysis at protein and nucleotide levels against *Brachypodium distachyon* genome in GenBank as well as Phytozome v12.1 and Ensembl Plant (<https://plants.ensembl.org/index.html>). High-scoring candidate genes in *B. distachyon* were identified as orthologs.

Table 2-2: Selection of wheat (*Triticum aestivum*) C and D class orthologs in *Brachypodium* and barley and its homoeologs (ABD).

Class	Gene name (Jian Ma, 2017)	Homoeologs	RefSeqV1.0 gene ID <i>Triticum aestivum</i>	TGACv1 Assembly Gene ID. <i>Triticum aestivum</i>	Protein length	UniProt	GenBank	<i>Brachypodium</i> ID (Wei et al, 2014)	<i>Hordeum vulgare</i> gene ID
C-class	<i>TaMADS3-A/ TaMADS50/ TaWAG2-A</i>	A	TraesCS3A02G31430 0	TRIAE_CS42_3AS_TGA Cv1_211344_AA06888 60	273 aa	COSTS7	mRNA:KT1887 82 Protein ALM58841	<i>BdMADS14</i>	HvMADS3 HORVU3Hr1G026650
	<i>TaMADS3-B/ TaWAG2-B</i>	B	TraesCS3B02G15750 0	TRIAE_CS42_3B_TGACv 1_220593_AA0710240	276 aa	A0A192YEI5	HG670306.1		
	<i>TaMADS3-D/ TaWAG2-B TaMADS71</i>	D	TraesCS3D02G14020 0	TRIAE_CS42_3DS_TGA Cv1_272039_AA09136 20	273 aa	A0A0S1W6 03 WAG-2m	KT188779.1 WAG-2g		
	<i>TaMADS58-A/ TaWAG1-A/ TaMADS8</i>	A	TraesCS1A02G12580 0	TRIAE_CS42_1AS_TGA Cv1_018970_AA00571 30	269 aa	A9J1W2 WM2		<i>BdMADS18</i>	HvMADS58 HORVU1Hr1G029220
	<i>TaMADS58-B/ TaWAG1-B/ TaMADS13-12</i>	B	TraesCS1B02G14480 0	TRIAE_CS42_1BS_TGAC v1_049537_AA015615 0	317 aa	Q8GTP4/A0 A1D5SJH4	Protein: BAC22939 mRNA: AB084577		
	<i>TaMADS58-D/ TaWAG1-D/ TaMADS22</i>	D	TraesCS1D02G12770 0	TRIAE_CS42_1DS_TGA Cv1_080255_AA02443 60	269 aa	G9JLx4	AM502863.1		
D-class	<i>TaMADS13-A/ TaMADS103</i>	A	TraesCS5A02G11750 0	TRIAE_CS42_5AS_TGA Cv1_393214_AA12698 60	254aa			<i>BdMADS2</i>	HvMADS13 HORVU1Hr1G023620
	<i>TaMADS13-B/ TaMADS111</i>	B	TraesCS5B02G11510 0	TRIAE_CS42_5BS_TGAC v1_423180_AA136941 0	252aa				
	<i>TaMADS13-D/ TaMADS115</i>	D	TraesCS5D02G11820 0	TRIAE_CS42_5DS_TGA Cv1_457116_AA14825 80	252 aa	TaAGL31/A GL9: Q1G161 WM3A: A0A0A0P41 8	mRNA:AM502 864.1 (Zhao et al. 2006b) (Paolacci et al., 2007) protein: CAM59042.1 (WM3A gene)		
	<i>TaMADS21-A/ TaMADS2</i>	A	TraesCS1A02G26270 0	TRIAE_CS42_1AL_TGAC v1_002892_AA004596 0	251 aa	W4ZPM0 WM27B		<i>BdMADS4</i>	Absent
	<i>TaMADS21-B/ TaMADS9</i>	B	TraesCS1B02G27330 0	TRIAE_CS42_1BL_TGAC v1_033317_AA013869 0	255 aa	WM27A A9J215	mRNA: AM502894.1 (WM27A) protein: CAM59072.1		
	<i>TaMADS21-D/ TaMADS17</i>	D	TraesCS1D02G26270 0	TRIAE_CS42_1DL_TGA Cv1_063298_AA02260 00	250 aa	A0A1D5SYQ 8			

Regarding *Triticum aestivum*, three homeologs for each ortholog from *Brachypodium* were identified which presented with very high similarity (>91%).

We proceeded to phylogenetic analysis to confirm the homology of *B. distachyon* candidate genes and show the relationship between *MADS* protein orthologs from different species in order to verify the ortholog of the candidate genes before detailed expression and functional analyses.

2.4.3 *MADS-box* transcription factors protein structure and conserved domain

Amino acid sequence alignments were performed, and derived phylogenetic trees for their functional groups were constructed to confirm the orthology of the *Brachypodium* candidates.

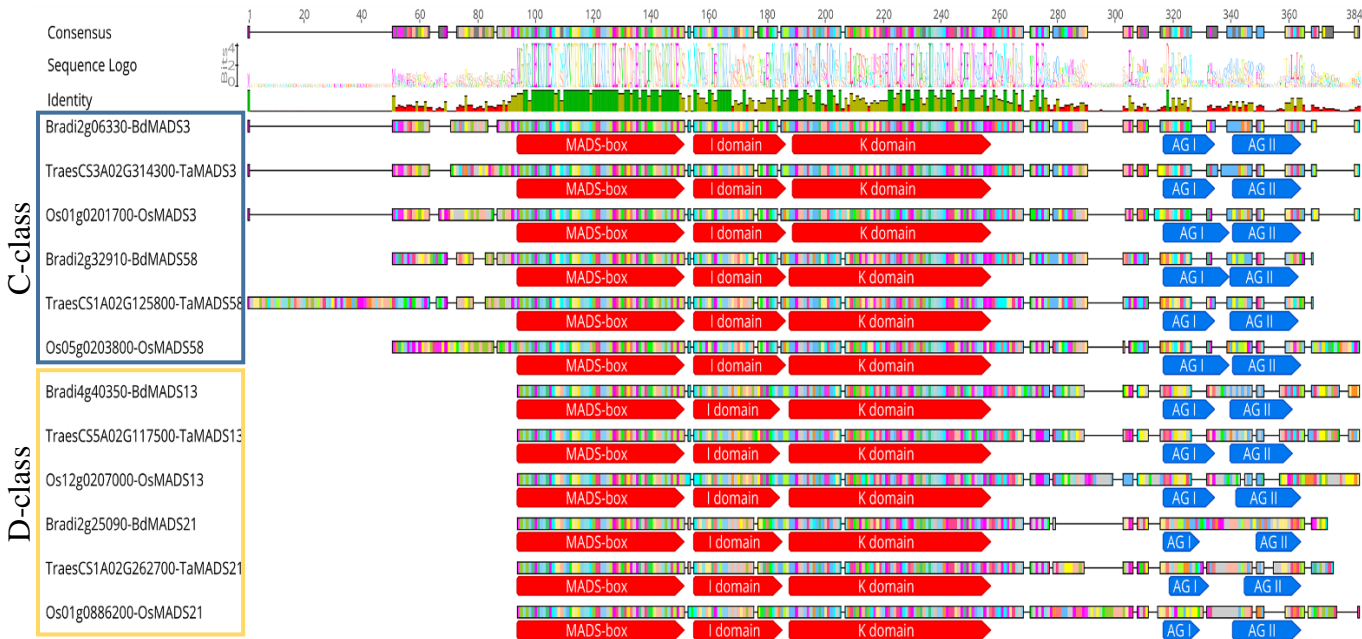


Figure 2-2: MAFFT alignment of C- and D-class *MADS* transcription factors from *Oryza sativa* and their orthologs in *B. distachyon* and *Triticum aestivum* showing the highly conserved domains *MADS-box*, I domain, K domain and two AG motifs in the C terminal.

MADS-box domain	100 110 120 130 140 150 MGRGRLEIKRIENTTIRQVTFCKRRNGLLKKAYELSVLCDAEVALTFSSRGRLEYYS NGGRLEIKRIENTTIRQVTFCKRRNGLLKKAYELSVLCDAEVALTFSSRGRLEYYS
I domain	160 170 180 NSVKATIERYKKANXDTSSNG-XVAEVNA-QY NSVKATIERYKKANXDTSSNG-XVAEVNA-QY
K domain	190 200 210 220 230 240 250 QESAKLRHQIQSLQNSN-RHLVGDSSVGTMSLXELKQLEGRLEKGLXKTRARKNELLEXETIYMQKREME QESAKLRHQIQSLQNSN-RHLVGDSSVGTMSLXELKQLEGRLEKGLXKTRARKNELLEXETIYMQKREME

Figure 2-3: The consensus sequences of *MADS-box* domain. I domain and K domain for the five species (Bd, Ta, At, Os and Hv) using MAFFT v7.308 (Kato & Standley, 2013) in Geneious R10.

A full-length sequences alignment of the transcription factors was performed and showed high homology across the *MADS-box* domain.

The alignment (Figure 2-2) revealed a very high similarity between the sequences from the three species including *MADS-box* domain in form of a helix which splits into two structure:

α -helix, involved mainly in DNA-binding, recognising CA-rich-G-box site or the so-called “Serum Response Element (SRE)-type CArG box” (Verelst, Saedler, & Munster, 2007), and two antiparallel β strands that facilitate the protein dimerization. It also contains an NLS motif KR[K/R]X₄KK that allow the TF to be actively transferred from the cell cytoplasm to the nucleus. This is followed by the Intervening domain that influences the specificity of DNA-binding protein-protein dimers and predicted to form an α -helix, the following is the Keratin-like domain involved in protein-protein interaction. After *MADS* domain, it is the most highly conserved domain of MIKC-type proteins, and it has been subdivided into three subdomains K1, K2, and K3 (Kaufmann et al., 2005). Finally, the C-terminal domain that is involved in transcriptional activation and forming higher-order multimeric *MADS* protein complexes (Gramzow & Theissen, 2010). The I domain and the C-terminal domain are less conserved than the *MADS* and K domains.

The last region of the protein sequences (C-terminal) presents two short highly conserved motifs which are *AGAMOUS* I (AGI) and *AGAMOUS* II (AGII) (Kramer et al., 2004). These motifs (Figure 2-2) primarily contain hydrophobic and polar residues and have no recognisable similarity to other known functional motifs. They do have some similarity in sequence to the conserved C-terminal sequences in many angiosperm AG-homologs (Kramer et al., 2004).

The alignment shows that the C-lineage present an N-terminal extension of around 44 residues before the *MADS* domain (Jager et al., 2003), which have not been seen in any of the D-lineage members. Previous functional study on *Arabidopsis* indicates that the N-terminal extension present in this protein sequences is not important to any main aspect of gene function (Mizukami et al., 1996).

In summary, the alignment confirms that our target genes are members of *AGAMOUS* (AG) subfamily of MIKC-type *MADS-box* genes. As well as illustrating the high similarity to their orthologs in wheat in terms of the major functional domains. It is a widely held view that the C-terminal region plays a key role in the diversity that distinguishes every single sequence regarding their specific function.

2.4.4 Phylogenetic, Identification and classification analysis of *MADS-box* genes in some monocots

To identify the putative *MADS-box* MICK-type proteins in *Brachypodium distachyon*, *Oryza sativa*, *Triticum aestivum*, *Hordeum vulgare* and *Arabidopsis thaliana*, a BioMart search using the Interpro ID (IPR033896), resulted in the identification of 326 proteins in total which has been used in the alignment and tree construction in order to cluster the families together and identify the unknown ones. The number was distributed as following; *Bd*: 40 TFs, *Os*: 43 TFs, *At*: 47 TFs, *Hv*: 41 and *Ta*: 52 genes with their copies in the three A, B and D homoeologues (155 in total) more than what has been found in (Zhao *et al.*, 2006) and (Paolacci *et al.*, 2007) together.

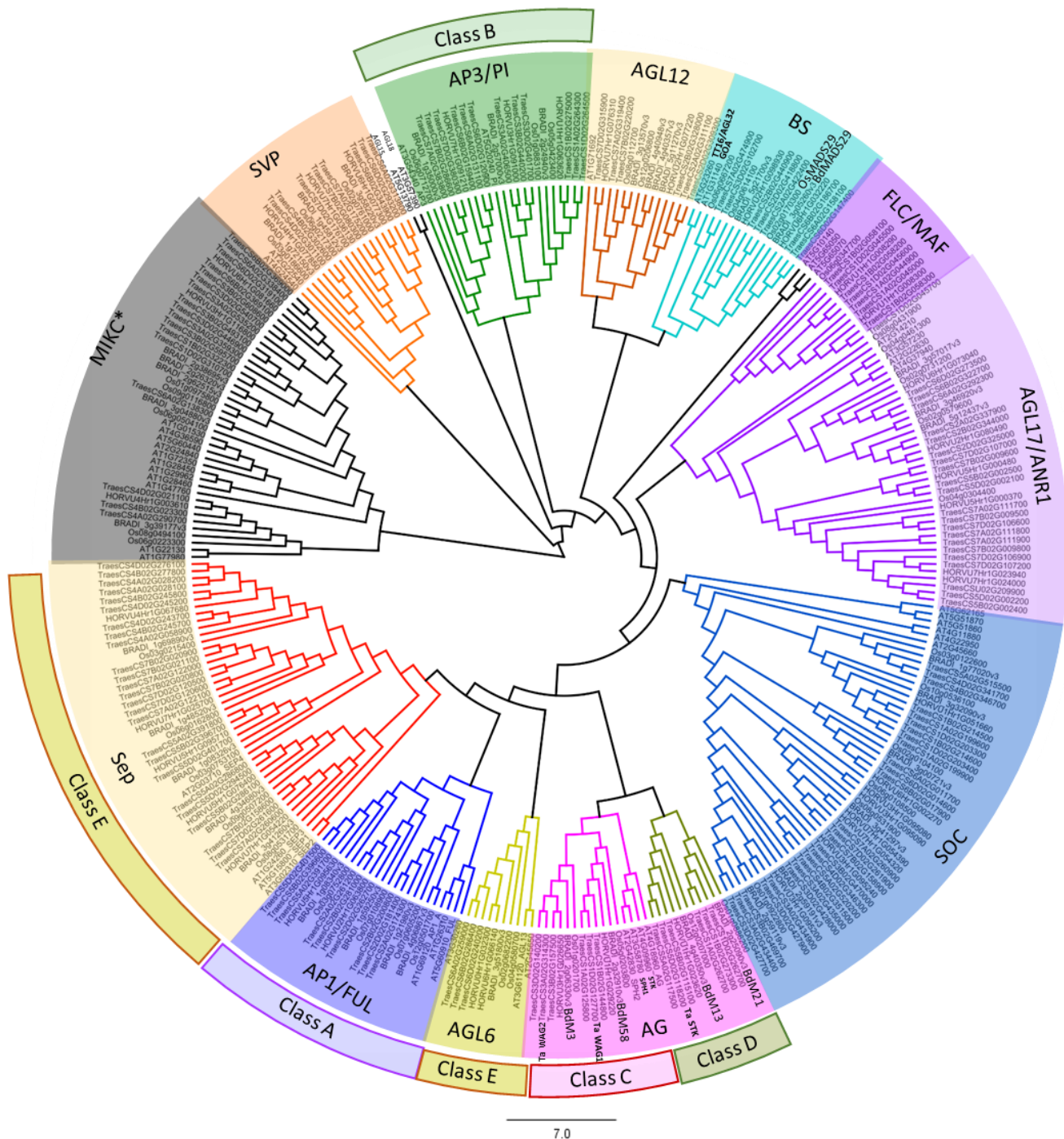


Figure 2-4: Maximum likelihood phylogenetic tree derived from the MAFFT alignment of *MADS-box* MIKC-type Protein (IPR033896) for *Triticum aestivum*, *Hordeum vulgare*, *Brachypodium distachyon*, *Oryza sativa* and *Arabidopsis thaliana* including orthologs of our target genes *BdM3*, 58, 13 and 21. Scale bar represents amino acid substitution per site.

To classify and categorise the *MADS-box* genes, a maximum likelihood phylogenetic tree of 326 genes from five species, based on full-length protein sequences was constructed to determine their phylogenetic relationships. In total, 285 genes were determined to be MIKC^C *MADS-box* genes (including; *Bd*: 35 TFs, *Os*: 38 TFs, *At*: 36 TFs, *Hv*: 38 TFs and *Ta*: 138 TFs) and 41 genes were MIKC* *MADS-box* (including; *Bd*: 5 TFs, *Os*: 5 TFs, *At*: 11 TFs, *Hv*: 3

TFs and Ta: 17 TFs). The MIKC^C genes were grouped into eleven sub-families based on their phylogenetic relationships with characterised orthologs in *Arabidopsis MADS-box* genes and on previously identified genes from *Oryza sativa*.

Since several whole-genome duplication events happened during angiosperm evolution, it is likely that the 47 genes in *Arabidopsis* is caused by a high duplication frequency, in combination with an increased retention of *MADS-box* genes that were subjected to neo-functionalisation and gained important functions in the core eudicots (Theissen *et al.*, 2000). In addition *Triticum aestivum* presents the higher number of *MADS-box* genes among the five species due to the evolution of hexaploidy as result of the hybridisation of wild and cultivated species to improve the agronomic traits.

To examine the phylogenetic relationships between *MADS-box* genes in detail, the MIKC^C was further classified into 11 clades, based on well-identified TFs from *Oryza sativa* and *Arabidopsis thaliana* (Figure 2-4). The cluster SOC consisted of the highest (16) number of *MADS* TFs, whereas clades AGL6, AGL12 and FLC/MAF had the lowest members. In the SOC clade contain similar number of genes in *Brachypodium* and rice (6 genes each), however, in barley and wheat this number rise to eight because of the duplication that occurred in *OsMADS56* (Os10g0536100) and *OsMADS50* (Os03g0122600) in wheat, where it was represented in two paralogues (Figure 2-4 blue highlight). *OsMADS37* (Os08g0531900) duplicated in two paralogs in barley as well, their ortholog in *Arabidopsis* found to function as a positive regulator for flowering (Liu *et al.*, 2008) and *OsMADS50* and *OsMADS56* function antagonistically in regulating long day dependent flowering in rice.

The Sep-like E-class seems to have expanded in only wheat with eight genes represented in this clade (Figure 2-4), while rice, *Brachypodium* and barley kept only five members each. In rice *OsMADS7* (Os08g0531700) found to be involved in stabilising rice amylose content at high temperatures and also play an essential role in grain filling (Zhang *et al.*, 2018). In this clade, we can also find the *OsMADS8* (Os09g0507200) that regulate stamen identity and tapetum development by activating bZIP transcription (Chen *et al.*, 2018), in addition to *OsMADS34* (Os03g0753100) that essential for rice inflorescence and spikelet development (Gao *et al.*, 2010). Their orthologs in *Brachypodium* and barley may have a similar function; however, for wheat, their duplication may lead to either sub-functionalisation or neo-functionalisation.

It has been anticipated that an ancient duplication after the divergence of angiosperms and gymnosperms gave rise to the C- and D-lineage *MADS-box* genes in the AG clade (Kramer, et al., 2004). Subsequently, random duplications resulted in additional paralogues in each lineage, the ovule-specific D-lineage and the well-characterised C-lineage, that generally involve in stamen and carpel identity as well as floral meristem determinacy. In the current study, we observed that there was no duplication event happened in *AGAMOUS* between *Brachypodium* and domesticated cereals (rice, barley and wheat). C-class genes were represented with two paralogues in four species (*TaWAG1-2*, *OsMADS3-58*, *BdMADS3-58* and *HvAG1-2*), and the same for D-class, however, barley appear to lost the orthologs of *MADS21* presenting only *MADS13* (Figure 2-4).

The B-sister clade appears to expand in rice and wheat, presenting three genes; the characterised and well-known *OsMADS29* (Os02g0170300) that regulates the degradation of the nucellus and the nucellar projection and activate the cysteine protease that involved in the PCD (Yin & Xue, 2012). *OsMADS29* was also found to regulate the vacuolar processing enzyme *OsVPE1* that plays a key role in the programmed cell death (PCD) process (Yang *et al.*, 2012). The second gene in the clade is *OsMADS30* (Os06g0667200) which has a function in root development during vegetative development (Schilling *et al.*, 2015) and then the third gene *OsMADS31* (Os04g0614100) that was predicted to be involved in salt tolerance at the rice germination stage - these three sub-clades split at least 45 million years ago (figure 2-4). On the other hand, *Brachypodium* and barley have only two genes, an ortholog of *OsMADS31* and another ortholog of *OsMADS29*. In this project we have investigated the expression of this ortholog in *Brachypodium* (Chapter 6).

Evolutionary analysis indicated that the rice and *Brachypodium* have approximately the same number of genes in most of the clades illustrated in the *MADS-box* MIKC-type family (table 2-3). On the other hand, barley and wheat appear to expand in the numbers of genes in most clades. This increase in numbers could be explained by the domesticated trait that characterises these cereals.

Table 2-3: Comparison of gene numbers in each *MADS-box* family in four species . There was an expansion in the SOC, AGL17 and Sep families in wheat and barley.

Families/species	<i>Triticum aestivum</i>	<i>Hordeum vulgare</i>	<i>B. distachyon</i>	<i>Oryza sativa</i>
SOC	8	8	6	6
ANR1/AGL17	7	6	3	5
Sep	8	5	5	5
FLC/MAF	3	2	0	1
AP1	3	2	4	4
B sister	3	2	2	3
AG	4	3	4	4
AP3	4	3	3	3
AGL12	2	2	3	2
SVP	3	3	3	3

2.4.5 Chromosomal locations of *MADS-box* gene family members in *Triticum aestivum*, *Oryza sativa* and *Arabidopsis thaliana*

The *MADS-box* MIKC-type gene was distributed in wheat among the three sub-genomes; A (48), B (53) and D (53) genes (Figure 2-5). The distribution reflects the presence of an ortholog of each gene in the three ancestral species of *Triticum aestivum*. This is also reflected among each of the seven homeologous.

Similarly, in *Oryza sativa*, *MADS-box* MIKC-type genes equally dispersed among the twelve chromosomes with some concentration on the regions between the centromere and the telomeres. A small variation was noticed in the distribution of *MADS-box* MIKC-type genes in *Arabidopsis*; in fact, it did not follow any remarkable pattern (Figure 2-5).

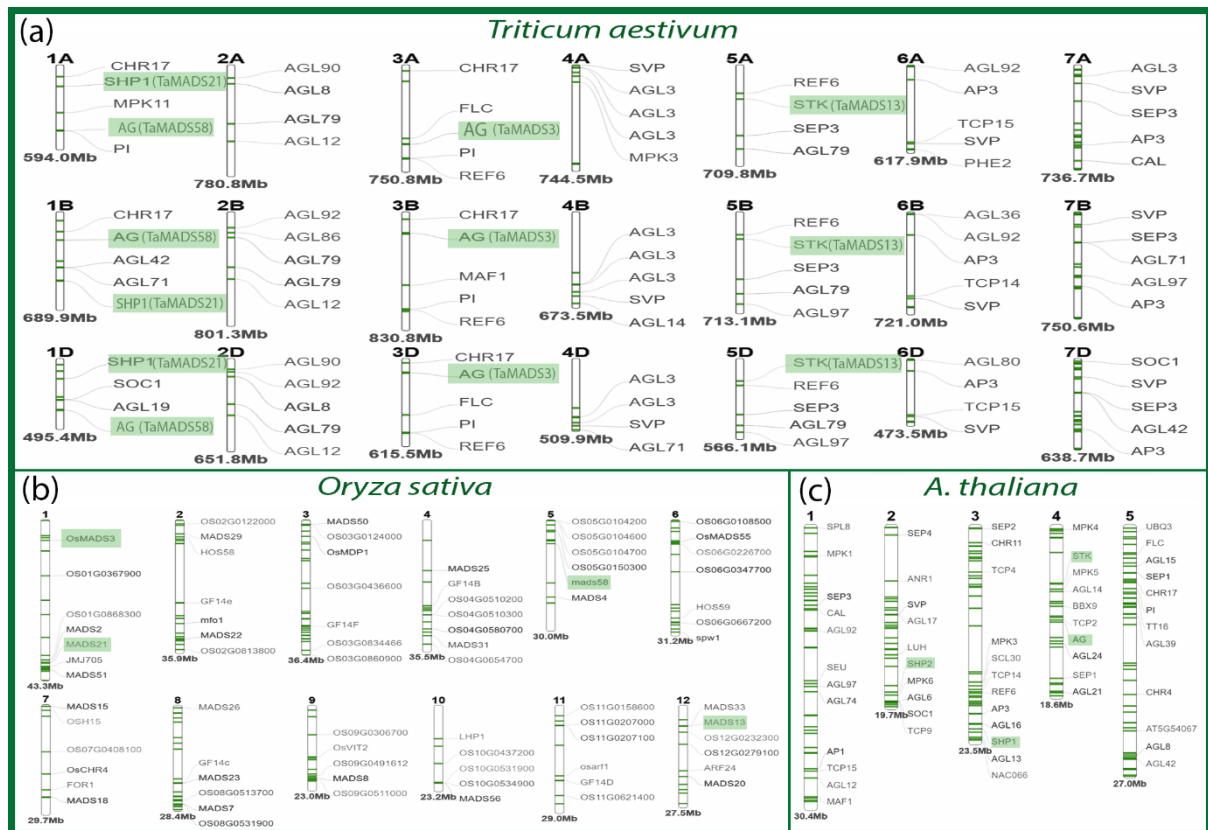


Figure 2-5: Chromosome location and distribution of *MADS-box* transcription factor protein in *Triticum aestivum*, *Oryza sativa* and *Arabidopsis thaliana* using (knetminer.rothamsted.ac.uk)

MADS-box MIKC-type genes are distributed across all chromosomes in the three species. In wheat specifically, MIKC-type genes were distributed based on their function. Genes that are responsible for the development and/or related to stress response such as, AGL17 that involve in root development, and FLC-like genes determine flowering time, have been found to be located in the highly recombinogenic distal chromosomal segments. On the other hand, genes related to highly conserve developmental function including photosynthesis, cell cycle and reproductive organ identity are enriched in proximal chromosomal segments such as AP3, AG-like genes (Ramírez-González et al., 2018).

2.4.6 Conserved motifs identification

The maximum likelihood phylogenetic tree generated, illustrates all the *MADS-box* families in five species; therefore, a further analysis was performed to identify the conserved motifs in the amino acid sequences shared among each class.

2.4.6.1 A-Class amino acid sequences alignment and motifs:

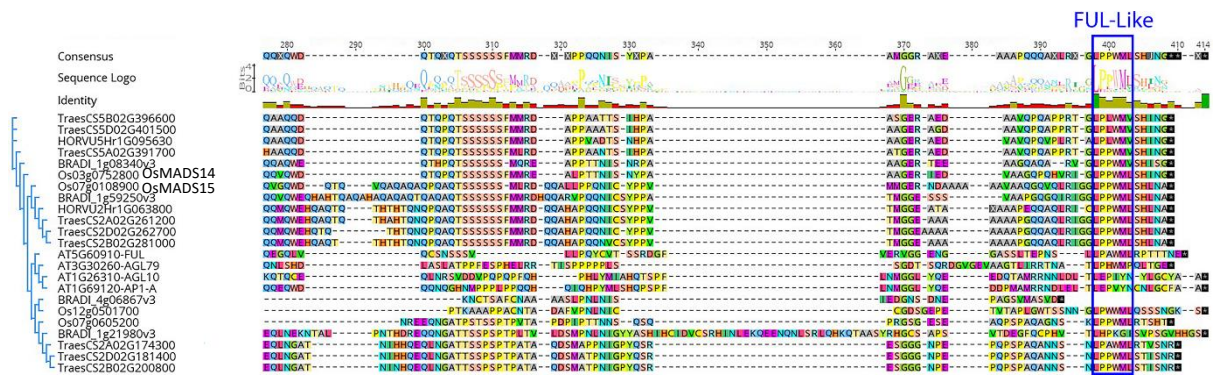


Figure 2-6: Alignment of the C-terminal regions of the A-class amino acid sequences for *Ta*, *Bd*, *Hv*, *Os* and *At* orthologs, MAFFT v7.308.

There are 23 members in this class A namely *SQUAMOSA*, containing sequences of *AP1/FUL*-like genes presented in all five species with variable numbers. Previous research on *Arabidopsis* and *Antirrhinum*, have found that the A-class genes *AP1* are responsible for floral initiation, control the identity of floral organs, and regulate the process involved in fruit maturation (Fornara *et al.*, 2004).

Their orthologs in monocots have some divergence due to the duplication events, in function, expression and sequences with some level of conservation (Zhang & Yuan, 2014). The well characterised four genes from *Arabidopsis* clustered with their orthologs in the monocot species and allowed the identification of three genes in wheat represented with their three homoeologues A, B and D. Two genes were found in *Hordeum vulgare* and four genes were found in *Brachypodium* and rice (Figure 2-6). Experimental researches were performed on the monocot model species *Oryza sativa* *MADS15* (class-A) gene and revealed that it is expressed in the apical floral meristem and increases in the palea, lemma, and lodicules after differentiation of the spikelet organs (Koyzuka *et al.*, 2000) which represent the first and the second whorl in the floral structure. Class A sequences have a highly conserved motif in the C-terminal called the FUL-like motif (LPPWML) (Litt and Irish, 2003) (Figure 2-6). Later this motif has undergone multiple duplication events followed by sequence divergence to give rise to euAP1 in core eudicot with two distinct conserved motifs: RRNa-LaLT/NLa and CFAT/A for instance, in *Brassica rapa* (BrCAL) and *Arabidopsis* (CAL), and euFUL in *Arabidopsis* (FUL) and *Antirrhinum* (AmFUL) (Litt and Irish, 2003). In the alignment Figure 2-6 only FUL-Like motif was shared among these sequences.

2.4.6.2 B-Class amino acid sequences alignment and motifs:

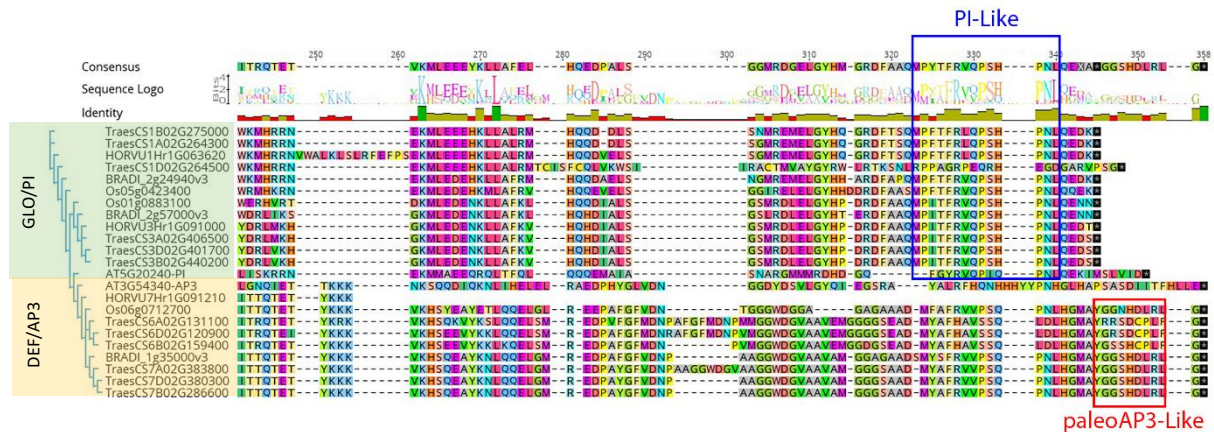


Figure 2-7: Alignment of the C-terminal regions of the B-class amino acid sequences for *Ta*, *Bd*, *Hv*, *Os* and *At*, MAFFT v7.308.

In DEFICIENS-/GLOBOSA- (class B), 23 genes were identified (*At*: 4, *Bd*: 3, *Ta*: 4, *Hv*: 3 and *Os*: 3) and clustered with the characterised PI and AP3 from *Arabidopsis*, four genes with their homoeologues A, B and D in *Triticum aestivum* (two *PI-like* genes and two *AP3-like* genes), and three genes in *Brachypodium*, barley and rice (figure 2-7). The floral homeotic genes expression of B-class results in the transformation of vegetative leaves into petals and determine the identity of stamens in eudicot model *Arabidopsis*. It has been believed that the duplication of the B-class genes led to the distinct sister clades of *AP3-like* from the gene subfamily *DEF* and *PI-like* from the gene subfamily *GLO* (Becker & Theissen, 2003).

Only one *DEF-like* gene has been found in *Bd*, *Os* and *Hv*, suggesting that no gene duplication event happened in AP3 clade, and it was highly conserved during evolution, except for wheat, where the gene duplication appears to have occurred in both clades DEFICIENS and GLOBOSA. However, in the *Bd*, *Os* and *Hv* the duplication was restricted only to the PI-like genes resulting in two orthologs in this clade.

In our target species, GLO-like proteins were characterised by a PI-motif (MPFTFRVQPSHPNL) found in the PI-like proteins C-terminal, while, DEF-like proteins were identified by a paleoAP3-motif (YGxHDLRLA) (Figure 2-7).

2.4.6.3 C-Class amino acid sequences alignment and motifs:

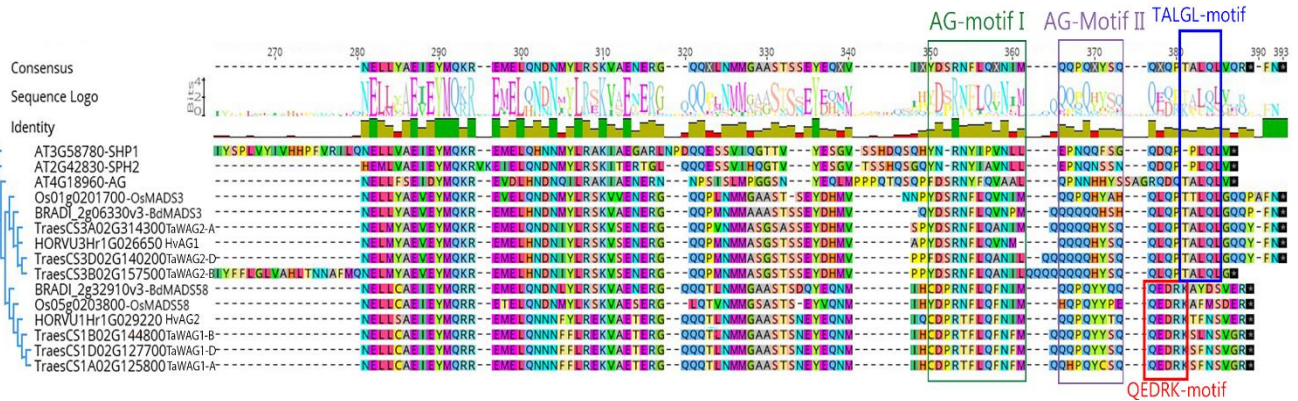


Figure 2-8: C-terminal region of the protein sequence alignment for *Ta*, *Bd*, *Hv*, *Os* and *At* belong to the AGAMOUS C-class. The rectangle indicates a conserved motif among this family, MAFFT v7.308.

The C-class branch from the AGAMOUS clade (Figure 2-8), contains two paralogues in *Brachypodium* *BdMADS3* and *BdMADS58* together with their orthologs in rice and barley. Similarly, wheat has two paralogues with their homoeologs from A, B and D genomes all clustering together with the well characterised *Arabidopsis* AG and SHP1/ SHP2. Although these genes have the same origin and conserve the two motifs AG-motif I and AG-motif II (Figure 2-8), the subsequent duplication contributed to functional diversification after the divergence of angiosperm, where C-lineage genes regulate stamen and carpel identity and control floral meristem determinacy, while D lineage genes more commonly control ovule development (Galimba & Di Stilio, 2015). Moreover, within the C-class the orthologs of *MADS3* showed a conserved motif that characterises them localised around 380 aa, called TALGL-motif, this motif was present in all four species (*Ta*, *Hv*, *Bd* and *Os*) (Figure 2-8). It has shown in previous studies on rice that these genes specify the plant reproductive organs stamens and carpel together with B-class and they are reported to suppress the expression of A-class as well (Wang *et al.*, 2015). On the other hand, the orthologs of *MADS58* across species conserved QEDRK-motif (Figure 2-8). Supporting to this theory studies on rice have shown that *OsMADS3* plays a predominant role in stamen specification, with high expression in stamen primordia, while *OsMADS58* in the floral meristem (Dreni & Kater, 2014).

Previous studies in wheat have been identified three homoeologues of WAG1 localised in the chromosomes (1A, 1B, and 1D) by Southern blot analysis (Meguro *et al.*, 2003). These genes were identified in our study as TaWAG1-A, TaWAG1-B and TaWAG1-D in the new assembly (RefSeq 1.1, 2019) (Figure 2-8) and showed high similarity to *BdMADS58* and

OsMADS58. While, (Wei *et al.*, 2011) has identified three homoeologues of WAG2 on chromosome two (2A, 2B, and 2D) that have a high similar to the *BdMADS3* and *OsMADS3*, named TaWAG2-A, TaWAG2-B and TaWAG2-D. WAG1 and WAG2 are classified as C-function genes in *Triticum aestivum* (Meguro *et al.*, 2003; Zhao *et al.*, 2006; Yamada *et al.*, 2009; Shitsukawa *et al.*, 2007; Hirabayashi & Murai, 2009; Murai, 2013; Callens *et al.*, 2018).

2.4.6.4 D-Class amino acid sequences alignment and motifs:

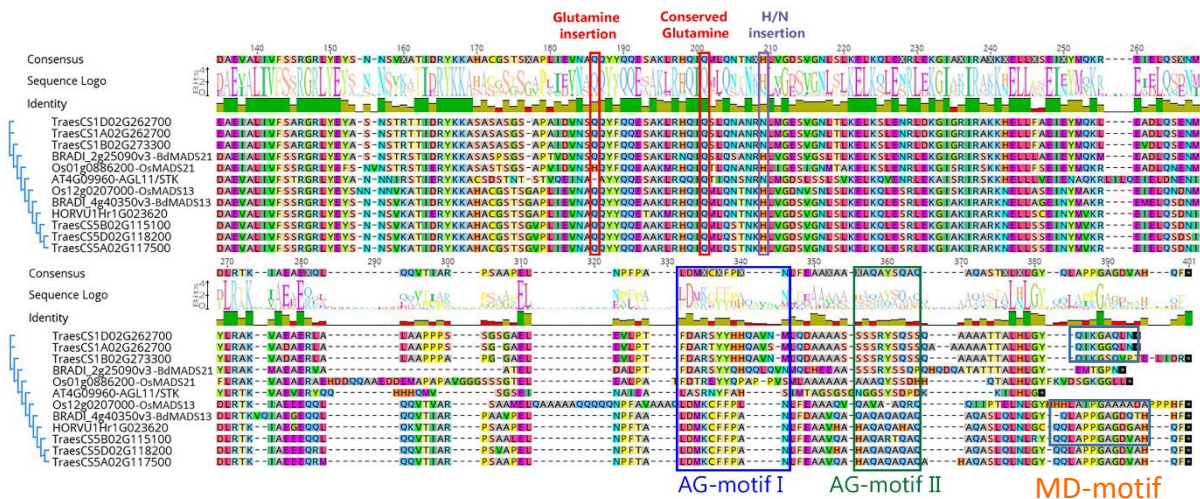


Figure 2-9: Alignment of the C-terminal regions of the D-class amino acid sequences for *Ta*, *Bd*, *Hv*, *Os* and *At*. MAFFT v7.308, scoring matrix BLOSUM62.

In the monocot D-lineage of the *AGAMOUS* clade, there is conserved glutamine at position 105 (in the alignment position 201) (Figure 2-9) in the protein sequence, which characterises the D-class and is absent in C-class (Kramer *et al.*, 2004; Dreni *et al.*, 2007). There is also a specific single glutamine amino acid insertion at position 90 (in the alignment position 186) in the protein's sequences and another single insertion of either histidine or asparagine residue at position 113 (in the alignment position 209). Both of these are absent in C-lineage proteins (Dreni *et al.*, 2007). Furthermore, at the C-terminal of the D-class proteins, there are two highly conserved motifs, AG I and AG II, among *AGAMOUS* clade (Kramer *et al.*, 2004) (Figure 2-9). Downstream of these motifs, the alignment revealed that the MD motif found in (Hsu *et al.*, 2010) study.

2.4.6.5 E-Class amino acid sequences alignment and motifs:

E-class genes belong to the *AGL2*-subfamily, was characterised by forming a high order protein complex with the class A, B, C and D in forms of floral quartets in order to specify flower organ identity (Becker & Theißen, 2003). Their function is essential in the specification of all flower whorls.

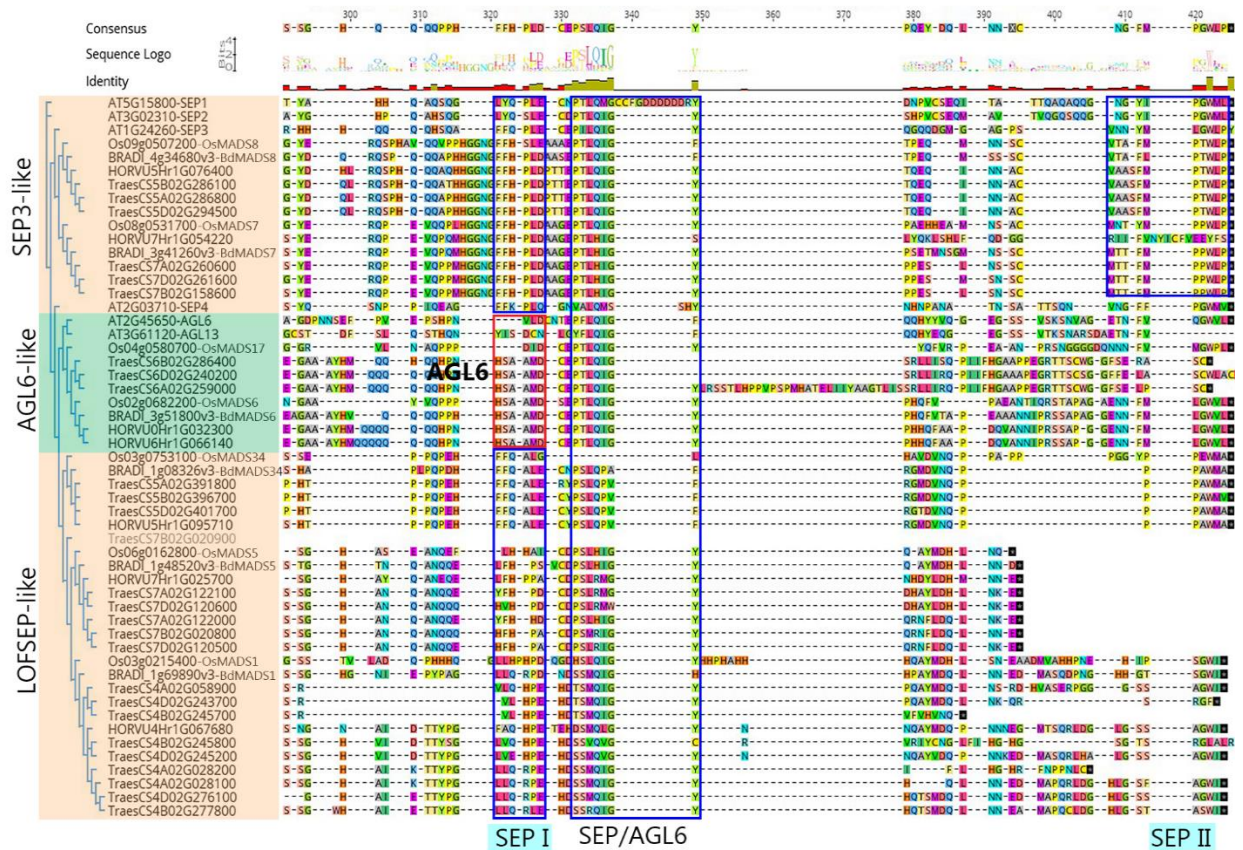


Figure 2-10: Alignment of the C-terminal regions of the E-class amino acid sequences for *Ta*, *Bd*, *Hv*, *Os* and *At*. The grey highlighted ID is a short wheat sequence of 165 amino acid; the orange box is *AGL2*-like proteins (LOFSEP and SEP3). MAFFT v7.308.

The C-terminal alignment of the E-class clade presents one conserved motif, SEP/AGL6 among all *AGL6*-like subclade and SEP-like family members (consensus: PSLQIGY) (Kanno *et al.*, 2006) (Figure 2-10). Our analysis revealed that gene duplication events divided the clade into SEP3-like and LOFSEP-like sub-families in grasses (figure 2-10). SEP3-like protein is characterised by the presence of two distinctive conserved motifs; SEP I (FFHPLD) and SEP II (NGFMPGWL) that consist of hydrophobic and polar residues (figure 2-10). The second sub-clade is the LOFSEP-like clade present only SEP I motif. This divergence could be explained by the results found in a recent study in rice where the SEP3-like genes (*OsMADS7-8*) presented restricted functions in specifying lodicule, stamens and carpel,

while the function of *LOFSEP* genes (OsMADS1-3-4-5) starts earlier, in the specification of meristem identity and determinacy of all spikelet and floral organs (Wu *et al.*, 2018).

In the AGL6 clade, there was the presence of the conserved motif SEP/AGL6 characterising the E-class and absence of the two motifs SEP I and SEP II. However, AGL6-like proteins have a highly conserved motif replacing the SEP I, which we called AGL6-like motif (HSAAMD) in this study. This motif is highly conserved among grass species, rice *Brachypodium* wheat and barley (Figure 2-10).

Our phylogenetic analysis revealed that there was an expansion in this clade in wheat. For instance, *MADS1* was represented with one copy in barley, rice, and *Brachypodium* but in wheat we found that there are three copies of *MADS1*, and their expression in Figure 2-14 showed that they were very active in the grain. This suggests that perhaps these genes gain new function or act redundantly with other copies of *MADS1* (Figure 2-14).

2.4.7 C/D-class *MADS-box* transcription factor expression in wheat

Vegetative and reproductive organs samples from different stages of flowering and grain development were collected for total RNA extraction that was converted to cDNA. The expression of our candidate genes from C and D lineage were tested in *Triticum aestivum* using RT-PCR then quantified using RT-qPCR.

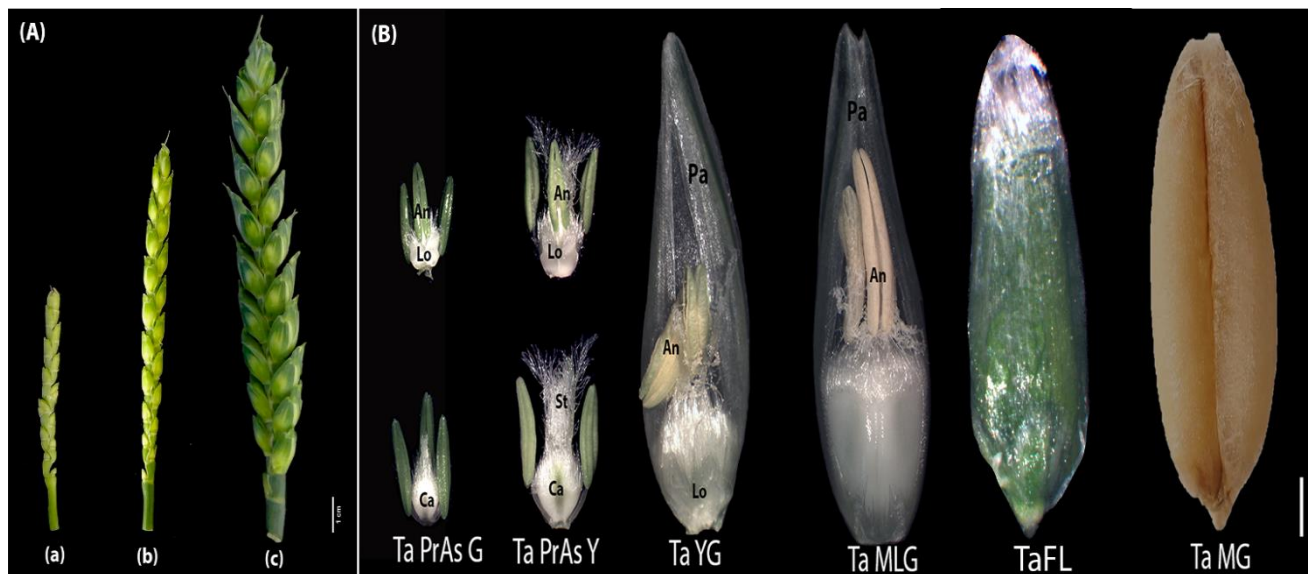


Figure 2-11: Physiological growth and development in wheat spike and grain. (A) terminal spike during floret formation and stem elongation phase (a) SPK 5 cm, (b) SPK 7 cm, (c) SPK 10 cm. (B) grain development stages, An: anther, Lo: lodicule, St: stigma, Ca: carpel, Ta PrAs G: Pre-anthesis green anther stage, Ta PrAs Y: Pre-anthesis yellow anther stage, TaYG: young grain, Ta MLG: mid-length grain, Ta MG: mature grain, scale bar, A: 1 cm, B: 1 mm

2.4.7.1 Expression pattern of C and D lineage genes using RT-PCR in wheat:

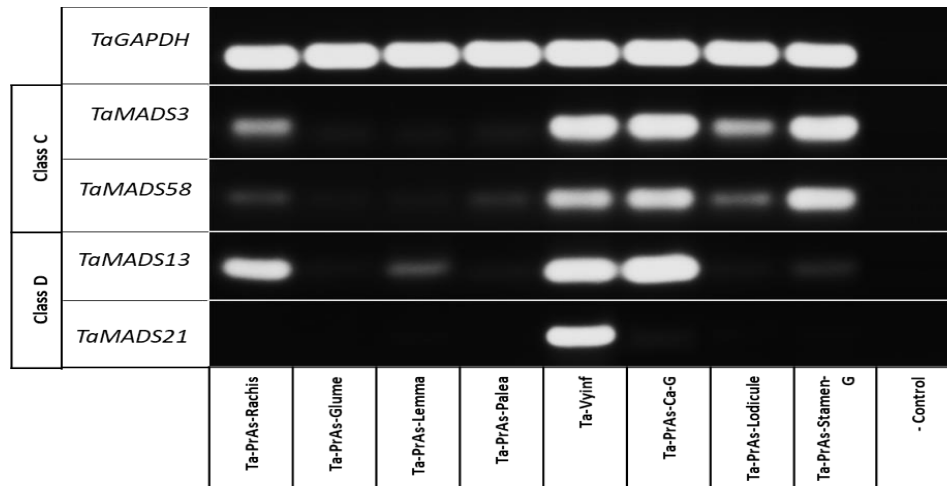


Figure 2-12: RT-PCR illustrating the expression profile of *TaMADS3*, 58, 13 and 21 in eight tissues samples .
Pr As: Pre-anthesis, VyInf: very young inflorescence, Ca: carpel, G: green.

The two paralogues of the C-class genes, *TaMADS3* and *TaMADS58*, displayed a similar pattern. There was a very low expression of the two genes in the glume, lemma and palea. However, there was a slight expression in the rachis and lodicules. The expression of the two paralogs peaks in the very young inflorescence, carpel and stamen at the pre-anthesis stage (third and fourth whorls).

On the other hand, the two D-class genes *TaMADS13* and *TaMADS21* did not illustrate a high similarity. The *TaMADS13* was expressed in the fourth whorl represented in the very young inflorescence, pre-anthesis carpel and rachis while the *TaMADS21* was expressed only in the very young inflorescence. These results showed that *TaMADS21* might have under-gone subfunctionalisation to express only in specific tissues that undergo differentiation.

2.4.7.2 Quantification of the expression of C and D lineage genes using RT-qPCR in wheat

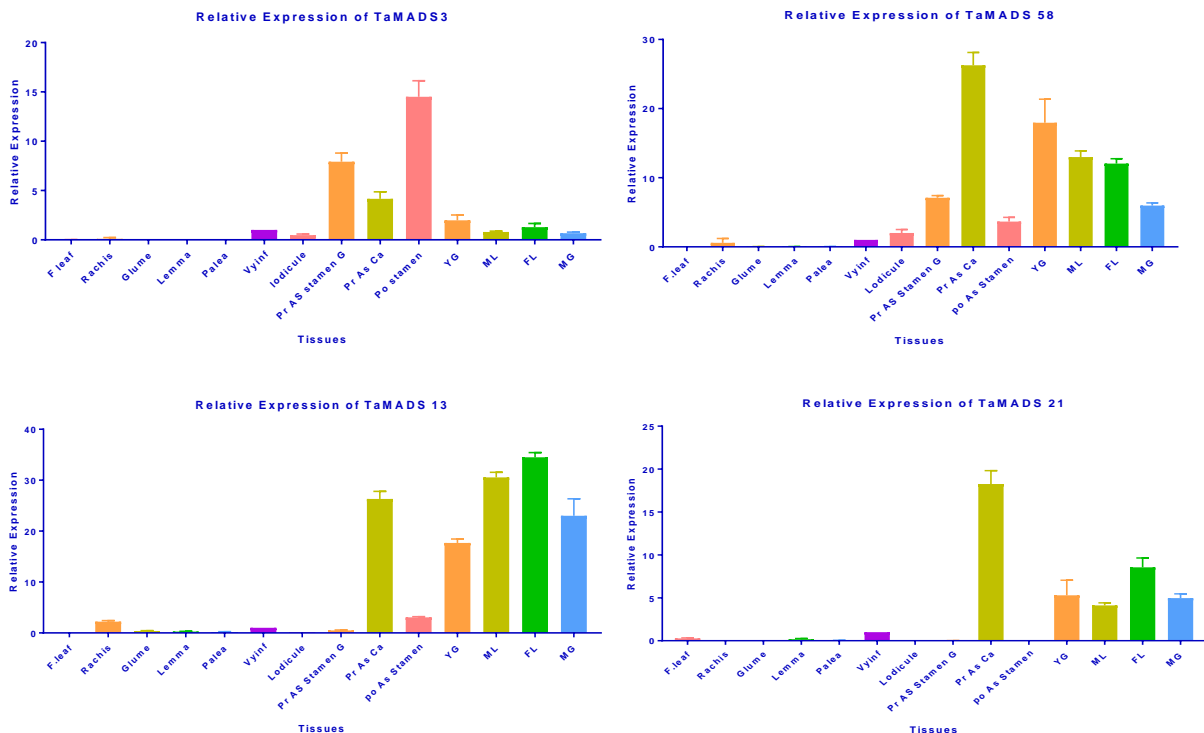


Figure 2-13: Quantification of the expression profile of *TaMADS3* (*TaWAG2-A*), *TaMADS58* (*TaWAG1-D*), *TaMADS13* (*A*), and *TaMADS21* (*A*) RT-qPCR in fourteen samples from vegetative and reproductive tissues *TaGAPDH* was used as a housekeeping gene. *VyInf* was used as calibrator. Error bars indicate \pm SD ($n = 2$). One-way analysis of variance (ANOVA) was used, there was significant difference of expression for all genes between tissue samples with an F ratio of F (13, 27) and $P < 0.0001$ with F= 109 for TaM3, F=104.85 for TaM58, F=311.51 for TaM13 and F=104.74 for TaM21.

In order to quantify the expression of the RT-PCR illustrated in Figure 2-12. RT-qPCR was performed using a wider range of samples, from vegetative and reproductive tissues. The two paralogues of C-class *TaMADS3* and *TaMADS58* (Figure 2-13) were not expressed in the vegetative organs including flag leaf, rachis glume lemma and palea. However, their expression starts with the formation of the floral organs in the very young inflorescence (*VyInf*). *TaMADS3* transcript was found abundant in the stamen (green) before anthesis with around 10-fold change compare to the very young inflorescence (*VyInf*) (Figure 2-13), then the expression peak in the stamen at anthesis stage (dehiscent anther) where pollen grains were mature and ready to pollinate the ovule. Its expression decreases but persists during grain development which suggests that they may have a role in grain filling and maturation. On the other hand *TaMADS58* gene (Figure 2-13), started also to be active in the young inflorescence and gradually increases in the lodicule pre-anthesis stamen (green), however, its highest expression was in pre-anthesis carpel with approximately 30-

fold change compared to the VyInf. Then, similar to *TaMADS3*, it continues to be expressed in the young, mid-length, full-length and mature grain with steady decrease (Figure 2-13).

The two paralogues belonging to the D-class *TaMADS13* and *TaMADS21* did not appear to be expressed in the vegetative organs and appeared to be exclusive to only the fourth whorl in the floret. *TaMADS13*, starting from pre-anthesis carpel with just more than 25-fold change compare to VyInf to decrease to 15-fold change in the young grain and increase gradually until reaching its peak in the full-length grain by approximately 35-fold change. In addition, *TaMADS21* did not show a similar pattern; the expression in the pre-anthesis carpel was higher and the expression in the young, mid-length, full-length and mature grain was relatively lower than *TaMADS13*.

There is a slight difference between the expression pattern in RT-PCR and the RT-qPCR due to the difference in primers that has been used in the experiments. However, they all highlight the presence of these genes mainly in the third and fourth whorl of the floret and also during grain development.

2.4.7.3 RNA-seq expression profile of all *MADS-box* MIKC-type transcripts in wheat

In recent years there were three generations of wheat genome assemblies project for *Chinese spring*, IWGSC with a size of 10.2 Gb in 2017 then (Clavijo *et al.*, 2017) with 13.4 Gb in 2017 and the latest IWGSC with 14.6 Gb in size in 2018 with its new annotation RefSeq 1.1 released in 2019. Along with these genome sequencing trials, there were many RNAseq studies in different organs and stages. We export expression levels of all *MADS-box* MIKC-type transcript from wheat expression web site (<http://www.wheat-expression.com>) contains more than thirty-six studies for transcriptome data for *Triticum aestivum* and Chinese spring (Table S8). The total genes were 155 including the A, B and D homoeologues, the transcripts that showed expression in the total tissue samples equal zero was removed, and therefore, the result showing in figure 2-14 illustrates only 113 genes. In order to identify their gene families, we BLAST them manually individually against the rice and *Arabidopsis* chromosomes (Figure 2-14).

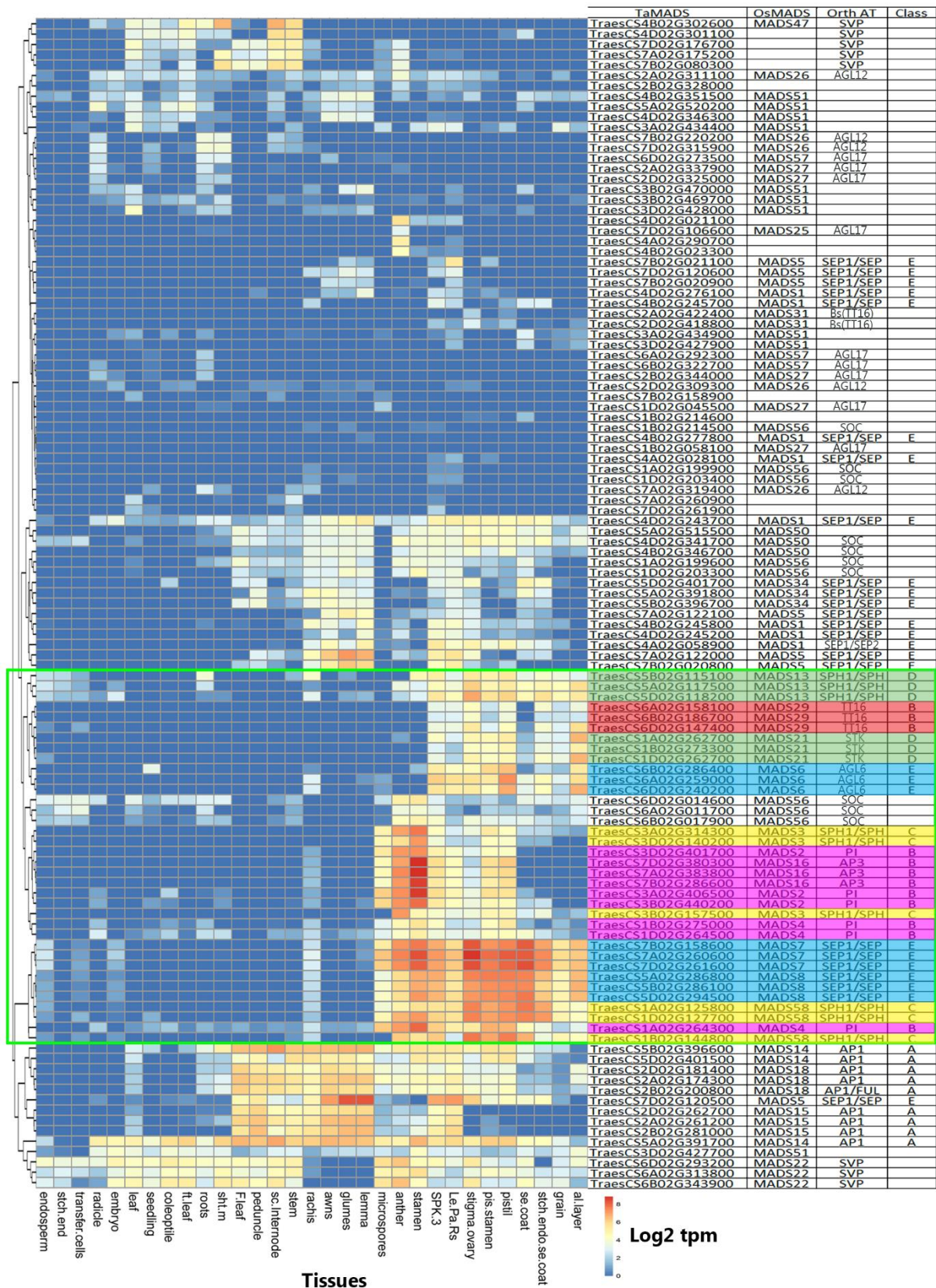


Figure 2-14: Expression Analysis of *MADS-box* MIKC-type genes transcripts in *Triticum aestivum*. Hierarchical cluster display of expression profile for 113 *MADS-box* genes showing expression in 31 different tissues samples in RNA-seq in different plant organs. Data were transformed to log₂ tpm from wheat-expression database, made in R with pheatmap package. al, aleurone, stch; starchy, se; seed, pist; pistillody, le; lemma, pa; palea, Rs; rachis, SPK3; spikelet top third, sc.int; second internode, sht.m; shoot meristem, ft leaf; first leaf, endo; endosperm, Os Orth; *Oryza* ortholog, At Orth; *Arabidopsis* ortholog.

Transcriptomic analyses from 36 studies in 31 tissues reveal that the classes of ABCDE model present some degree of clustering on the basis on their expression pattern.

C-class genes were highlighted with yellow and showed specific expression to flower and grain development (Figure 2-14). The orthologs of rice *MADS3* were more abundant in the microspores (pollen grain), anther and stamen presenting the 3rd whorl, while the orthologs of the rice *MADS58* showed a peak during grain development especially in the carpel and seed coat as well as the stamen. In addition, it can be seen very clearly in Figure 2-14 green rectangle, that the C-class genes co-express with B class genes, wheat *TaPI*-like and *TaAP3*-like genes (highlighted with pink) present almost identical expression pattern to wheat *TaMADS3*-like genes. Similarly, E-class genes (highlighted with blue), wheat *TaSEP1*-like co-expressed with wheat *MADS58*.

On the other hand, D-class genes expression cluster in a specific clade. Showing a relatively high expression in the carpel (pistil-4th whorl), endosperm and aleurone layer. These expressions coincide with the expression of members of B-class *TaMADS29*-like which know to promote nucellus PCD in rice by regulating cysteine proteases (Yin & Xue, 2012), and also its ortholog in *Arabidopsis* TT16 was shown to intervene in the nucellus degeneration (Xu *et al.*, 2016). The wheat *TaAGL6*-like gene from E-class was co-expressed with D and B-class genes, suggesting that it may have an interaction with these genes.

A-class genes expression cluster in one clade (Figure2-14), presenting a similarity in their expression pattern that covers vegetative and reproductive organs. The sub-clade that contain *TaAP1*-like appear to be a slightly abundant in the flag leaf, peduncle, stem, rachis, awns, glume and lemma, while, C-class gene where absent in these specific tissues, similar to what it was found in *Arabidopsis* where *AP1*, *AGL24* and *SVP* directly repress the expression of the B-class and C-class floral homeotic genes (Gregis *et al.*, 2009). This class (A-class) clusters with one member of E-class, *TaSEP1*-like, suggesting an interaction between the two classes.

2.4.8 C/D-class *MADS-box* transcription factor expression in barley

In order to investigate the expression of our candidate genes in barley, samples were collected from different stages, including vegetative and reproductive tissues and grain development.

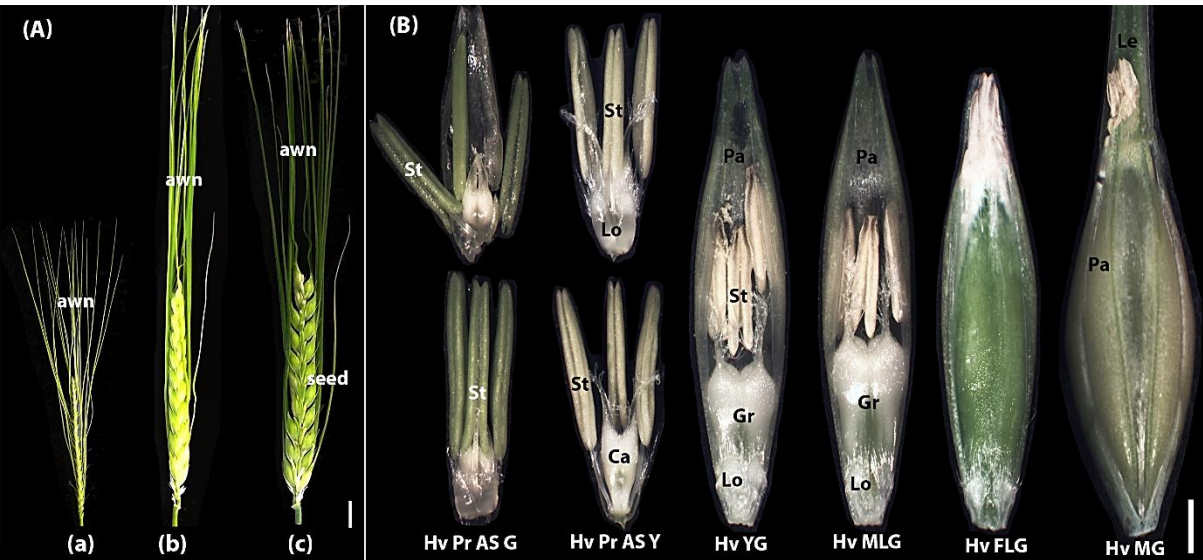


Figure 2-15: *Hordeum vulgare* developmental stages that samples were collected from. (A) Spike development, (B) grain development. HvPrAs G: Pre-anthesis, G: green, Y: yellow, HvYG: 2 DAA young grain, HvMLG: 6 DAA mid-length grain, HvFLG: full-length grain 10 DAA, HvMG: mature grain 20 DAA, DAA: days after anthesis, Scale bar A: 1 cm, B: 1 mm.

2.4.8.1 Expression pattern of C and D lineage genes RT-PCR in barley

After extraction the total RNA, it was converted to cDNA and tested for the presence of our candidate genes *HvMADS3*, 58, 13 and 21 mRNA using RT-PCR with appropriate primers.

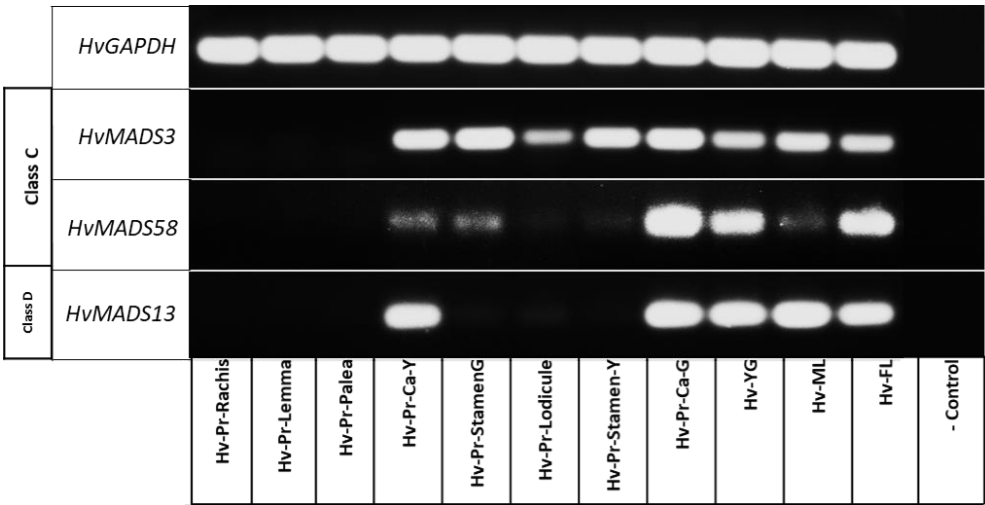


Figure 2-16: Expression pattern of *HvMADS3* (*HvAG1*), *HvMADS58* (*HvAG2*) and *HvMADS13* in barley using RT-PCR of in tissues in pre and post-anthesis stages, including, rachis, Lemma, palea, pre-anthesis carpel when the stamen was yellow, pre-anthesis stamen green pre-anthesis lodicule pre-anthesis stamen yellow, pre-anthesis carpel when the stamen was green, young grain (2 DAA) (YG), mid-length grain (6 DAA) (ML) and full-length grain (10 DAA) (FL). *HvGAPDH* was used as a housekeeping gene.

The figure (2-16) illustrate the expression of the two paralogs *HvMADS3* (*HvAG1*) and *HvMADS58* (*HvAG2*). The transcripts of these genes were absent in the vegetative tissues rachis, lemma and palea and very abundant in the reproductive tissues. *HvMADS3* was observed in whorl 2 (lodicule), whorl 3 (stamen) all stages and whorl 4 (carpel and ovule)

with slight variation, however, *HvMADS58* was restricted to only whorl 3 and 4 represented by stamen and carpel (Figure 2-16). These results demonstrate a clear difference in the expression pattern in these paralogs.

The only member of the D-class *HvMADS13* showed a restricted expression pattern to the fourth whorl starting prior and after pollination with transcript abundance in the carpel. This expression persists during young grain and grain filling; mid-length and full-length grain (Figure 2-16).

2.4.8.2 Quantification of the expression of C and D lineage genes using RT-qPCR in barley

The expression pattern observed in the RT-PCR was quantified using RT-qPCR in a wide range of tissue samples.

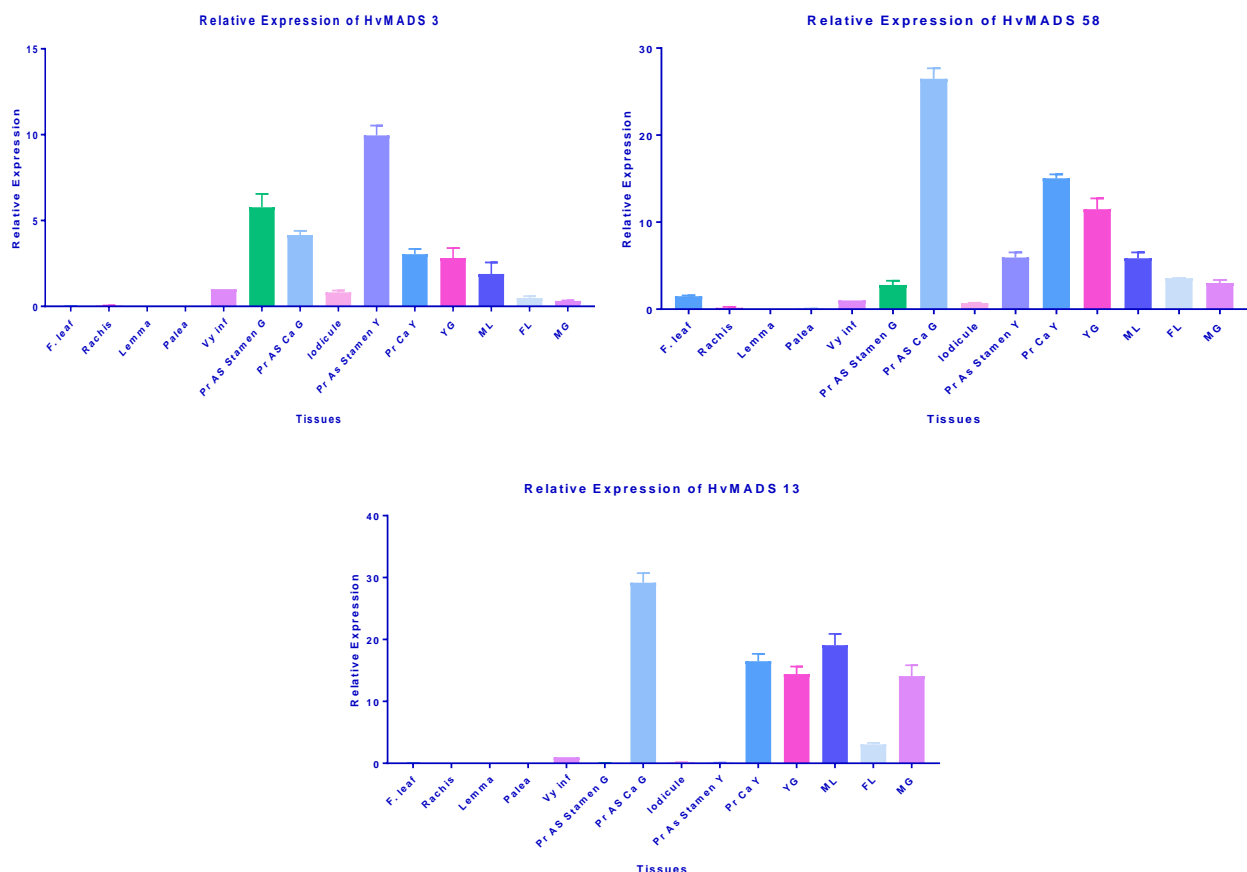


Figure 2-17: Quantification of the expression profile of *HvMADS3* (*HvAG1*), *HvMADS58* (*HvAG2*) and *HvMADS13*, using RT-qPCR in fourteen samples from vegetative and reproductive tissues in barley. *HvGAPDH* was used as a housekeeping gene. Vyinf was used as calibrator. Error bars indicate \pm SD (n = 2). Pr As Ca: Pre-anthesis carpel, Y: yellow, YG: young grain, ML: mid-length grain, FL: full-length grain, MG: mature grain, F: flag, VyInf: very young inflorescence, G:green.

The expression of the four candidate genes was quantified using RT-qPCR in numerous tissues from vegetative to reproductive stages (Figure 2-17). *HvMADS3* (*HvAG1*) starts been active at differentiation of the floral organs in the very young inflorescence, this expression increases in the green stamen and carpel at pre-anthesis, then peaks in the mature yellow stamen just prior to the pollination. The expression persists in the carpel; however, it gradually decreases during grain development, reaching its lowest level at maturity. These results suggest that *HvMADS3* could have a role either in the pollen grain formation or anther wall structure (Figure 2-17). Nevertheless, its paralog *HvMADS58* did not follow the same expression pattern, and the transcript was very abundant at nearly 30 fold-change compared to the very young inflorescence in the carpel prior to anthesis and decreases gradually during grain formation. There was some expression in the green stamen which increases slightly when the yellow stamen matured, however, this expression is just 5 fold change. These could be interpreted by the theory of functional divergence of this paralog where its functions become focused more on the fourth whorl.

The only member of the D-class (*HvMADS13*) showed expression pattern dominating only the fourth whorl prior and after the pollination, the gene stays active during grain development with a considerable 17 fold change, indicating its possible role in grain filling.

2.4.8.3 RNA-seq expression profile of the available *MADS-box* MIKC-type transcript in barley

For barley (*Hordeum vulgare* L.), there was an RNA-seq study that shows the expression of some *MADS-box* genes expression in eight tissues samples (International Barley Genome Sequencing Consortium, 2012). After selection all *MADS-box* MIKC-type IDs from *Hordeum vulgare* (41 genes), a heat-map was generated to illustrate the expression profile of these transcription factors. Unfortunately, only 26 genes were found in this study.

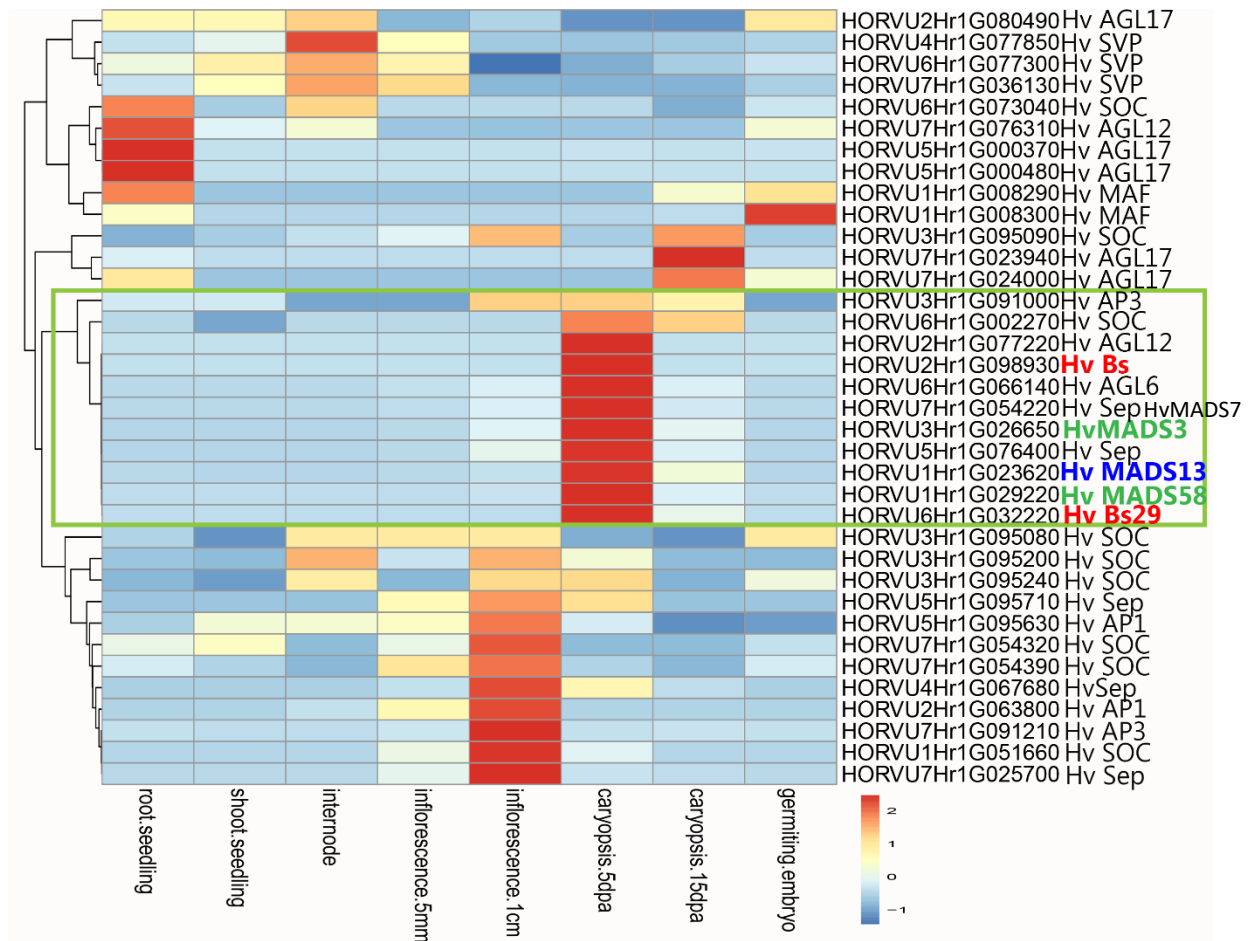


Figure 2-18: Expression Analysis of *MADS-box* MIKC-type genes in *Hordeum vulgare*. Hierarchical cluster display of expression profile for 26 *MADS-box* genes showing expression in eight different tissues samples throughout the life cycle of barley cultivar. Data was extracted from deep RNA sequencing (RNA-seq) (International Barley Genome Sequencing Consortium, 2012) then transformed to Log2. Unit variance scaling is applied to rows. DPA: days post-anthesis. Both rows and columns (not shown) are clustered using correlation distance and average linkage. 26 rows, 7 columns. The unit is FPKM (fragments per kilobase million).

In previous research, only seven *MADS-box* gene were identified (Schmitz *et al.*, 2000). Our study successfully identifies the expression of 26 genes in seven tissues in various stages of the life cycle of barley. The green rectangle (figure 2-18) illustrate the co-expression of Class-C (HvAG1 and HvAG2), class-D (HvMAD13-like) and class-E (HvSep, AGL6). The genes appear to have the same pattern of expression across the samples tested, they express slightly in the inflorescence 1cm, then peak in the carpel five days after anthesis to decrease in carpel 15 DPA. These suggest that the interaction of these genes result in some cell differentiation at the early stage of inflorescence and floral organ development as they expressed highly after anthesis.

Interestingly this co-expression included the two paralogues of the characterised B-sister *OsMADS29* that shown to activate *OsVPE1* (vacuolar processing enzyme) that play a key

role in the programmed cell death (PCD) during grain development (Yang *et al.*, 2012). It was shown also that *OsMADS29* is a direct regulator of *Ospap31* (Cysteine protease) (Os02g0715000), from papain family C1, involved in PCD in rice (Yin & Xue, 2012), which is close ortholog to *Hvpap23* (HORVU1Hr1G003940) in barley. A similar observation was seen in *Brachypodium* transcriptome analysis (Chapter 3, Figure 3-11). These results could be interpreted in two ways either that C- and D-class intervene in the PCD process or that their expression just co-occurs in the same organ as we do not have a wide range of samples from numerous studies. In addition, this clade includes members of E-class, AGL6-like and Sep-like and the same observation was noticed in *Arabidopsis* in the floral quartet model where the interaction between members of C-class (AG), D-class (SHP, STK) and E-class (SEP) contribute to the ovule identity in the ABCDE model (Theißen *et al.*, 2016).

2.5 Discussion

For over a decade, many attempts have been made in order to predict the function of some *MADS-box* genes based on their sequence similarity to the model system *Arabidopsis*, however, the expansion of some orthologs in different species made it challenging to predict the function. In this study, we have presented a comprehensive expression profiling for members of C and D-class *MADS-box* genes in wheat and barley along with an account of phylogenetic relationships analysis with rice and *Arabidopsis* of all *MADS-box* MIKC-type genes in five species.

The analysis was successfully identified 326 *MADS-box* MIKC-type proteins in five different species. The phylogenetic tree illustrates relative expansion in some families in barley and wheat compared to *Brachypodium* and rice for instance in SOC family there were six genes in *Brachypodium* and rice and eight copies in barley and wheat (Figure 2-4, table 2-3). That could be explained by the fact that wheat and barley shared a common ancestor with *Brachypodium* and rice approximately 30 to 40 million years ago and during the hybridisation and domestication these crops develop more copies of the ancestral genes that may contribute their economic traits. One example was *MADS1* that was represented by one copy in rice and *Brachypodium*, there was a duplication event that happened shortly after 4 million years that rose new copies of *MADS1* (Figure 2-14). In order to know if these copies came from domestication or duplication, we analysed the *Aegilos tauschii* that is a

wild ancestor of wheat and found that it contains these duplicated copies as well. Further studies are needed to confirm this supposition.

However, some families almost conserved their copy number across the five species, including *AGAMOUS* and B-class (AP3-like). In *AGAMOUS*, barley appears to lose the ortholog of *MADS21* and retain *HvMADS13* as the only member in the D-class.

The expression profile and its quantification, in addition to the RNA-seq, provide a coherent expression pattern in the C and D-class genes. Previous studies (Yamada *et al.*, 2009) showed that WAG1 (*TaWAG1*) and WAG2 (*TaWAG2*) are detected in the stamen, carpel, and ovule. Our investigation illustrated a comprehensive study of this expression in a wide range of samples and revealed that there was a difference in the expression pattern of the two paralogs: *TaMADS3* (*TaWAG2*) showed high expression in the pre and post-anthesis stamen while the *TaMADS58* (*TaWAG1*) was more abundant in the pre-anthesis carpel and grain. Similarly, barley presented the same pattern except that it lacks *MADS21* ortholog. These results suggest functional divergence that occurred differentiating the two paralogs *Hv/TaMADS3* and *Hv/TaMADS58*.

Brachypodium distachyon presented similar number of genes to wheat and barley (Figure 2-4) with two copies in C class (*BdMADS3* and *BdMADS58*) and two genes in the D class (*BdMADS13* and *BdMADS21*). Due to the close position to these crops in the phylogenetic tree (Figure 2-4) it is expected to have similar expression pattern throughout the plant development.

In general, wheat and barley showed high conservation of all major sub-families in *MADS-box* MIKC-type transcription factors. This conservation is played to the expression (Figure 2-13, 2-17) that could reflect the function as well. In these crops the sub-family AGL17, SOC, FLC and SEP were expanded due to gene duplication that resulted in new copies which have different expression pattern in some genes (Figure S14).

The expression data analysis and the list of IDs generated in five species are expected to help in the selection of appropriate candidate genes for further functional characterisation.

2.6 Conclusion

Wheat and barley were largely expanded in the AGL17 (ANR1), FLC (MAF), SOC and SEP-like with additional paralogs in these families. Some of the new genes are expression in

different organs than their ancestors suggestion that this genes duplication resulted in neofunctionalisation. Although barley lacks *MADS21* ortholog, both species wheat and barley showed a similar pattern of the C and D-class with an evident divergence of *MADS58* which showed high expression in the fourth whorl, unlike its paralog *MADS3*. The mRNA of our candidate's genes were initiated in the very young inflorescence during floral organ primordia then accumulates with high activity in the floral organs, suggesting that these genes are involved in floral organ development and differentiation.

Chapter 3

Investigating the role of *AGAMOUS* genes during grain development in *Brachypodium distachyon*

Chapter 3. Investigating the role of *AGAMOUS* genes during grain development in *Brachypodium distachyon*

3.1 Abstract

MADS-box transcription factors (TFs) are central to the ABCDE model for flower development where C-lineage genes are involved in stamen and carpel specification, while D-lineage genes play essential roles in ovule specification and development. In our project, we are studying two members within each of the C and D classes (*BdMADS3*, *58*, *13* and *21*) from *Brachypodium distachyon*. *B. distachyon* is an annual grass that has been established as a reference model to study temperate grasses and is closely related to economic valuable crop species, including, wheat and barley. The main goal of this chapter is to investigate the gene expression patterns (temporal and spatial) throughout development and the phylogeny of *BdMADS3*, *58*, *13* and *21* transcription factors in *Brachypodium distachyon* in comparison to other cultivated cereals.

Our work provides a comprehensive expression analysis of C and D lineage genes during flower and grain development in temperate grasses/cereals and reveals points of similarity and difference between the two lineages (C and D-class). Our study also suggests that the variation between the two members of the family could suggest functional divergence within the *Poaceae*.

In addition, here we revealed that the C-class and D-class expression is beyond the flowering stage in different layers of grain development. Therefore, our results suggest that these genes could play a significant role in defining the final size, shape and the yield in cereals.

3.2 Introduction

Domestication of cereals, like many other staple plants, has been linked to the evolution of human civilisation, and as a source of food for at least last 3000 years, the oldest example dating back to 1200 BC. It is estimated that wheat is the third largest staple food crop and a key ingredient in cereal production around the world (15% of the world's calorie intake), after maize and rice (<http://www.worldatlas.com>) in 2019. In addition, wheat accounts for more than 149 million tonnes in the European Union in 2019 and the demand is expected to rise in the next years (<http://faostat.fao.org>). This crop produced as core constituent in a range of products for different markets. Adding to the fact that the total annual turnover from cereals in the UK in 2018 amounted to approximately £3.2 billion (<http://www.statista.com>) and symbolised as a significant segment of the British economy.

It is widely believed that global food resources are under pressure because of the rapidly growing population (predicted to be 9.1 billion by 2050). The resultant impact of population growth on food supply is probably going to drive industries around the world to produce more food than they ever have in modern human history. The demand for resource-intensive products like wheat, rice, maize and barley is likely to increase dramatically. Hence, it is fair to envisage that in not too distant future, the food supply and security will face unprecedented challenges the most pertinent being sustainability and impact of population growth. Likewise, there is ever-increasing anticipation from scientific research to address some of the issues related to crop diseases and pathogens that are becoming alarmingly resistant to pesticides due to their unregulated use or misuse.

When it comes to studying the genetic makeup of wheat, it presents several complications as well as challenges as an experimental system. To elaborate further, the polyploidy and genome size of wheat species and its constituent number of paired nucleotides totals to a large number, to be exact, 17,000,000,000 base pairs (17 GB). This is approximately more than five times the amount of DNA in the human genome, which is 3 billion bp and 40 times the size of the rice genome.

Moreover, wheat-bread is classified as a allohexaploid genome with the complete set numbering 42 chromosomes, presented in three homoeologues ($2n = 6 \times = 42$, AABBDD) which is a result of hybridisation of different ancestral wheat species including, *Einkorn wheat* AA, *Triticum searsii* BB and *Aegilops tauschii* DD. As a consequence, hexaploid wheat has three distinct and closely related subgenomes in one plant.

Brachypodium distachyon was suggested as a reference species for the Pooideae family in 2001 (Draper *et al.*, 2001), and currently considered as an essential tool for dissecting numerous biological aspects of grasses due to the expanding availability of genetic resources including a fully sequenced genome and an extensive collection of TILLING lines and T-DNA mutants (Vogel & Hill, 2008). *Brachypodium* offers a convenient reference for studying cereal grain development because its mature size is less than a third of the size of cereal crops such as wheat, maize and rice, and therefore it is more amenable to phytotron and greenhouse research (Chochois *et al.*, 2012). It provides a valuable resource for studies on economically significant members of the world's major temperate zone cereals, biofuel and forages crop species such as switchgrass and ryegrass.

3.2.1 *Brachypodium distachyon* as a Reference:

The *Pooideae* subfamily forms a prominent and important sub-group of the *Poaceae*, which have particular importance in food security and applications as biofuel substrates. This sub-group has, over time, diversified in the temperate climate and contains the *Poeae*, *Aveneae*, *Triticeae*, and *Bromeae* tribes as further delineations. The sub-groups represent the cereals valuable to the UK economy; such as wheat, barley, oat, and rye. As well as important forage grasses such as *Lolium* (Hubbard, 1954).

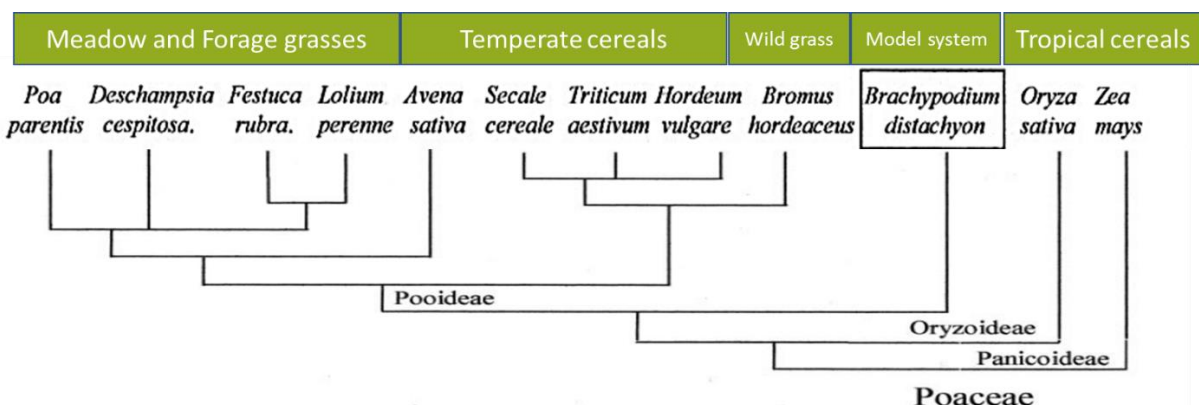


Figure 3-1: Schematic phylogenetic relationship of *B. distachyon* to other *Poaceae* adapted from the data presented by (Catalán *et al.*, 1995; Catalán *et al.*, 1997; Catalán & Olmstead, 2000).

Previous researches focussed on rice, maize, and sorghum; in terms of genome sequencing and functional analysis with the aim of improving quality of domesticated cereals quality, however, these species cannot reflect the core *Pooideae* group as they are situated in a distant clade, which presents many differences in their phenotype, genotype and ecology. The phylogenetic position of *B. distachyon* as an intermediate between rice and the *Triticeae* (Kellogg, 2001) made it a candidate for consideration as a sister group to this important section of the core *pooideae* (*Bromeae*, *Aveneae*, *Triticeae*, and *Poeae*). Therefore, it has been used as a relevant reference system for temperate cereals (Draper *et al.*, 2001) with the aim of investigating and studying cereal quality and development (Paterson *et al.*, 2009; Schnable *et al.*, 2009; Vogel *et al.*, 2010). In addition, it should also provide insights into the effects of cultivation and domestication as this is the first wild grass to be sequenced.

This self-pollinating annual grass (*B. distachyon*) provides numerous advantages that make it an excellent model system for studying cereals. It presents the practicality and convenience to work with on a small-scale lab environment, including, short life cycle of 10-18 weeks compared with 12-40 weeks wheat, small physical size of 30 cm, simple growth requirement, relatively small genome size ($(2n = 10) \approx 355$ Mbp) which present 2% of wheat genome and availability of transformation. All these benefits allow us to analyse and assess its features compared with neighbouring taxonomic groups, both wild and cultivated.

One of the most important steps in the life cycle of plants is reproductive organ development. It has been shown in studies using eudicot species like *Arabidopsis thaliana* and *Antirrhinum majus*, that MADS-box transcription factors belonging to the *AGAMOUS* (AG) subfamily regulate the identity of floral organs and floral meristem determinacy and its orthologues in many species play an essential role the floral organogenesis.

In this study, we also highlighted the role of these genes during grain development as it has an economical importance in the value of the crops, especially the grain filling and maturation stages that determine the yield and control the nutrient amount in the grain.

3.2.2 ABCDE Model of Flower Development

In the last 20 years, intensive studies have been carried out in *Arabidopsis thaliana*; the model dicot species, to understand and investigate floral organ development. The results of these studies have led to the establishment of a general model for the development of floral organs in higher plants, which is called, the A, B, C, D and E model.

This model suggests that five classes of homeotic genes identify flower development (Rijkema *et al.*, 2010). According to the floral quartet models of specification floral organ (Smaczniak *et al.*, 2012), the A- and E-class protein complex are specifically involved in the development of the sepals in the first floral whorl, the A-, B- and E-class protein complex identify petals in the second whorl, the B-, C- and E-class protein complex are required for stamens development in the third whorl, and the C- and E-class protein complex specify gynoecium in the fourth whorl. The D-class proteins interact in a larger complex with the E-class proteins to specify ovule identity (Figure 3-2).

In rice, similar schematic representation was made by (Li *et al.*, 2011) illustration the genes involved in specifying each floral organ. Additional studies revealed that duplication and subsequent sub-functionalisation is believed to have occurred, resulting in paralogues in some classes including, MADS3 and MADS58 for C-class (Yamaguchi *et al.*, 2006) and MADS13 and MADS21 for D-class controlling the 3rd and 4th whorls. This study will be focussing on these families in particular (C-D class).

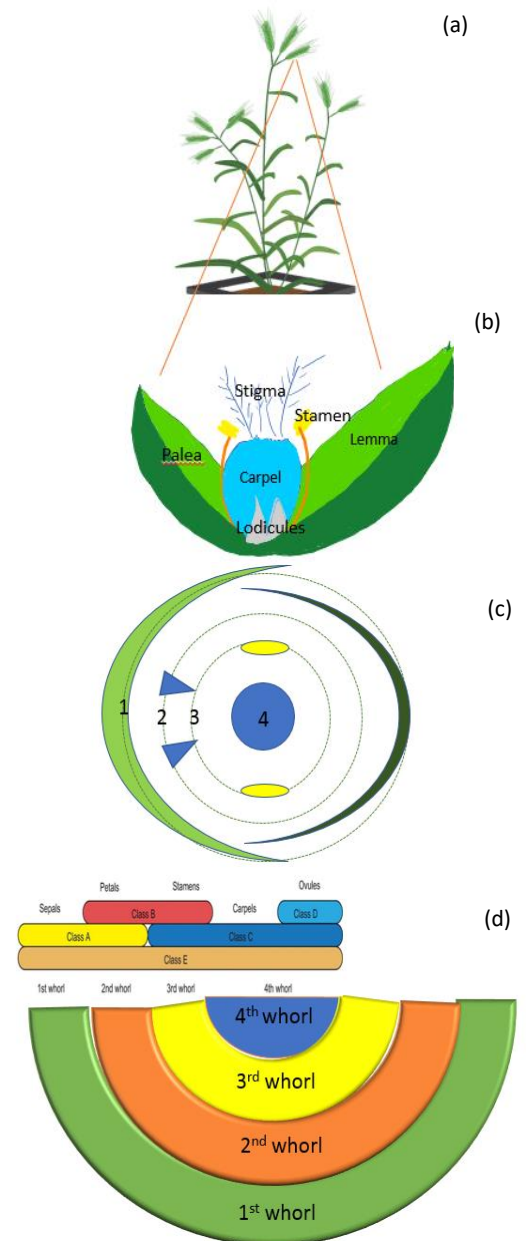


Figure 3-2: Schematic representation of (a) *B. distachyon* plant, (b) *Bd* Floret, (c) *Bd* whorls organisation of the floret and (d) the A, B, C, D and E model in *Arabidopsis* (Theissen, Melzer, & Rümpler, 2016).

3.2.2.1 C and D-Class *MADS-box* genes functions

AGAMOUS (*AG*) in *Arabidopsis* split into two sub-clades euAG with only one gene *AG* and *PLENA* (*PLE*) with two genes *SHP2* and *SHP2*. *AGAMOUS* (*AG*) encodes a C-class *MADS-box* transcription factor, which is responsible for meristem determinacy and carpel identity (Barton, 2010; Sun & Ito, 2010). Previous research has documented that the *ag* mutant in *Arabidopsis* showed an accumulation of *WUSCHEL* (*WUS*) gene in floral meristem and remains to the late stage of flower development, and concludes that in the meristem formation stage of the wild type, *WUS* in combination with *LFY* induces *AG* expression. Then after carpel formation, *AG*, together with other factors represses *WUS*, terminates stem cell fate and allows carpel differentiation (Lenhard *et al.*, 2001; Lohmann *et al.*, 2001). Moreover, similar studies highlighted that the *ag* mutant gave rise to undetermined flowers with repetitive organs in the 2nd and 3rd whorl (Bowman *et al.*, 1989; Yanofsky *et al.*, 1990). Therefore, in the floral meristem (FM), *AG* regulates meristem determinacy by repressing *WUSCHEL* (*WUS*) (stem meristem gene) in two different pathways, either by direct binding to the promoter of the *WUS* during stage 4 of flowering and resulting in immediate repression. Alternatively, by using an intermediate repressor *KNUCKLES* (*KNU*) gene that encodes transcription factors which suppress *WUS* in relatively delayed process in stage 6 (Liu *et al.*, 2011). Furthermore, in previous studies in *Arabidopsis*, it was shown that homeotic gene *AP2*, is a particular negative regulator of *AG*. It expresses in the first and the second whorl to repress the formation of carpel-like and stamen-like organs (Drews *et al.*, 1991). The second sub-clade of the C-lineage is *PLENA* (*PLE*) with two genes *SHP2* and *SHP2*, which were revealed to be involved in the organisation of cell divisions in gynoecium and seed coat endothelium and for endosperm development, together with *B* sister and *SEEDSTICK* in *A. thaliana* (Ehlers *et al.*, 2016).

C-class *MADS-box* genes duplicated and increased in number during the evolution of the grasses. Rice, for instance, has two *AG* orthologues, *OsMADS3* and *OsMADS58*, and their functions are diversified (Yamaguchi *et al.*, 2006). Whereas the floral meristem (FM) determinacy is partially compromised in an *Osmads3* mutant and severe loss of determinacy was observed in knockdown lines of *OsMADS58*. In these lines, a set of floral organs (lodicules, stamens and partial carpels) is repeatedly formed in flower, and a floral meristem-like structure remains, even in the mature flower (Yamaguchi *et al.*, 2006).

Despite the fact that reverse genetic analysis revealed that an *OsMADS58* mutation had little effect on floral phenotypes, but dramatically enhanced the indeterminate phenotype of the *Osmads3* single mutant, indicating the importance of *OsMADS58* in floral meristem (FM) determinacy in rice (Dreni *et al.*, 2011). Considering all of this, the evidence suggests that C-class *MADS-box* genes play crucial roles in regulating the determinacy of the FM in both *Arabidopsis* and rice.

In wheat two paralogues WAG1 and WAG2, belonging to the C-lineage were identified and found to be expressed in the third and fourth whorl, in addition to the first whorl presented by the lemma and palea from bolting to heading stage. These genes were associated with carpel formation (Meguro *et al.*, 2003; Zhou *et al.*, 2008; Paolacci *et al.*, 2007)

In rice two D-lineage genes, *OsMADS13* and *OsMADS21* were isolated and believed to have arisen before the diversification of *Poaceae* around 66–70 MYA (Kramer *et al.*, 2004; Xu & Kong, 2007). Reverse genetics have shown that *OsMADS13* is essential for specification of the ovule in the 4th whorl (Lopez-Dee *et al.*, 1999; Dreni *et al.*, 2007; Yamaki *et al.*, 2011). Its mutants are female sterile, and ovules are converted into carpel-like structures, creating indeterminate flowers with carpels inside carpels (Dreni *et al.*, 2007). *OsMADS21* is likely a paralogue of *OsMADS13* and shows high sequence similarity. However, it lost roles in ovule identity determination, suggesting that *OsMADS21* was a subject of functional diversification of these two D-lineage genes since they arose by duplication and has probably lost its function in rice (Dreni *et al.*, 2007; Xu & Kong, 2007). Two *MADS-box* genes belonging to the D-class, FBP7 and FBP11 in *Petunia* was revealed to be involved in the regulation of ovule organ identity (Colombo *et al.*, 1995; Angenent *et al.*, 1995). In *Arabidopsis*, only SEEDSTICK (STK) was identified as D-lineage that determines the identity of ovules and involve in the regulation of several metabolic processes during seed development. One of the processes is epicatechin synthesis, which is essential in plant protection and flavour enrichment (Mizzotti *et al.*, 2014).

In *Brachypodium*, the D-lineage gene *BdMADS21* (*BdMADS4*) was found to be lowly expressed in the glume, palea and carpel. While its paralog *BdMADS13* (*BdMADS2*) showed a similar pattern except that it expresses highly in the carpel. Overexpressing *BdMADS21* (*BdMADS4*) gene in *Arabidopsis* result in significant phenotype with curly rosette, cauline

leaves and short filaments with sterile stamen compared to the over-expression of *BdMADS13* (*BdMADS2*) (Wei *et al.*, 2013). These results suggest that after the ancient duplication prior the divergence of angiosperms and gymnosperms giving rise to the C and D-lineage *MADS-box* genes in the AG clade, another species-specific genome-wide or segmental duplications resulted in additional paralogues in each lineage. In rice, *OsMADS21* lost its function while in *Brachypodium* conserves its function. No functional studies were performed on *BdMADS3* (*BdMADS18*) or *BdMADS58* (*BdMADS14*) to our knowledge, however, their expression showed different patterns where the *BdMADS3* (*BdMADS18*) expresses in lodicule, stamen, carpel and young grain while *BdMADS58* (*BdMADS14*) expression was restricted to the stamen (Bo Wei *et al.*, 2014).

In wheat, only one member of D-lineage was identified called *WSTK/TaAGL2*, which is an ortholog of STK in *Arabidopsis* and *OsMADS13* in rice and found to be express in the carpel and forms a complex with the class E protein WSEP (Yamada *et al.*, 2009). However, in the current study, we demonstrate that wheat has two paralogues in the D-class, *TraesCS5A02G117500*, *TraesCS5B02G115100*, *TraesCS5D02G118200* homoelogs present high sequence similarity to *OsMADS13* and *TraesCS1A02G262700*, *TraesCS1B02G273300*, *TraesCS1D02G262700* homoelogs which seems to be orthologues of *OsMADS21* in rice (Chapter 2).

To date, a dynamic description of C and D gene expression has not been established. In the current study, we selected the orthologues of the *OsMADS3* and *OsMADS58* belonging to the C-class and *OsMADS13* and *OsMADS21* belonging to D-class in *Brachypodium distachyon* and named them based on their homology to their orthologues in rice. The expression results revealed that the members of C and D-class expression start in very early stages of flowering phase and gradually increases in floral organ and persist in the different stages of grain development, these results indicate that these homeotic genes have a role beyond flowering regulation.

3.3 Materials and Methods

3.3.1 Plant materials

In this study we used *Brachypodium distachyon* (diploid inbred line Bd21) obtained from Dr Sinead's lab. This line is fully sequenced and has a diploid genome of ~355 Mb. The inbred line *Bd21* seeds were de-husked and grown as described (Draper et al., 2001), seeds were vernalised on moist filter paper in a 9mmØ Petri dish at 5°C for 48h in order to break the dormancy and stimulate flowering, then moved to room temperature for 48h. The germinated seedlings were then sown in a soil mixture of two-part of the soil, one vermiculite and one sand supplied by Levington® then transferred to a phytotron (controlled environment) under the following conditions: photosynthetic photon flux density (PPFD) of 400 $\mu\text{mol m}^{-2} \text{s}^{-1}$, 20 h light/4 h darkness, with a day/ night temperature of 24°C under 52% relative humidity. Plants were watered by flooding under tray every other two days.

3.3.2 Tissue Collection and storage

In this study, all plant materials from *Brachypodium distachyon* used for PCR and RT-PCR were collected in specific growth stages from the life cycle of the plant including; vegetative tissues: root 3, seedling 3, leaf 18, glume 18, 2nd leaf 26, stem 32, flag leaf 49; reproductive tissues: very inflorescence 33, emerging spikelet 34, pre-anthesis spikelet 34, lemma 34, palea 34, carpel 30, stamen 30, lodicule 30; grain development: young grain, mid-length grain, full-length grain, mature grain, they were immediately weighed and flash-frozen in liquid nitrogen after collection and stored in Eppendorf tubes at temperature of -80°C. These developmental stages were collected from two biological replicates.

3.3.3 Nucleic acid Analysis

3.3.3.1 Isolation of total RNA from the plant tissue samples

Total RNA was extracted from frozen ground tissues using either TRIsure reagent (BIOLINE) for RT-PCR following the manufacturer instructions, or the ISOLATE II RNA extraction kit (BIOLINE) for the quantitative RT-PCR due to its excellent purification of the RNA from any inhibitors. In the first method, 1ml of TRIsure reagent was applied to finely ground tissue in a mortar and left covered to return to room temperature for 5 min to prevent any RNase contamination then pipetted into Eppendorf tubes. 200 μl of phenol-chloroform

was added to eliminating proteins then vortexed and left at room temperature for 2 minutes. The cells debris were precipitated by centrifuging the tubes 11,400 rpm for 10 minutes at 4°C, 500 µl of the nucleic acid in the supernatant were transferred to a new Eppendorf tube and an equal part (500 µl) of isopropanol was added to precipitate the nucleic acid then left to stand 5 minutes on ice before centrifuging another 10 minutes at 11,400 rpm, 4°C. The supernatant was discarded, and the pellet was washed with 70% Ethanol and left to dry for ~15 minutes at room temperature, then dissolved in ~20-30 µl of RNase free water. The nucleic acid was then treated with 2 µl of NEB DNase I enzyme (in a 1x reaction buffer (10mM Tris-HCl, 2.5mM MgCl₂, 0.5 mM CaCl₂, pH 7.6) for each 20 µl by a 10 minutes activation at 37°C followed by another 10 minutes deactivation at 75°C and finally stored at -80 °C for RT-PCR.

Quantitative RT-PCR required extraction of the total RNA using the ISOLATE II RNA extraction kit from (BIOLINE) due to the purification of the RNA from any inhibitors that could cause false sensitivity in qRT-PCR experiments. The process started by mixing 3.5 µl of β-mercaptoethanol and 350 µl of the lysis buffer then pour the mixture to finely ground tissue in the mortar and left to thaw at room temperature before transferring it to the supplied filter column then centrifuged to remove any cell debris. The nucleic acid was collected and mixed with 350 µl of 70% ethanol by pipetting up and down 5 times. The mixture was transferred to the supplied filter column then centrifuged 30 seconds at 11,000 xg, the membrane was treated with 350 µl of desalting buffer and centrifuged at 11,000 xg for 1 minute. In the meantime a DNase I enzyme was prepared by adding 10 µl of DNase I to 90 µl of its reaction buffer (RDN) and applying 100 µl for each sample then left it for 20 minutes at room temperature. The membrane was washed once with 200 µl RW1 buffer then twice with 600 µl RW2 buffer before eluting the total RNA in a new Eppendorf with 20-30 µl RNase-free water.

Visualization of total RNA was performed using gel electrophoresis (1% agarose mixture with TBE X1). 2µl of total RNA, 2 µl of loading dye (6X) and 8 µl dH₂O was loaded on the gel and run at 80V for 30 min. The RNA integrity was assessed by finding two distinct RNA band on the gel. The samples were stored at -80°C until required.

3.3.3.2 Quantification of the nucleic acid

To identify the concentration of the RNA/DNA, the absorbance at 260nm and 280nm was measured then calculate the ratio between the absorbance values (Abs260: Abs280) using the spectrophotometer NanoDrop™ 2000 UV-Vis (Thermo Scientific) and NanoDrop™ 2000/2000c software (Thermo Scientific) to calculate the concentration for each type; 1 Abs260 unit equates to 40 µg/ml of RNA or 50 µg/ml of double-stranded DNA. First, 1 µl for the blank (RNase-free water for the RNA and TE for the DNA) was 1.5 µl of the sample was pipetted onto the pedestal of the spectrophotometer platform then the samples were measured successively. The software then calculated the concentration of RNA or dsDNA based on the absorbance ($260/280 \approx 1.8$, $260/230 \approx 2.0$) and the default extinction coefficient. The result table was saved as excel file for the subsequent cDNA synthesis step.

3.3.3.3 First Strand cDNA Synthesis

Approximately 700ng RNA was used in 20µl cDNA synthesis reactions. The following first-strand cDNA synthesis reaction mix was then added:

Table 3-1: First-strand cDNA synthesis reaction mix

First-strand cDNA synthesis reaction mix	Quantity
Total RNA	700ng
10 µM oligodT primer	0.5 µl
0.1% DEPC water	Up to 12 µl
Incubation at 70°C for 5 minutes	
Total RNA with annealed oligodT primer	12 µl
0.1% DEPC water	2.1 µl
5x Transcriptase enzyme Buffer	4 µl
10 mM dNTPs	1 µl
RNase inhibitor (NEB)	0.4 µl
Tetro Reverse Transcriptase enzyme (200u/ µl) (BIOLINE)	0.5 µl
Incubation in the thermocycler at 42°C for 30 min	
Incubation in the thermocycler at 95°C for 5 min	

The first-strand cDNA was diluted 1:5 and stored at -20°C until required.

3.3.3.4 Genomic DNA Isolation

Genomic DNA extraction was performed using approximately 300ng of young leaf tissue, which was collected into a 1.5ml Eppendorf tube, with glass beads (Sigma-Aldrich) with a size of 425-600µm for genomic DNA extraction and frozen in liquid nitrogen. The grinding of the tissues was performed using the tissue laser Amalogamator ZoneRay® for 12 sec. The ground sample was left to thaw at room temperature for 5 minutes then 300 µl of the extraction buffer (200 mM Tris-Hcl pH 7.5, 250 mM NaCl, 25 mM WDTA, 0.5% SDS) was

added to each sample. The mixture was incubated at 50°C for 30 minutes then vortexed and 1 volume (300 µl) of Phenol:Chloroform (SIGMA) was added before centrifuging at maximum speed for 5 min. The supernatant was transferred to a new Eppendorf, and 1 Vol (300 µl) of isopropanol was added for precipitation. This was followed by incubation for 30 min at 20 °C, then centrifuging at maximum speed for 10 min. The pellet was washed with 600µl 70% ethanol and left drying at room temperature before dissolving it in 20-30 µl Tris-EDTA buffer. The DNA was quantified using a NanoDrop™ 2000 UV-Vis spectrophotometer (Thermo Scientific) and stored at -20°C ready to be used for PCR.

3.3.3.5 Reverse transcription polymerase chain reaction (RT-PCR)

Oligonucleotide primers were manually designed to amplify gDNA and cDNA products using SnapGene Viewer 2.8.3. The primers were designed with a length ranging between 21 - 25 bp and annealing temperature (T_m) between 55 - 65°C. Sequences for all primers used in this study are shown in the appendix (table S1). Primers were obtained from SIGMA and were diluted (1:10) to 10µM concentration for all reactions. Reactions were performed using Dream Taq Green PCR Master Mix (Thermo Scientific).

PCR/RT-PCR reaction mix	Volume
Dream Taq Green PCR Master Mix(2X)	5 µl
5 µM forward primer	0.5 µl
5 µM reverse primer	0.5 µl
dH ₂ O	3 µl
Templet cdna 1µg	1 µl

PCR conditions were typically similar for both genomic and cDNA amplification, being 1µl of template in 10µl reactions with the following cycles: 95°C 6 minutes for denaturation, then 30-35 cycles of 95°C 30 seconds, the annealing temperature vary between 55-65°C for 30 seconds, 72°C for 1 minute followed by 72°C for 6 minutes for final extension. Reaction conditions were varied in some cases to according to product size, primer T_m and individual gene copy numbers.

3.3.3.6 Gel Electrophoresis

The separation of nucleic acids by agarose gel electrophoresis was performed using molecular grade agarose powder (Bioline), which was melted in 1X TAE buffer (40 mM Tris base, 20 mM glacial acetic acid, 1 mM EDTA), cooled to ~50°C after which ethidium bromide was added at a final concentration of 0.2 µg/ml (0.5 nM). 1% w/v agarose as a concentration was adopted for all the amplicons. A molecular weight ladder 100bp/1Kb

(NEB) was also loaded. The loaded gel was run at 100 v-500mA for ~30 minutes using Labnet Enduro E0303 Model 300V Power Supply. A UV transilluminator NuGenius (5 million pixel camera, lens of 8 – 48mm f/1.2 and a sensor of 1/2.5 inch) was used visualise the gel. The size(s) of DNA or RNA was determined using the molecular weight ladder (100bp/1Kb).

3.3.3.7 Quantitative RT-PCR

Quantitative RT-PCR reactions were carried out using SYBR Green (SensiMix SYBR Low-ROX Kit) (BIOLINE) on Applied Biosystems™ 7500 Real-Time PCR Systems with 7500 Software v2.0.6, on 96well MicroAmp® Fast Optical 96-Well Reaction Plate with the following reaction mix components:

Reaction mix composition	Volume
2x SensiMix™ SYBR® Low-ROX	5 µl
25 µM Forward Primer	0.1 µl
25 µM Reverse Primer	0.1 µl
DNase-free H ₂ O	3.8 µl
cDNA template	1 µl

The thermal cycling conditions started by polymerase activation at 95°C for 10 min followed by 40 cycles of 95°C for 15 seconds then the annealing temperature of 65°C for 15 seconds flowed by 72°C extension for 15 seconds.

Twenty-one tissue samples from root three days after germination (DAG) to mature grain 78 DAG, were used to quantify the expression of the target genes as well as the control. The stock of the primers is prepared to be 100µM (100pg/µl) then diluted 1:10 to have a working solution of 10µM (10pg/µl). In order to know the primer efficiency, a standard curve has been generated base on 5-fold dilution (1:5, 1:25, 1:125, 1:625, 1:3125) with three replicates of each sample. Quantitative values for gene expression were calculated using *BdACT7*, *BdGAPDH* and *BdUBC18* as a reference gene; the efficiency was 97.63%, 92.35% and 107.24% respectively. The same was done for *BdMADS3*, *58*, *13* and *21* and their efficiency was 110%, 106%, 102% and 94% respectively (the range between 90 and 110% considered acceptable for qPCR reactions). In order to calculate the efficiency of the primer, a standard curve was generated to get the slope value from the following equation:

$$y = -slope(x) + y-intercept$$

y-intercept: the point where the line crosses the y axis

Once the slope has been calculated, the efficiency of each primer was generated base on the flowing equation:

$$efficiency = (10^{-1/slope} - 1) * 100$$

The T_m ° of the primers was 65°C, GC content of 45%, no dimer nor secondary structure were formed. From the standard curve a C_t value that between 15 and 30 (Yuan *et al.*, 2006) has been selected to determine the working concentration because it can be reproducibly measured leading to a dynamic range of 10^5 , which is within the 4 to 8 logs dynamic range reported in other studies. Therefore, in our case, it was a dilution of 1:5 (0.2).

To overcome the significant variation of the c_t values (Hong *et al.*, 2008) between the different developmental stages, we select the best housekeeping gene by performing statistical analysis using NormFinder software. *BdUBI18* and *BdACT7* showed high stability in the C_t values throughout the stages.

Reference genes	Gene stability value
<i>BdUBI18</i>	0.010
<i>BdACT7</i>	0.015
<i>BdGAPDH</i>	0.032

Based on these results, the *BdUBI18* was used as a reference gene, and a combination of *BdACT7* and *BdUBI18* was used as well.

As described in the Livak method (Livak & Schmittgen, 2001) the ΔC_t for each sample was calculated as follows:

$$\Delta C_T(sample) = C_T(target\ gene, sample) - C_T(reference\ gene, sample)$$

The very young inflorescence sample was chosen as a calibrator to calculate the $\Delta\Delta C_t$:

$$\begin{aligned} \Delta C_T(calibrator) &= C_T(target\ gene, calibrator) - C_T(reference\ gene, calibrator) \\ \Delta\Delta C_T(sample) &= \Delta C_T(sample) - \Delta C_T(calibrator) \end{aligned}$$

Relative Quantitation of Gene Expression (normalised expression fold-change) was calculated with the following equation:

$$RQ = 2^{-\Delta\Delta C_T(sample)}$$

3.3.4 mRNA *In situ* Hybridisation.

The protocol used in this study follows (Drea *et al.*, 2005).

3.3.4.1 Samples collection and preservation

For mRNA *In situ* Hybridisation experiments, different tissue samples and developing grains stages were freshly collected and fixed in formalin–acetic acid–alcohol (FAA) (3.7% formaldehyde, 5% acetic acid, 50% ethanol) and vacuum infiltrated under the pressure of

400 mbar (KNF LABOPORT diaphragm lab vacuum) for 5 minutes then left overnight at 5°C. Three washes with 70% ethanol was performed before the samples goes through an increasing concentration of ethanol for dehydration then gradually removing the ethanol using an increasing dilution of the clearing agent Histo-Clear (National Diagnostics) starting from 25%, so the samples can infiltrate the paraffin wax (Formula 'R' Paraffin, supplied by Leica Biosystems) and prepared for the embedding process as follows:

- Successive bath of ethanol for one hour 70%, 85%, 95%, 100% at 4 °C
- Successive bath of Histo-clear four 30 min, 25%, 50%, 75% at room temperature
- 100% Histo-clear for overnight at room temperature
- At 60°C:
 - 100% Histo-clear for 1h
 - 50% Paraffin /50% Histo-clear 6h
 - 100% Paraffin overnight
 - 100% Paraffin 8h
 - 100% Paraffin overnight

All steps were performed manually using HISTOSETTE® Tissue Processing/Embedding Cassettes.

3.3.4.2 Tissues embedding and sectioning.

The paraffin wax infiltrated samples were then transferred to the Tissue KEDEE embedding console (KD-BMIV tissue embedding centre) for embedding in paraffin blocks. Wax sections of 8µm thickness were cut on a mechanical Microtome (KEDEE, KD2258), and organised on permanent positive charged slides (Superfrost® plus, Thermo scientific) and left overnight on a slide warmer (Thermo scientific hot plate) at 40°C.

3.3.4.3 Probe Transcription and Labelling.

The probe transcription performed using gene-specific fragments (*BdMADS3*, 58, 13 and 21) (table S2) which has been amplified from cDNA using T7 reverse (GAATTGTAATACGACTCACTATAGGG) appended primers It was amplified using DreamTaq PCR Master Mix 2X (Thermo Scientific) then purified using E.Z.N.A.® Cycle Pure Kit (Omega Bio Tek) according to the manufacturer's instructions to get higher recovery of purified amplicon. The product was used to generate digoxigenin-labelled RNA probes.

The transcription was performed in-vitro in 20µl reactions for 2 hours at 37°C in the following reaction: 5 µl from 0.1% diethylpyrocarbonate (DEPC), 2 µl of dNTPs (10 µM), 2 µl from digoxigenin-11-UTP 200 nmol (57µl) (Roche), 1 µl of RiboLock RNase Inhibitor (40 U/µL) (Thermo fisher Scientific-2500U), 2 µl from T7 RNA polymerase 50U/µl (New England

Biolabs), 2 µl of 10x RNA polymerase reaction buffer, finally 6 µl from *BdMADS3/58/13/21* Probe template (Purified DNA). The transcription reactions were precipitated in 4 M Lithium chloride (LiCl) and 3 volumes of absolute ethanol (60µl) overnight at -20°C. Reactions were spin for 10 minutes at 11,400 xg at 4°C and pellets re-suspended in 20µl TE (100 mM Tris, 10 mM EDTA) buffer.

The prober were tested using the dot blot by dilution of 1:100 as following; on nitrocellulose membrane 1µl of the diluted probe were spotted, along with three serial dilution (1:50, 1:100 and 1:200) of known concentrations for dot-blot probe testing then baked for 3 min at 50°C, followed by a brief wash with 1X TBS (100 mM Tris-Hcl, 500 mM NaCl), nitrocellulose membrane was treated with the blocking solution (0.1g milk powder in 1xTBS) for 30 min , followed by another 30 min in 1:5000 Anti-Digoxigenin-AP (Fab fragments) 150U from Sigma-Aldrich (1µl antibody in 5ml 1x TBS). After that, the excess of the antibody was washed for 5 min with 1x TBS, then AP-buffer (100mM Tris-HCl, 100mM NaCl, 5mM MgCl₂, 0.05% Tween 20, pH 9.5) for 5 minutes to equilibrate to AP environment, finally, it was developed in AP-Buffer containing NBT (nitroblue tetrazolium:0.1 mg/ml) and BCIP (5-bromo-4-chloro-3-indolyl phosphate-p-toluidine salt: 0.075 mg/ml) in darkness. Blots were developed until strong colour products could be seen and probe strength was compared to the three dilution control mad (1:50, 1:100 and 1:200)

3.3.4.4 Slide pre-treatment.

The tissues on the slides were deparaffinised for 10 minutes in two histo-clear clearing agent baths (National Diagnostics) then rehydrated through a decreasing series of ethanol (100%, 80% then 50%), followed by two 5 minute rinse with 1x PBS (1.8 mM KH₂PO₄, 10 mM Na₂HPO₄, 137 mM NaCl, 2.7mM KCl). In order to eliminate all proteins, 30 min of pre-warmed proteinase K treatment (100ul from proteinase K enzyme added to 250ml buffer (100 mM Tris, 10 mM EDTA pH7.5)) was performed for at 37°C, then 2 minutes in glycine (5ml 10% glycine, 250ml 1xPBS). PBS rinse 5 min; acetic anhydride (0.5% in 0.1 M triethanolamine pH 8) 10 min with stirring; PBS rinse 5 min; then ethanol increasing series, 50%, 80% and 100%. Slides were air dry and stored at 4°C until hybridisation.

3.3.4.5 Hybridisation and washing.

All probes were used at a 1:50 dilution, therefore, the amount used was 32µl of the probe in 1600µl of hybridisation solution (HS) for 8 slides. The HS contains; 200µl of 10X *in situ*

salts, 800 µl of Deionized formamide, 400 µl 50% w/v dextran sulfate, 40 µl 50X denhardts solution, 20 µl tRNA and 1600 µl DPEC H₂O. The 10X *in situ* salts (100mM Tris-HCl, 100mM Na₃PO₄, 50mM EDTA, pH 8.0), 50% v/v deionized formamide (Sigma Aldrich) , 50% dextran sulfate, 50X Denhardts solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone) and 0.1 mg/mL tRNA). Immediately prior to hybridisation probes were denatured at 80°C for 3 minutes and instantly on ice. ~200 µl of probes mixture were applied to slides and covered with glass coverslips and hybridisation was performed overnight in a 50°C incubator. 0.2% SSC was used to remove the coverslips and slides were washed through the following three washes of 0.2% SSC (Saline Sodium Citrate) preheated to 50°C with gentle agitation, 1% blocking solution in 1 x TBS for 1hr, 1% BSA (Bovine Serum Albumin) in 1 x TBS + 0.03% TRITON for 1 hr, 1 x TBS containing a 1:3000 dilution of anti-digoxigenin–alkaline phosphatase and 0.05% Tween 20 for 1 h, followed by 3 x 10 minutes washes in 1 x TBS with gentle agitation and a final 5 min in AP-buffer.

3.3.4.6 Slide development and Mounting.

The colour reaction was developed in AP buffer containing NBT (0.1 mg/mL) and BCIP (0.075 mg/ mL) for up to 48 hrs. Once the visible colour was developed, slides were washed for several times in distilled water to stop the reaction, followed by sequential washes in 50%, 70% and 100% ethanol. Slides were then allowed to air dry and permanently mounted in DPX mounting medium (Fisher Scientific).

3.4 Results

3.4.1 Expression pattern analyses of C and D lineage genes in *Brachypodium*

Grass inflorescences are characterised by the formation of several spikelets that are arranged in alternate distichous form. Each spikelet contains several florets that are composed of the lemma, palea, lodicules, stamen and carpel. In *Brachypodium distachyon* each spike contains two to three spikelets with cleistogamous florets (self-pollination).

With the aim of performing a comprehensive study and investigating the expression patterns of our target genes from C-lineage *BdMADS3-58* and D-lineage *BdMADS13-21*, samples were collected across the life cycle of the *Brachypodium distachyon* panning vegetative tissues, reproductive tissues and grain development, encompassing seedling to the mature grain.

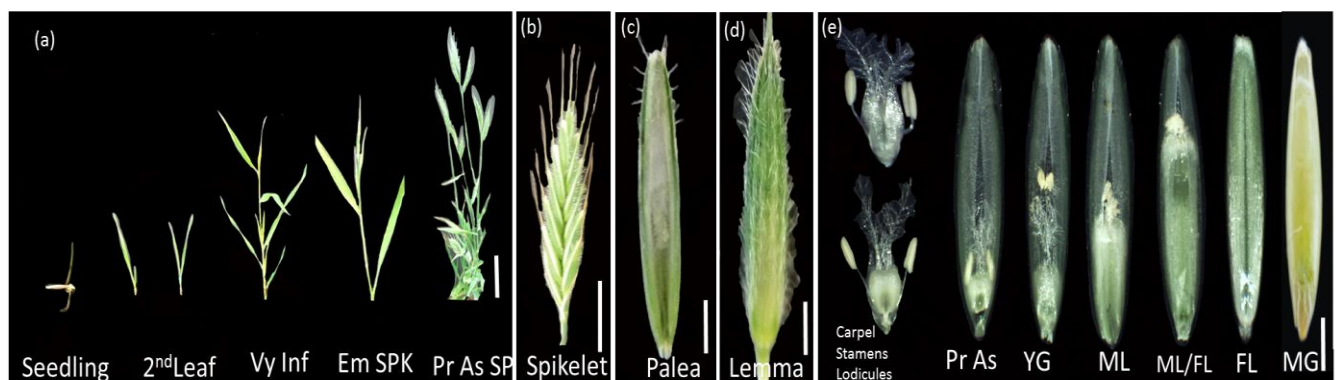


Figure 3-3: Sampling stages of *B. distachyon*. (a) Developing plant phenotype, Very young inflorescence (VyInf), Emerging spikelet (Em SPK), pre-anthesis spikelet (PrAs SP), (b) mature spikelet in full grains, (c-d) palea and lemma of the floret, (e) developing stages of the grain, young grain (YG), mid-length grain (ML), full-length grain (FL), mature grain (MG), scale bars (a)5 cm,(b) 5mm, (c-e) 1 mm.

3.4.1.1 The expression pattern of the C and D-lineage using RT-PCR

After extraction the RNA from the samples, cDNA was synthesised, and RT-PCR was performed using primers for the two candidates gene form C-lineage and the D-lineage the figure 3-4 below shows the amplified transcript from each sample.

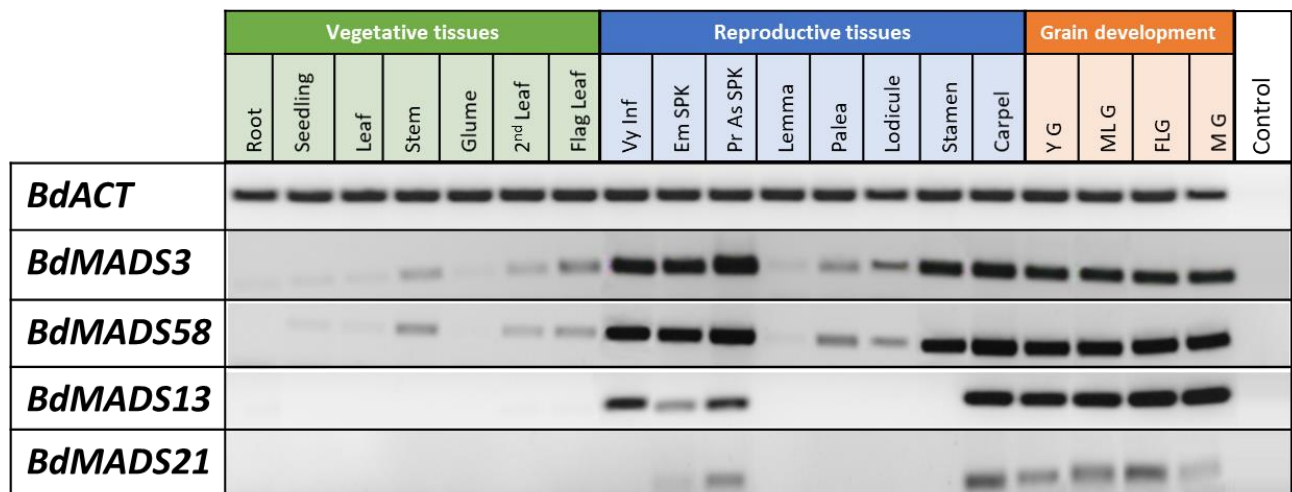


Figure 3-4: Expression patterns of *Brachypodium* *MADS3*, *58*, *13*, and *21* TFs using Reverse transcription-PCR in different plant stages; vegetative, reproductive and grain development tissues, the housekeeping genes is β -actin (*BdACT*) for 30cycles, F. leaf; Flag leaf, Vy Inf SPK; very young inflorescence, Em SPK; emerging spikelet, Pr As; Pre-anthesis, YG; young grain, MLG; Mid-length grain, FLG; full-length grain, MG; mature grain.

The results (Figure 3-4) show remarkably high expression during the flowering stage in the reproductive tissues for all our target genes (*BdMADS3*, *58*, *13* and *21*) and low expression in the vegetative tissues. However, the expression persists during grain development which suggests that they have a crucial role beyond the flowering stage.

The paralogues *BdMADS3* and *BdMADS58* showed a high similarity in their expression patterns starting in the very young inflorescence where we find the floral organ primordia and then focused more in the third and fourth whorls from pre-anthesis until mature grain with no remarkable expression in the lemma, palea and lodicules in the first and second whorls. The two D-lineage paralogues *BdMADS13* and *BdMADS21* were expressed during the formation of the floral organs; however, their activity was concentrated in the fourth whorl represented by ovule and carpel.

Previous studies on eudicots highlighted the homeotic function, for instance in *Arabidopsis*, *AGAMOUS* (AG) acts to specify stamen and carpel identities and floral meristem determinacy (Bowman *et al.*, 1991). Studies in rice considered that *OsMADS3* and *OsMADS58* might have distinct functions in specifying stamen identity, with *OsMADS3* playing a more important role (Yamaguchi *et al.*, 2006).

3.4.1.2 Expression patterns of C and D-lineage genes using RNA *in situ* Hybridisation

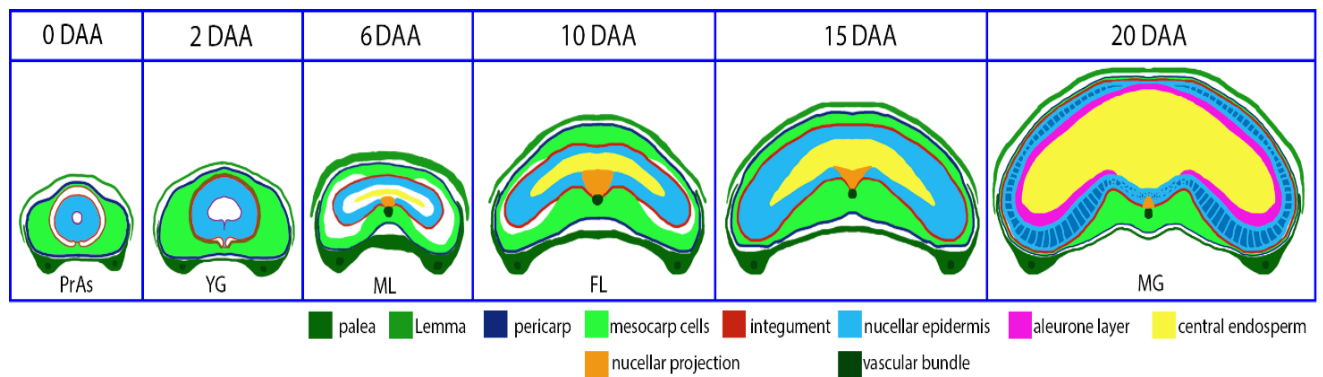


Figure 3-5: Schematic summarising *Brachypodium distachyon* grain development illustrating the different grain layers growth. DAA days after anthesis.

In major core *pooids* including wheat and barley, there are four main cell types that form the endosperm; transfer cells, aleurone cells, starchy endosperm cells, and embryo-surrounding region cells. The endosperm is enclosed with layers of maternal tissues such as the nucellar epidermis/projection, which compress and degenerate at mature stages. However, these layers in wild grass *Brachypodium* persist (figure 3-5) (Hands *et al.*, 2012). Moreover, observation of the aleurone layer showed that there was no differentiation of peripheral and modified aleurone regions. This contrasts with members of the *pooids* that present large-celled endosperm in lobed structure that encloses the crease and nucellar projection (Figure 3-5) (Opanowicz *et al.*, 2011).

In our project grain development was staged using pollination event; pre-anthesis was assigned to 0 days after anthesis (0DAA), young grain stage is 2 days after anthesis (2DAA), mid-length stages is 6 days after anthesis (6DAA), full-length stage is 10 days after anthesis (10DAA), and mature grain stage is 20 days after anthesis (20DAA).

In order to determine spatial and temporal expression patterns within the tissues, where gene expression was observed with RT-PCR, mRNA *in situ* hybridisation (ISH) is performed using labelled complementary RNA (antisense probe) from the divergent 3' ends of the cDNAs.

The mRNA *in situ* hybridisation analysis (Figure 3-6) detected the expression of *BdMADS3* during the early flowering stage in the spikelet meristem and the carpel primordium when the floral organs are initiating. The mRNA of *BdMADS3* was identified at the pre-anthesis

stage in longitudinal section of the carpel along the layer of the endocarp (figure 3-6 (a)) in addition to some expression in the lateral side of the carpel in the mesocarp cells. Subsequently, the expression increases in the young grain stage in the same areas then focuses more on the outer integument and differentiating cells of the nucellus parenchyma cells which retained a considerable degree of cellular integrity, consisting of multiple cell layers (figure 3-6 (c)). After the degradation of most of the peripheral nucellar parenchyma (mesocarp) cells at mid-length stage, *BdMADS3* transcript expression was detected in the remaining pericarp. As the grain start elongating a signal of mRNA appears with the development of the endosperm at full-length (figure 3-6 (e)) and mesocarp, whereas no transcript was detectable in the nucellar epidermis. This expression persisted until the mature grain stage and concentrated more in the aleurone layer as the endosperm expanded rapidly while the mesocarp continued to break down. (Figure 3-6 (h)) shows a strong signal of the transcript of *BdMADS3* in the mature pollen grain whereas there was no expression at this stage in the anther wall.

These results are consistent with the RT-PCR previously shown and suggest that *BdMADS3* is expressed in the tissues and cells that are multiplying and dividing, for instance, the integument and nucellus as well as expression in the cells that are degenerating and shrinking like the mesocarp (figure 3-6 (e)).

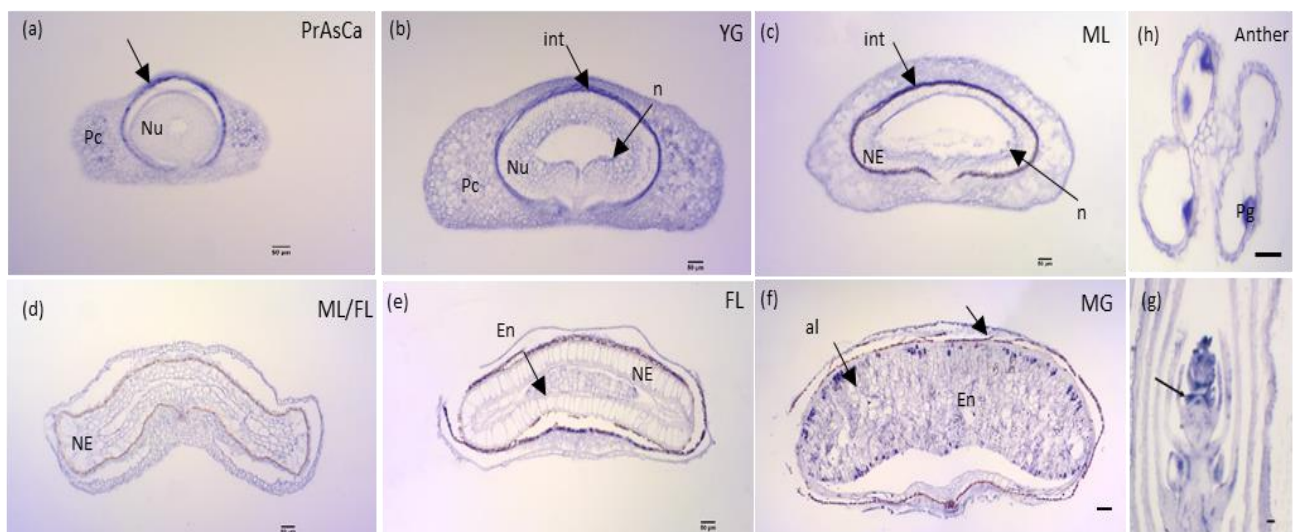


Figure 3-6: The mRNA expression pattern of *BdMADS3* in meristem and grain longitudinal sections. (a) Pre-anthesis carpel (PrAsCa) Arrows indicate inner-carp, pericarp (Pc), (b) young grain (YG) Arrows indicate outer integument, , nucellus parenchyma cells (Nu) and nucellar lysate (n), integument (Int), (c) Mid-length grain (ML) Arrows indicate outer integument and nucellar epidermis (np), nucellar epidermis (NE), (d-e) Mid and full length stage (ML/FL) Arrows indicate endosperm, (f) mature grain (MG) Arrows indicate aleurone

(al) and outer integument, (g) spikelet meristem (SPK Mr), (h) mature anther, pollen grain (Pg), Scale bar: 50 μ m. For negative control see Figure S3 Appendix.

The other C-class paralogue *BdMADS58*, similarly was expressed at the apical side of the meristem and in the carpel and floral organ primordia at the early floret emerging stage (Figure 3-7 (a)). After this stage, the round mass of actively dividing cells in the main bulk of the primordium becomes the nucellus just adjacent to the differentiating integument where the transcript of *BdMADS3* accumulates at the pre-anthesis stage with some signal in the mesocarp (Figure 3-7 (b)). The spatial expression then distributes in the pre-anthesis carpel in addition to a small signal in the pollen grain (Figure 3-7 (b)), after pollination the carpel expands and expression is focused more in the nucellus chlorenchyma cells and some endocarp mesocarp (Figure 3-7 (b)) In subsequent stages the signal migrates to the outer integument and persist in the mesocarp (figure 3-7 (c)). During maturation periods in the full length and mature grain, the signal could be detected in the aleurone layer in addition to the palea and lemma, specifically in the bundle sheath around the xylem (figure 3-7 (f-g)), indicating possible roles in these organs as well. The pollen grain showed high expression of the *BdMADS3* mRNA. Therefore, our results suggest that the activity of *BdMADS58* could also be involved in grain development, especially storage of the lipid and enzymes in the aleurone layer during maturation.

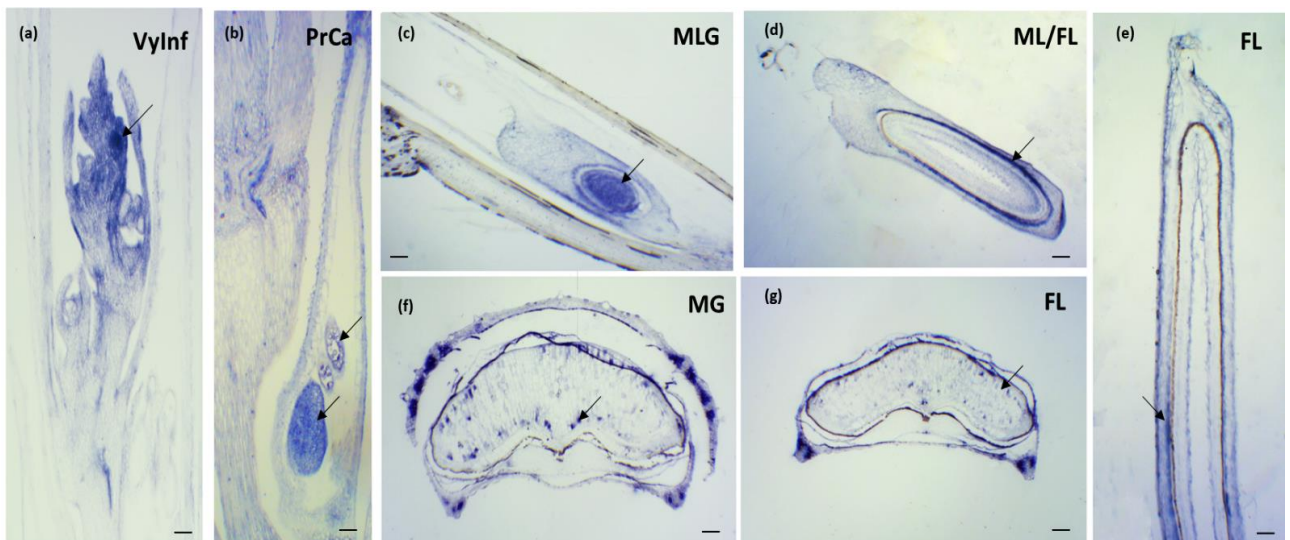


Figure 3-7: *BdMADS58* mRNA in different stages of *Brachypodium distachyon*. Pre-anthesis spikelet (Pr As SPK), Mid-length grain (ML grain), Full-length grain (FL grain). Scale: (a-b) 25 μ m, (c-d) 50 μ m.

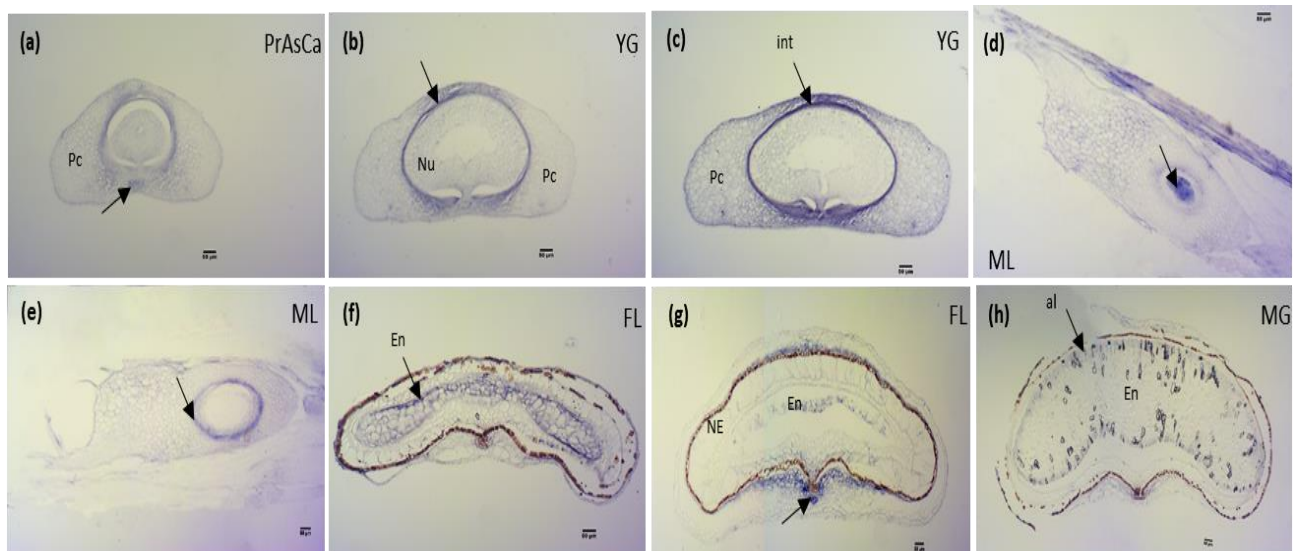


Figure 3-8: Spatial and temporal expression of *BdMADS13* mRNA in a longitudinal and transversal grain sections of *Brachypodium distachyon*. (a) Pre-anthesis carpel, (b) young grain, (c-e) Mid-length grain longitudinal and transversal section (ML), (f-g) full-length grain (FL), (h) mature grain (MG), arrow indicates vascular bundle in (a), the chlorenchyma cells in (b-e) and the endosperm (En) in (f), mesocarp and vascular bundle in (g) and aleurone layer (al) in (h), pericarp (Pc), nucellus parenchyma cells (Nu), nucellus epidermis (NE), integument (Int), scale bar: 50 μ m.

The mRNA of *BdMADS13* (Figure 3-8) was detected in the longitudinal section of tissues at different developing stages. Starting with the first stage of grain development, a weak signal was observed around the chlorenchyma cells of the endocarp and concentrated in the abaxial side of the carpel where the cells are larger (Figure 3-8 (a)). The signal of *BdMADS13* mRNA increases after anthesis during the young grain and mid-length stages in the same tissue, however, it appears weak in the pericarp in the adaxial region near the stigma (Figure 3-8 (d-e)). Moving on to the full length stage at 10 DAA, the mRNA seems to accumulate in the chlorenchyma cells of the mesocarp, only a thin layer of the adaxial side of the grain and around the vascular bundle (Figure 3-8 (g)), in the subsequent stages (15DAA), a very strong signal was detected in the developing endosperm which may indicate that *BdMADS13* positively regulates the endosperm cellularisation and proliferation genes.

mRNA could not be detected in the nucellar epidermis at the full-length stage (Figure 3-8 (g-f)). The *BdMADS13* transcript expression focuses more on the aleurone layer with a weak signal at the maturity stage where the grain stores nutrient and starch in the endosperm, the vascular bundle maintains some expression as well (Figure 3-8 (h)).

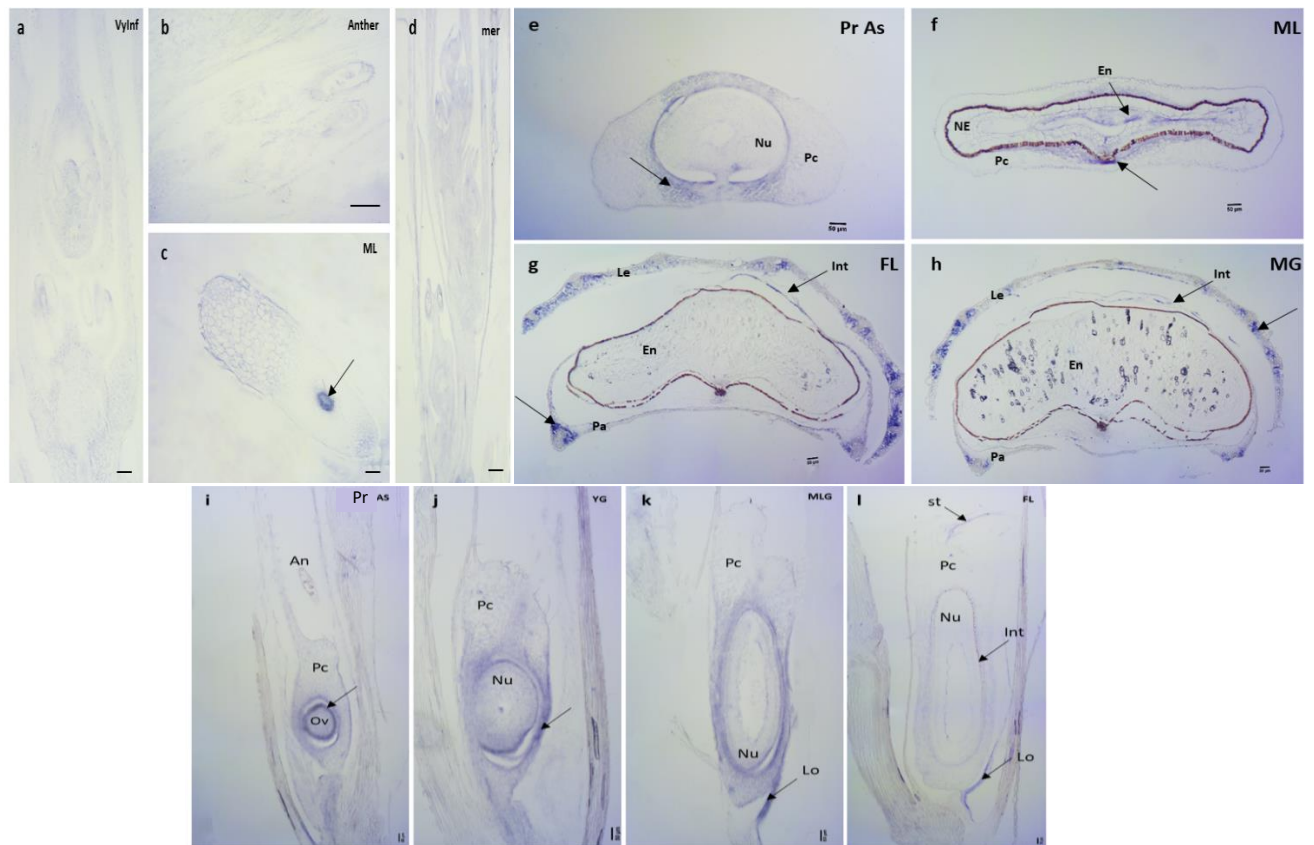


Figure 3-9: The mRNA expression pattern of *BdMADS21*, (a) very young inflorescence (Vy Inf), (b) Pr As mature anther, (c) mid-length grain (ML), (d) spikelet meristem (mer), (e) Pre-anthesis carpel (PrAs), (f) mid-length grain (ML), (g) full-length grain (FL), (h) mature grain (MG), (i-l) longitudinal section in developing Caryopsis, An; anther, Pc; pericarp, Nu; nucellus, lo; lodicule, Int; integument; Ov; ovule, NE; Nucellar epidermis, En; endosperm, Pa; palea, Le; lemma, St: stigma scale bar: 50 µm

The *BdMADS21* mRNA expression was not detected in the spikelet meristem where the floral organs start differentiating (Figure 3-9 (d)), nor very young inflorescence (Figure 3-9 (a)), nor in the anther at pre-anthesis stage (Figure 3-9 (b)). Only a weak signal could be observed in the abaxial and adaxial side of the carpel at the pre-anthesis stage especially in the large chlorenchyma cells this could be seen very clear in mid-length stage around the nucellus (Figure 3-9 (c-k)). This expression persists with a weaker signal in addition to some expression in the developing endosperm during the mid-length stage (Figure 3-9 (f)). No mRNA signal could be detected in the endosperm in full-length and mature grain stage (Figure 3-9 (g-h)), On the other hand, there was an expression in the lemma and palea during these two stages especially in the bundle sheath around the xylem and cortex. The embedded longitudinal section (Figure 3-9 (i-l)) supports the previous results shown in the transverse sections with an additional illustration of the lodicule signal which appears in the mid-length and full-length stages, and a weak signal was observed in the stigma as well (Figure 3-9 (k)).

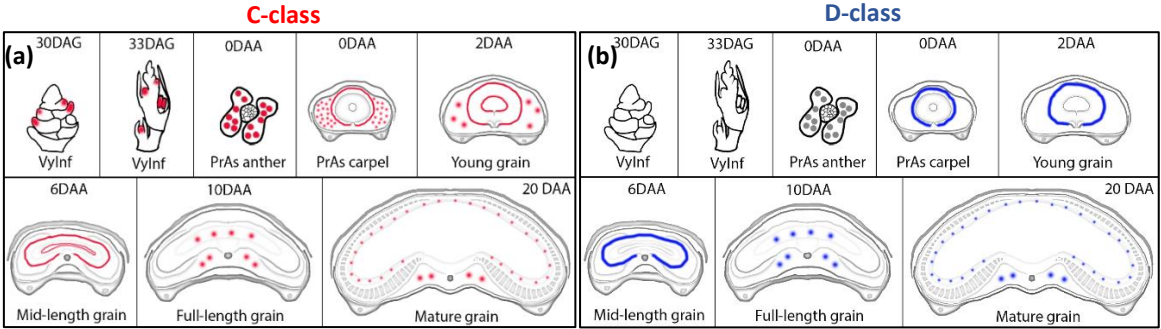
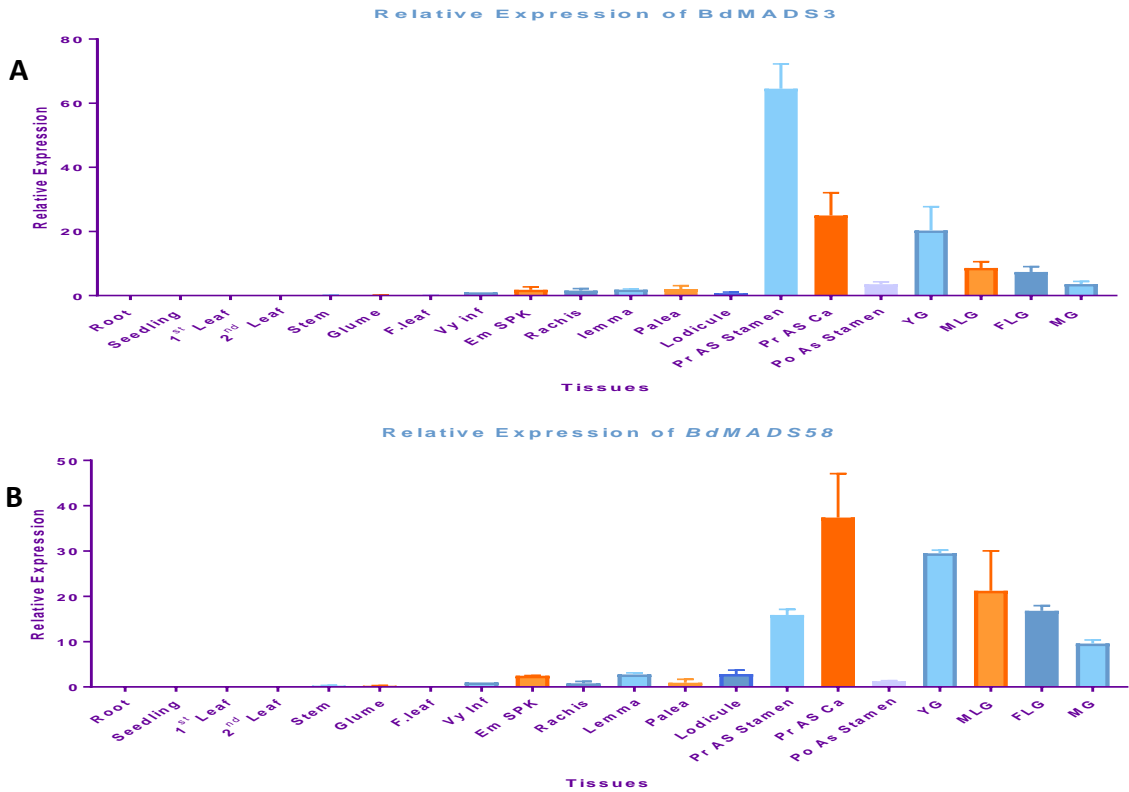


Figure 3-10: General illustration of the special localisation of the mRNA signal of (a) C-class genes and (b) D-class genes. DAG: days after germination, DAA: days after anthesis.

3.4.1.3 Expression pattern analysis for *BdMADS3*, *BdMADS58*, *BdMADS13* and *BdMADS21* in *Brachypodium distachyon* (21) WT using RT-qPCR.

Due to the high sensitivity of qPCR, precise quantification of mRNA levels is necessary; therefore, highly stable expressed reference genes were chosen, because, genes with high variation between developing stages are more likely to introduce false results. A comparative test was achieved on three different housekeeping genes; glyceraldehyde-3P-dehydrogenase (*BdGAPDH*), ubiquitin (*BdUBI18*) and actin (*ACT*) on a wide range of plant developing stages tissues, *BdACT7* showed high stability in the Ct values across all growth stages. The stability was calculated by NormFinder software.



different in between samples, $F(19, 20) = 27.9$, and $P < 0.0001$. Error bars indicate \pm SD ($n = 3$). Flag leaf (F. leaf), Very young inflorescence (Vylnf), Emerging spikelet (Em SPK), pre-anthesis (PrAs), carpel (Ca), post-anthesis (PoAs) young grain (YG), mid-length grain (MLG), full-length grain (FLG), mature grain (MG).

The RT-qPCR results illustrate the expression patterns of *BdMADS3* and *BdMADS58* in different development stages (Figure 3-11), from the root at three days after germination to mature grain 78 DAG. The relative expression patterns between the two genes were almost similar and followed approximatively the same patterns. The very young inflorescence was used as a calibrator (fold change = 1) since it is the first organ to show expression. The qRT-PCR detected no prominent expression of *BdMADS3* and 58 in the early stages and vegetative organs (root, seedling, 1st leaf, 2nd leaf, stem, glume and flag leaf) (Figure 3-11). The expression of the two paralogues is first detected in the very young inflorescence stage where the floral organs begin to differentiate. (Yamaguchi *et al.*, 2006) indicates that the expression of the rice *OsMADS3* is detectable in stamen primordia when the lemma and palea primordia initiate, this was similar to *Brachypodium*. The expression peaks in the stamen pre-anthesis with 60-fold change and around 20 fold change in *BdMADS3* and *BdMADS58* respectively, also in the carpel when the full floret forms prior to anthesis marked approximately 25-fold change in *BdMADS3* and just less than 40-fold change for *BdMADS58*, with a noticeable presence of the mRNA of *BdMADS58* in the lodicule as well. The expression dropped and almost disappeared in the stamen after anthesis where the mature anther is open, and the pollen grains are released (Figure 3-11). However, it persists in the young grain with a steady decrease during mid-length and full length to its lowest level at mature grain This trend was matched in both paralogues with different levels of relative expression.

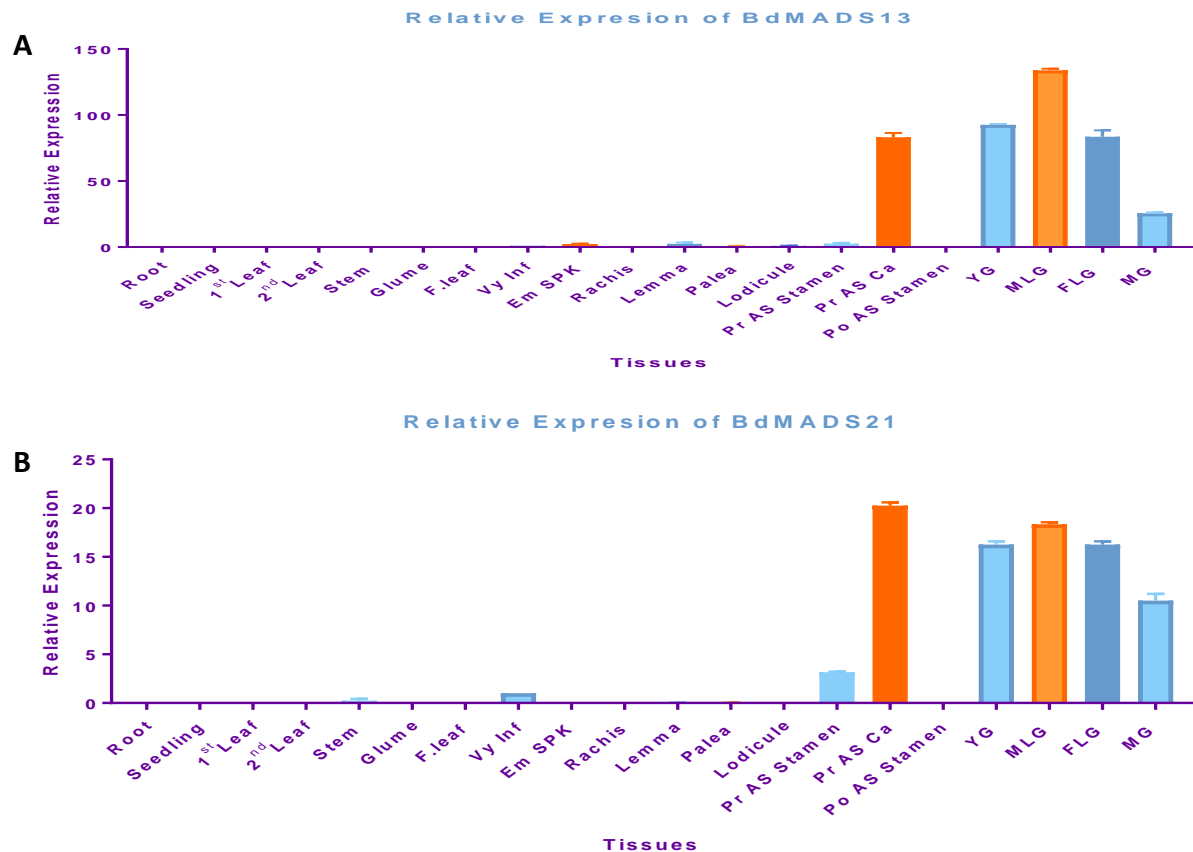


Figure 3-12: The relative expression of *Brachypodium MADS13*(A) and *MADS21*(B) mRNA using RT-qPCR, (Vy Inf is calibrator), the expression was calculated using the delta-delta Ct method, one-way analysis of variance (ANOVA) was used, there was a significant expression difference of *BdMADS13* in between samples $F(19, 20) = 169$, $P < 0.0001$. Similarly, the expression of *BdMADS21* with an $F(19, 20) = 224270$ and $P < 0.0001$. Error bars indicate \pm SD ($n = 3$). Flag leaf (F.leaf), Very young inflorescence (VyInf), Emerging spikelet (Em SPK), pre-anthesis (PrAs), carpel (Ca), post-anthesis (PoAs) young grain (YG), mid-length grain (MLG), full-length grain (FLG), mature grain (MG).

The two members of the D class lineage (Figure 3-12) show similar expression patterns; however, there was no detectable presence of *BdMADS13* and 21 mRNA in the vegetative tissues even in the stage of inflorescence formation, emerging spike, lemma, palea and lodicule. At the pre-anthesis stage, the stamen shows expression for both paralogues at different levels. The expression peaks in the carpel prior to anthesis, particularly high for *BdMADS21*, then decreases gradually after anthesis in the young grain, mid-length, full-length, to reach around 340-fold for *BdMADS21* in the mature grain. However, *BdMADS13* fluctuates little bit with a small peak during full-length then decreases at maturity. Remarkably, the *BdMADS13* mRNA relative expression is the highest among all four genes (*BdMADS3*, 58, 13 and 21) (Figure 3-12) and also it peaks in MLG stage showing its importance during grain filling.

3.4.1.4 Transcriptome analysis: the expression profile of all *MADS-box* transcription factor in *Brachypodium*

After investigating the expression of our target genes with RT-PCR and *in situ* levels. An RNA-Seq transcriptome (Array Express: E-MTAB-7607) was used in order to extract the expression data for these genes. Eight stages across the life cycle of the plant were used in generating the transcriptome (Pre-anthesis ovary Pr As Ov 0DAA, Young grain YG 2DAA, Mid-length grain ML 6DAA, Full-length grain FL 10 DAA, Mature grain MG 20 DAA, Germination Germ, and Seedling SDL) to give us a complete picture about the pattern of the expression of these genes.

In the RNA-Seq data set, we extracted the expression of only the type II *MADS*-MEF2-Like genes, which contains the four domains (MIKC) giving us 40 genes (Figure 3-13). These genes were selected using the InterPro scan for the identifying domains (IPR033896).

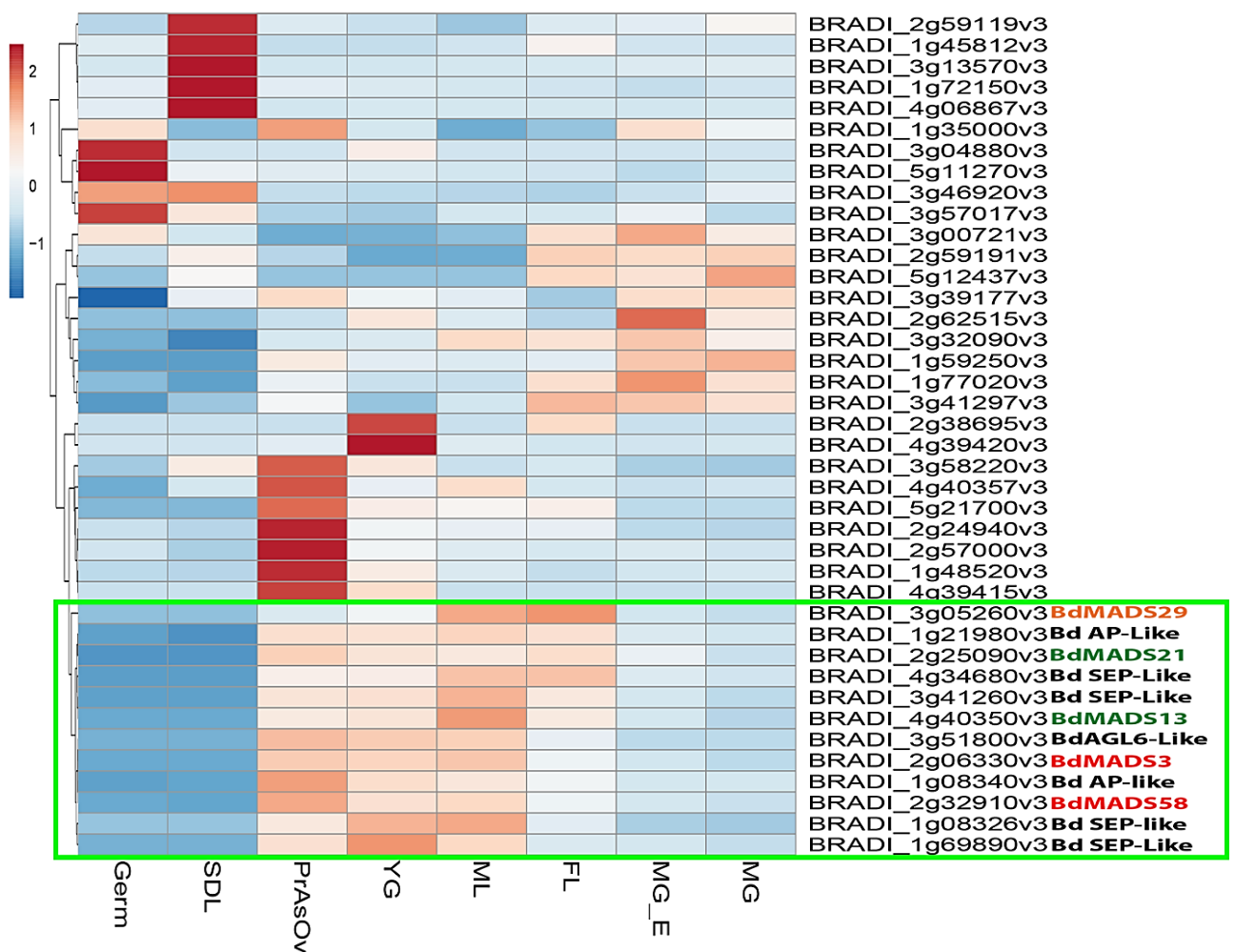


Figure 3-13: Heat map illustrating all *MADS* MIKC-type in *Brachypodium* from eight different tissues samples, RNAseq DESeq2 normalised counts visualised with (<https://biit.cs.ut.ee/clustvis>), the correlation was used for clustering distance, and clustering method was average. Germ: germination, SDL: seedling, PrAsOv: Pre-anthesis ovary, YG: young grain, FL: full-length grain, ML: Mid-length grain, MG-E: mature grain

without the embryo, MG: mature grain, (Yusoff, Kourmetli et al, unpublished) (Array Express: E-MTAB-7607).

In our lab transcriptome (Yusoff et al, unpublished), eight developmental key stages have been chosen to investigate RNA expression. These stages start from germinating seed until mature grain. In the heat-map generated (Figure 3-13) we could observe that the majority of *MADS-box* genes are highly transcribed in reproductive organs and only five were specific to the seedling during the early stage of the plant cycle (vegetative tissues).

These results illustrate the expression of our candidate genes from C and D-lineage, which were clustered (Figure 3-13 green rectangle). Their expression was correlated and focused in the pre-anthesis stage up to full-length grain. It also co-expressed with some specific genes that belong to *SEP-like* (E-lineage) and *AP3-like* (B-lineage) families which suggest that these transcription factors interact together to form the quartet in order to bind to the CArG-box motif in the downstream genes. Yeast two-hybrid assays showed that both *BdMADS13* (*BdMADS2*) and *BdMADS21* (*BdMADS4*) could interact with *BdSEP3* (Bradi_4g34680v3), but *BdMADS2* can additionally interact with *BdAP1* (Wei *et al.*, 2013), these genes are clustered with *BsMADS3* and *BdMADS58* (Figure 3-13 green rectangle). In addition, there was *BdMADS29* gene that highly expressed during mid-length and full-length grain. This gene is investigated in chapter 6. *BdMADS13* and 21 gene expression peak at mid-length and full length suggesting their role in the grain filling.

3.5 Discussion

Our analysis using the reference plant *Brachypodium* demonstrates that the C-class and D-class genes expression is highly conserved between tropical grasses (*Oryza sativa*) and temperate grasses (*Triticum aestivum* and *Hordeum vulgare*).

In the current study we revealed that both paralogues *BdMADS3* and *BdMADS58* were shown to be expressed from early inflorescence in the floral primordia at the early stages (Figure 3-4). These results are similar to rice where it was shown that *OsMADS3* and *OsMADS58* are redundantly regulating floral meristem determinacy in rice (Dreni *et al.*, 2011). In subsequent stages, the two paralogues diverge, as the *BdMADS3* showed some specification to the third whorl, while *BdMADS58*, was shown to be more abundant in the

carpel, especially in the mesocarp (Figure 3-7, 3-10). There was no noticeable expression in lemma and palea in the first whorl or the lodicule in the second whorl. After anthesis and as the endosperm cellularises and expands, both genes continue to be expressed with variable levels in various stages of the developing grain. *BdMADS3* shown to be expressed in the integument, nucellar lysate and the lateral side of the mesocarp (Figure 3-6), whereas, the *BdMADS58* was expressed more in the outer integument at mid-length stage (Figure 3-7). In the following maturity stages, the expression of both genes decreases gradually and is focused in the aleurone layers which is known to store large quantities of oils and lipids that are used as a nutrient during seed germination.

In our study, we found that the expression pattern of the two members of the D-class (*BdMADS13* and *BdMADS21*) was slightly different from the C-class gene expression. *BdMADS13* starts expressing in the very young inflorescence where the floral organ primordia start to form. In contrast, its paralogue *BdMDAS21* did not show any evident expression at early stages of flowering as the *in situ* hybridisation showed no expression in the meristem, pre-anthesis carpel or stamen, confirmed with RT-qPCR (Figure 3-11). Both genes, *BdMADS13* and *BdMADS21*, were active and abundant in the pre-anthesis carpel and continue to express throughout seed filling and maturation stages. This expression was localised using *in situ* hybridisation for *BdMADS13* in the endocarp and the vascular bundle (Figure 3-8) then is focused in the mesocarp in young grain, and in the integument and aleurone layers in the later stages.

There are no data reported on the expression localisation in the grain of the C-class and D-class in the monocot species. However, some members of M α -type were shown to be expressed in the peripheral endosperm after cellularisation and involved in endosperm proliferation including, AGL62, AGL40, AGL95 and AGL64 in *Arabidopsis* (Bemer *et al.*, 2010). Two *MADS-box* genes belonging to the D-class lineage, FBP7 and FBP11 in *Petunia*, were shown to be involved in the regulation of ovule organ identity (Colombo *et al.*, 1995; Angenent *et al.*, 1995). These genes present high similarity in sequence to the *Arabidopsis* AGL11/STK gene (Rounsley *et al.*, 1995) and *OsMADS13* in rice which was shown to be required in the same process. The paralogues of the D-class *BdMADS13* and *BdMADS21* genes in maize are *ZAG2* and *ZMM1* respectively. It was reported that *ZAG2* is expressed during carpel formation (Schmidt *et al.*, 1993) and at the endosperm development stages from 10 to 28 days after anthesis (Liu *et al.*, 2015). *ZmMADS1a* was identified and characterised in (Dong *et al.*, 2019) and has high sequence similarity to *ZAG2*. It was found to play a positive regulatory role in the starch biosynthesis pathway by positively regulating several starch biosynthesis-related genes which enhance the sugar content.

It has been suggested that ancient gene duplications were critical in shaping the evolution of the subfamily. Before the divergence of extant angiosperms (*Amborella* and *Nymphaeales*) (Xi *et al.*, 2014), a duplication event produced the ovule-specific D-lineage and the well-characterised C-lineage, whose members typically promote stamen and carpel identity as well as floral meristem determinacy. In our study, it appears that subsequent duplications in the C-lineage resulted in independent instances of paralog sub-functionalisation (*MADS3* and *MADS58*) and maintained functional redundancy. It has been stated that the increase in morphological diversity is due to the multiple sub-functionalisation events that have occurred in this subfamily.

Our analyses show that the expression of the C-class and D-class genes extends beyond the flowering stage in different stages of the grain development. Therefore, our results suggest that these genes could play a significant role in defining the final size, shape and composition of the grain in grasses.

3.6 Conclusion

Our survey of expression shows that C-class genes, *BdMADS3* and *BdMADS58*, were active in the floral meristem and organ primordia. This expression persists in the mesocarp and integument and during caryopsis growth, the expression focuses more in the aleurone layers (peripheral endosperm). Similarly, the D-class genes *BdMADS13* and *BdMADS21* were expressed in maternal tissues and in grain development with no expression detected in the emerging floral meristem or stamens.

To conclude, the spatial and temporal distribution of the C-class and D-class genes demonstrates that they are acting beyond flowering initiation and might have a role in during grain development. Especially, in the mesocarp during the degeneration, in the endosperm during cellularisation and in the aleurone layer during enzyme storage. These findings provide new insights into the genetic and molecular control of the formation of floral organs and grain development.

Chapter 4

Functional analysis of *BdMADS3* and *BdMADS13* in
Brachypodium

Chapter 4. Functional analysis of *BdMADS3* and *BdMADS13* in *Brachypodium*

4.1 Abstract

MADS-box genes are core members of the ABCDE model for flower development, where C-lineage genes are specifically involved in the development of stamen and carpel, respectively while the D-lineage genes play essential roles in ovule identity determination. We report here the cloning and functional characterisation of two *MADS-box* genes, *BdMADS3* from C-lineage and the *BdMADS13* from D-lineage from *Brachypodium distachyon* using RNAi gene-silencing. Gene-silencing of *BdMADS3* caused an alteration in floral spikelet structure, male sterility and delay in flowering time, while the RNAi lines of *Bdmads13* showed a reduction in the floret number in the spikelet, anther wall malformation and sterility. Quantitative real-time PCR revealed a noticeable decrease in the expression levels in the RNAi lines, especially in the floral organs of *Brachypodium* for both genes. These phenotypes are suggested that *AGAMOUS* genes are responsible for a broad spectrum of floral biology besides the floral organ identity and meristem determination previously found in rice and *Arabidopsis*. Moreover, the mutation caused male sterility that is a beneficial trait for hybrid vigor utilisation and hybrid crops seed production.

4.2 Introduction

The elegant ABCDE model categorises the homeotic genes based on their function in forming the four flower layers of ring-shaped regions called whorls. In rice generally, the E-lineage genes act together with all genes of the model to specify the floret whorls (Li Hu *et al.*, 2011). A-lineage genes are specifically involved in the development of the perianth organs (lemma and palea) in addition to specifying meristem identity (Wu *et al.*, 2017; Kobayashi *et al.*, 2012), whereas lodicules are identified by the combined activities of A-lineage and B-lineage genes in the second whorl of the floret. The stamens (Zhang & Wilson, 2009), located in the third whorl are specified by the combination of B- and C- lineage genes. The latter appear to identify the carpel, while the D-class proteins specify ovule identity (Yoshida & Nagato, 2011).

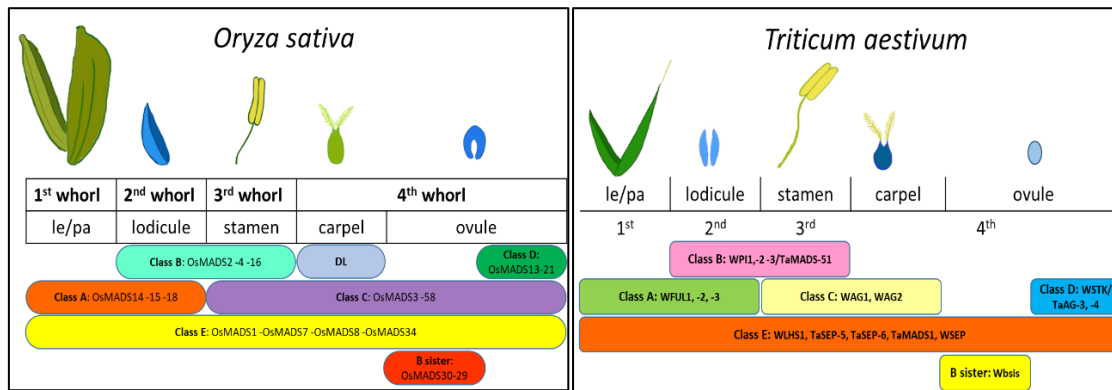


Figure 4-1: ABCDE floral model suggested for rice and wheat, the model was adopted from (Kater *et al.*, 2006) for wheat, the model was adopted from (Callens *et al.*, 2018).

According to this model in *Arabidopsis*, the specification of the stamens in the third whorl is controlled by genes from class C, from *AGAMOUS* (AG) family (Yanofsky *et al.*, 1990) and class B together with genes from class E. An *ag* flower develops petals in the third whorl replacing the stamens, in addition to floral replication in the fourth whorl instead of a carpel (Bowman *et al.*, 1991). Another members of *AGAMOUS* are *SHATTERPROOF1* (SHP1) and *SHP2* genes that found be involved in controlling carpel identity (Pinyopich *et al.*, 2003) and also in carpel dehiscence zone formation (Liljegren *et al.*, 2000).

Similarly in rice (*Oryza sativa*), the third world is specified by genes from three classes B-, C- and E. *OsMADS4* from B-class is needed to determine the identity of stamens (Kang *et al.*, 1998), in addition to *OsMADS3* and *OsMADS58* which show redundancy in stamen specification (Yamaguchi *et al.*, 2006). *OsMADS3* is essentially expressed in stamen, carpel, and ovule primordia, then its expression decreases when these organs differentiate. However, *OsMADS58* is expressed in stamen, carpel and ovule primordia at early floral stages, then its activity remains during the differentiation and maturation of these organs, suggesting that *MADS58* has a dominant role in carpel morphogenesis (Yamaguchi *et al.*, 2006). Together, *OsMADS3* and 58 have functional similarity to AG.. On the other hand, there is apparent functional diversification between these paralogs, giving them predominant functions in different whorls.

OsMADS3 has a stronger role in repressing lodicule development in whorl two and in specifying stamen identity, whereas *OsMADS58* contributes more to floral meristem determinacy and to regulate carpel morphogenesis (Kater *et al.*, 2006). In addition, E-class

genes, represented by *OsMADS7* and *OsMADS8*, were found to be required for stamen specification in the third whorls (Cui *et al.*, 2010).

In rice, three B class genes, two C class genes, and five E class *SEP-like MADS*-box genes have been reported to be involved in anther development (Hu *et al.*, 2011; Yamaguchi *et al.*, 2006; Yao *et al.*, 2008; Ciaffi *et al.*, 2011; Li *et al.*, 2011; Khanday *et al.*, 2013). Specifically, the floral homeotic gene from C-class, *OsMADS3*, that is expressed in stamen and carpel primordia during early floral development, is essential for stamen specification (Kang *et al.*, 1998; Yamaguchi *et al.*, 2006).

The fourth whorl is specified by a complex of genes from C-, D- and E class (Honma & Goto, 2001). In *Arabidopsis*, the class D gene *STK* is restrictedly expressed in the ovule, however, all four members of the *AGAMOUS* family, *STK*, *SHP1*, *SHP2*, and *AG* are involved in specifying ovule identity. In rice, two members of D-class have been identified, namely *OsMADS13* (Lopez-Dee *et al.*, 1999) and *OsMADS21* (Lee *et al.*, 2003). *OsMADS13*, but not its paralogue *OsMADS21*, was found to be involved in the specification of ovule identity from the formation of ectopic carpels without ovules in *Osmads13* mutants (Dreni *et al.*, 2007) showing that it affects the meristem determinacy as well.

According to the ABCDE model (Figure 4-1), *DROOPING LEAF* (*DL*) (Nagasawa *et al.*, 2003; Yamaguchi *et al.*, 2004) an orthologue of the *Arabidopsis* gene *CRABS CLAW* (*CRC*) contributes to the specification of carpel identity together with C and E class. It also controls floral meristem determinacy, and the antagonistic regulation with class B genes (Kater *et al.*, 2006).

Another group of genes called B-sister, phylogenetically related to the B-class genes, was found to be mainly expressed in ovules (Becker *et al.*, 2002). Two *Arabidopsis* B-sister *MADS*-box proteins were characterised, *TT16*, that is necessary for inner integument differentiation and controls seed coat development (Nesi *et al.*, 2002) and *GOA* which is expressed in ovule primordia and the outer integument and also controls organ size via cell expansion (Prasad *et al.*, 2010). In rice, B-sister clade has three members, *OsMADS29*, 30 and 31. *OsMADS30* was found to be expressed in all organs in the plant (Yang *et al.*, 2012), while, *OsMADS29* seems to positively regulate cysteine proteases and *OsVPE1* genes that are involved in programmed cell death (PCD) by directly binding to the *CArG*-box motif in their promoter (Yin & Xue, 2012; Yang *et al.*, 2012).

To date, no functional study is available on *Brachypodium MADS-box* genes. There is only overexpression of two genes (*BdMADS2*, *BdMADS4*) in *Arabidopsis* that caused curly rosette leaves, small sepals and petals, and early flowering (Wei *et al.*, 2013). Our investigation provides a functional analysis of one member of C-lineage (*BdMADS3*) and one from D-lineage (*BdMADS13*) in *Brachypodium* using RNA interference lines. Our study Compare the genetic regulatory mechanism between wild type and *BdMADS3*, *BdMADS13* transgenic lines by phenotype and genotyping in order to establish an understanding of the role of these genes expression.

As figure 4-1 illustrates, these genes; *BdMDAS3* and *BdMADS13* are express in the male and female organs respectively in grass floret, therefore they were chosen to be studied for each class in this project.

4.3 Material and methods

4.3.1 Transgenic materials and growth conditions

In this study we used *Brachypodium distachyon* (diploid inbred line Bd21) obtained from Dr Sinead's lab, as wild type reference. The seeds of *Brachypodium distachyon Bdmads3* and *Bdmads13* RNA interference lines were kindly provided by Dr Matthew Moscou's lab at the Sainsbury Lab, Norwich. The constructs were made previously in Dr Sinead's Lab using a complementary sequence of the target transcript (Figure 2-4), then transferred to *Agrobacterium tumefaciens* and send to Dr Moscou's lab to transform *Brachypodium distachyon* (inbred line Bd21) callus culture and develop into a full plant with seeds following the protocol described in (Vogel & Hill, 2008). The seeds were de-husked and vernalised on moist filter paper in a 9mmØ Petri dish at 5°C for 48h in order to break the dormancy and stimulate flowering, then moved to room temperature for 48h. The germinated seedlings were then sown in a soil mixture of 2:1:1 (compost: vermiculite: sand) supplied by Levington® then transferred to a phytotron (controlled environment) under the following conditions: photosynthetic photon flux density (PPFD) of 400 µmol m⁻² s⁻¹, 20 h light/4 h darkness, with a day/ night temperature of 24°C under 52% relative humidity. Plants were watered by flooding under tray every other two days.

4.3.2 Plant sampling

The RNAi lines were planted at the same time under the same growth conditions with the wild type lines. Genomic DNA was collected from 10 lines from *Bdmads3* RNAi for

screening. All 150 plants were carrying the construct; however, only one line showed a severe phenotype. The line 2: BDT_00104 was called *Bdmads3-2*.

Similarly, *Bdmads13* RNAi lines were germinated and screened for the construct,

Line 1: BDT_00181 (Callus ID: 12-4) named *Bdmads13-1*, **Line 2:** BDT_00182 (Callus ID: 12-5) named *Bdmads13-2*, **Line 3:** BDT_00183 (Callus ID: 12-6) named *Bdmads13-3*, **Line 4:** BDT_00184 (Callus ID: 14-1) named *Bdmads13-4*, **Line 5:** BDT_00185 (Callus ID: 15-1) named *Bdmads13-5*.

Only three lines (*Bdmads13-2*, *Bdmads13-3*, *Bdmads13-5*) displayed alterations in the flowering time, height and number of tillers with complete sterility. Plant samples were manually collected from different flowering stages, and then total RNA was collected for cDNA synthesis.

4.3.3 Characterisation of the transgenic phenotypes

A drastic phenotype was observed in one out of ten lines germinated for F2 from *Bdmads3* RNAi lines (*Bdmads3-2*). *Bdmads13* RNAi showed phenotype in plants in all the five lines grown, but only severe phenotypes were selected (*Bdmads13-2*, *Bdmads13-3* and *Bdmads13-5*). The plants were measured and photographed using NIKON Coolpix digital camera.

4.3.4 Morphological analysis

Photos were taken at corresponding stages of the life cycle of the wild type and the RNAi lines in order to compare the morphological differences between the lines, pictures were taken using Surface Pro 4, Nikon Coolpix and dissection microscope camera; GXCAM (GT vision Ltd), with Zeiss Stereo Microscope 47-50-57, the samples images were previewed and saved using GX Capture 7 software.

4.3.5 Expression clone and genotyping

4.3.5.1 *Bdmads3* and *Bdmads13* RNAi lines expression clone map

Reverse genetics was used to silence *BdMADS3* and *BdMADS13* genes (post-transcriptional gene silencing). Constructs containing *BdMADS3* and *BdMADS13* cDNA in the pIpKb007 vectors were made and sent to Dr Matthew Moscou's lab at the Sainsbury Lab, Norwich for transformation.

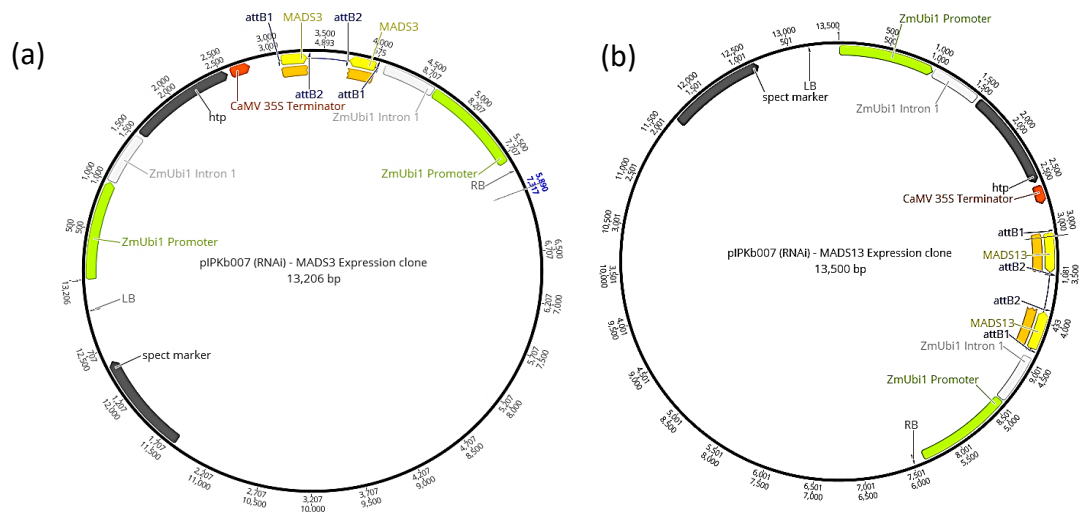


Figure 4-2: The pIPK007 RNAi expression clone generated using GATEWAY® technology. (a) *BdMADS3* cDNA sequence is cloned in two opposite orientations between the attB1 and attB2 recombination sites, resulting in the production of complementary transcripts that can hybridise generating double-stranded RNA. (b) Similarly, *BdMADS13* cDNA sequence was cloned in pIPK007 RNAi expression.

Two expression clones (Figure 4-2) were prepared, one for each gene (*BdMADS3* and *BdMADS13*), by recombination using the Gateway technology (Invitrogen). Each pIPKb007 vector contains two expression units, the first unit contains a hygromycin selectable marker gene driven by the maize Ubiquitin1 promoter, and a bacterial streptomycin resistance gene is also present in the vector backbone. The second expression unit contains the gene-specific cDNA sequence (259 bp) for *BdMADS3* and 406 bp from *BdMADS13* target genes under the control of ZmUbi1 Intron 1 and ZmUbi1 promoter. The chimeric molecule was introduced to *Brachypodium* plants using the Agrobacterium-mediated transformation method. The transcription of target gene inverted sequences results in forming a dsRNA stem with a single-stranded terminal loop that triggers the degradation of the plant mRNA. The Genotyping was performed by Screening for the plants that carry the pIPKb027 by amplifying the Hygromycin region using Hpt2 primers. The sequence was confirmed by sequencing and only the plants that carried the construct and showed the most severe alterations had been selected for further studies.

4.3.6 Reverse Transcriptase PCR Analysis

Primers sequences are listed in the appendix (table S1) and the method in chapter 3 material and methods section.

4.3.7 Quantitative reverse transcription PCR Analysis

The pRT-PCR primers were designed to amplify inside of the region that has been used in the RNAi construct. Sequences of the primers are in the appendix (table S3), and the

method is described in chapter 3 material and methods section. Corresponding stages were compared using ΔC_t method.

4.3.8 Scanning electron microscopy (SEM)

Freshly collected samples (meristem and floret) were fixed in FAA (3.7% formaldehyde, 5% acetic acid and 50% ethanol) and vacuumed under pressure of 400 mbar for 5 min, then put in increasing concentration of ethanol for dehydration. Two washes of 100% analytical grade ethanol was performed before going through two critical point drying with liquid CO₂ using a Balzers CPD030 critical point dryer. After that, the samples were mounted onto 13mm aluminium stubs (Agar Scientific Ltd) using sticky carbon tabs and coated with palladium in a Quorum Q150 TES coating unit. Samples were viewed on a Hitachi S3000H Scanning Electron Microscope with an accelerating voltage of 10kV, and the photos were saved on BMP file format. The process was performed in the Electron Microscopy Facility at university of Leicester by the help of Natalie Allcock and Ania Straatman-Iwanowska.

4.3.9 Light microscopy and histological analysis

4.3.9.1 Cytochemical staining with toluidine blue

The samples were collected, preserved and embedded as described in chapter 3 in section 3.3.4.1 and 3.3.4.2, then they were sectioned 10 μ m thick and mounted on permanently positive charged slides then dewaxed with two washes of histo-clear (National diagnostics) for 10 minutes. The slides were rehydrated using decreasing ethanol dilutions (%100, %80 and %50) for 5 minutes each then were deep for 1 minute in 0.05% Toluidine blue in 0.1M phosphate buffer, pH6.8, and rinsed with deionised water. The slides were mounted (Merck's Entellan® Mounting Media-1.07961.0100) after air-dried and observed under a compound microscope (GX Microscopes, L3200B). Images were taken using a camera (Euromex CMEX 5.0MP) and previewed with software (ImageFocus 4 v2.9) then saved as tiff file format for high resolution.

4.3.9.2 Pollen grain viability analysis with Alexander staining

Anthers were collected freshly from the floret buds at maturity stage then disrupted using a dissecting needle and gently opened on a microscope slide using a dissecting microscope (Motic SMZ-168 Microscopes). Drops of freshly prepared Carnoy's fixative solution (6 part alcohol, 3 part chloroform and 1 part acetic acid) were put on the pollen grains 3 times for 5 minutes. 10 μ l of Alexander staining solution (5ml 95% alcohol, 0.5ml 1% (in EtOH)

Malachite green, 25 ml distilled water 12.5ml glycerol, 1% (in H₂O) Acid fuchsin, 0.25ml 1%(in H₂O) orange G, 2ml glacial acetic acid and 2.25 ml H₂O up to 50ml final volume) (Peterson *et al.*, 2010) was put on the pollen then covered with coverslip and observed under compound microscope (GX Microscopes, L3200B).

4.3.9.3 Fluorescent stain with DAPI

Paraformaldehyde 4% (w/v) fixative was used to fix freshly collected pollens, which were released from the anthers with dissecting needle then transferred in Eppendorf tube and left in the fixative solution overnight to clear the cells. The pollen grains were collected by brief centrifuging and pipetting out ~10µl and put on a microscope slide to let dry for 3 minutes before applying the fluorescent stain. DAPI (4', 6-diamidino-2-phenylindole) staining solution (Park *et al.*, 1998) was used in order to label the nucleus of the pollen grains. The DAPI stock solution was made by dissolving DAPI powder (Sigma) in distilled water (0.4mg/ml DAPI) and GUS buffer was made following (0.1 M sodium phosphate pH 7.0, 1 mM EDTA, 0.1% (v/v) Triton X-100). 8µl of DAPI (10 µg/ml) was mixed with 10µl of GUS buffer and applied on fixed pollen grain. 10µl of DAPI stain was applied on the pollen, sealed with nail varnish to prevent the sample from drying, and immediately observed under a fluorescence microscope (Nikon ECLIPSE 80i, Japan). Pollen was imaged using epifluorescence microscopy DAPI with excitation peak: 360/40 and range of 340-380. A mercury lamp was used as the excitation source for the upright microscope Nikon ECLIPSE 80i, which is fitted with an LED illumination system (CoolLED, UK) and Plan Fluor 40x and Plan Apo VC 60x oil immersion objectives. The pictures were taken by Nikon DS-QiMC camera (Y-QT, Japan). Photos were observed, captured and treated using NIS-Elements Basic Research version (BR Ver4.13.05 32bit edition) software in JPEG 200 format then converted to TIFF using the same software.

4.4 Results

The analysis started by a detailed, comprehensive comparison of different floral organs between the wild-type and the silenced lines (*Bdmads3* and *Bdmads13* RNAi lines) in order to understand the effect of the silencing on the plant generally rather than focusing on one organ. Genotyping was performed on all the available RNAi lines, screening for the plant carrying the construct. The results revealed that all the tested lines were accommodating the hygromycin resistance gene that was in the RNAi construct, however, plants from all

the five lines showed severe phenotype in *Bdmads13* RNAi line and one line out of ten in *Bdmads3* RNAi line.

4.4.1 Comparing and contrasting the expression profile between the RNAi lines and the wild-type

4.4.1.1 *BdMADS3* and *BdMADS13* mRNA comparison to wild-type using reverse transcription-polymerase chain reaction (RT-PCR)

It is essential to determine the effectiveness of gene silencing in the RNAi lines so that the phenotypes can be attributed to gene function. In this specific experiment, F2 generation tissues were used with germinating wild type and RNAi line at the same time under the same growth conditions and collecting organs at equivalent stages.

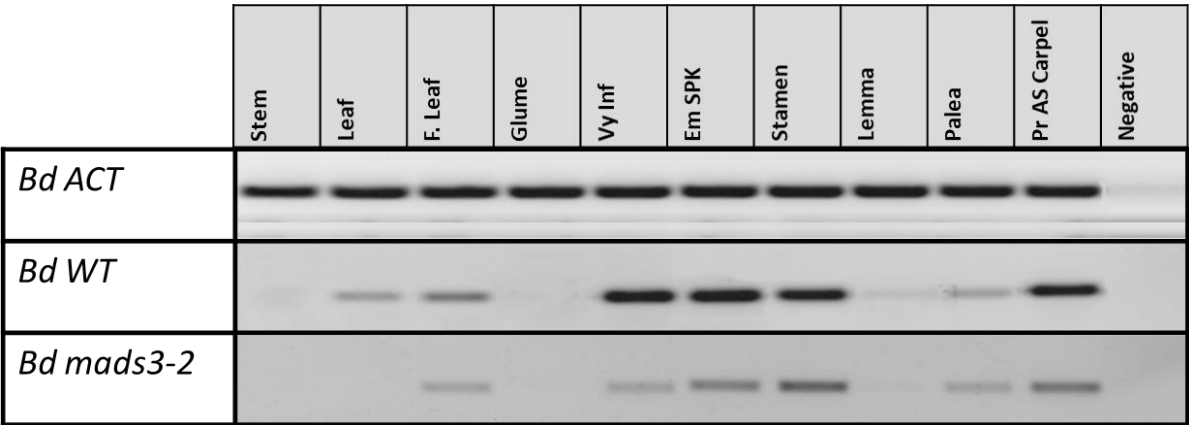


Figure 4-3: RT-PCR comparing the different expression pattern between *Bd21* (WT) (B) and *Bdmads3* RNAi line (A), the intensity of the band on gel reflects the amount of mRNA present in the tissue, 30 cycles.

Taking into consideration that the *Bdmads3* RNAi line presents a delay of the flowering stage of about 10 days (Figure 4-10), the equivalent stages were compared using RT-PCR. The results (Figure 4-3) highlighted that there were some differences in the intensity of the band on the gel reflecting probably the low amount of the *BdMADS3* mRNA in the interference line tissues. For instance, the carpel cDNA amplicon band of the wild type was more intense then the *Bdmads3* RNAi line and the same applies to the stamen, young inflorescence, and emerging spike. The mRNA could not be detected in the stem, glume, and lemma in both lines. However, reduced expression was detected in the RNAi line in the flag leaf and palea. Overall, there was less *BdMADS3* mRNA present in the RNAi tissues compared to the wild type.

Similarly, *Bdmads13* RNAi line had a delay in the flowering of around 20 days (Figure 4-10). Plants from all the five lines displayed severe phenotype were selected for *Bdmads13* RNAi,

samples were collected at equivalent stages in wild-type from different stages and tissues for RNA extraction, they were called *Bdmads13-2* (2C3), *Bdmads13-3* (3B3) and *Bdmads13-5* (5C1) (figure 4-4).

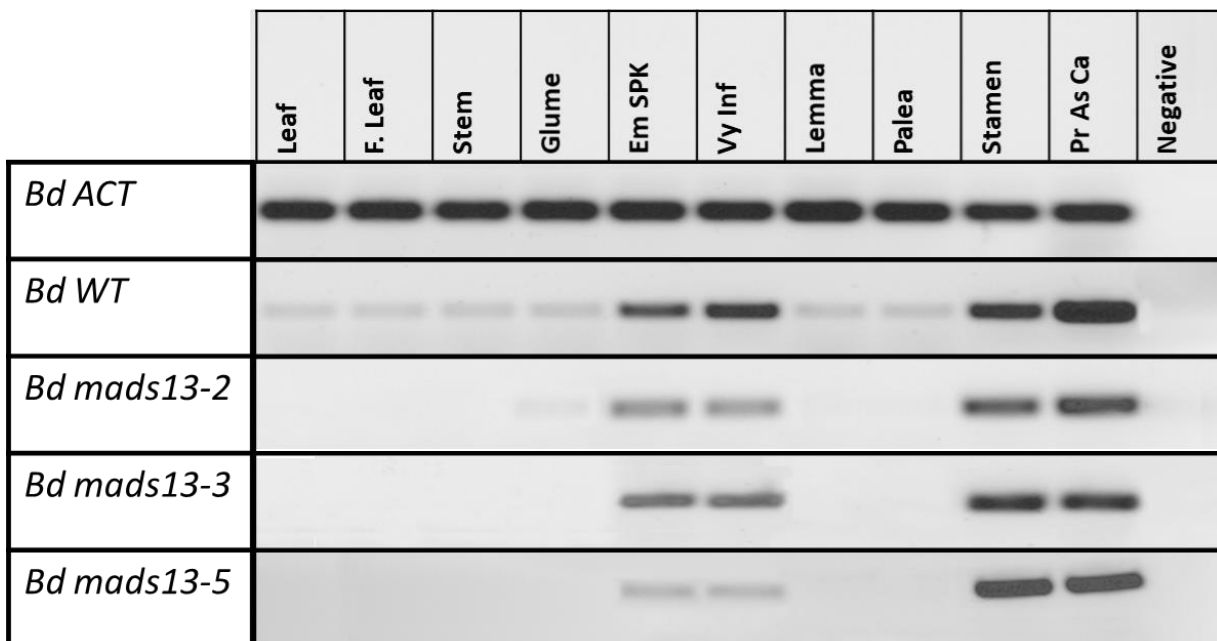


Figure 4-4: RT-PCR comparing the different expression pattern between *Bd21* (WT) and the three *mads13* RNAi lines (*Bdmads13-2*, *Bdmads13-3* and *Bdmads13-5*), the intensity of the band on gel reflects the amount of mRNA present in the tissue, 30 cycles.

RT-PCR in the wild-type showed some traces of *BdMADS13* mRNA in the leaf, flag leaf, stem, glume, lemma and palea while there was strong expression of the *BdMADS13* mRNA starting from emerging spike, increasing equally in the very young inflorescence and stamen until having its peak in the carpel at pre-anthesis stage. On the other hand, *Bdmads13* RNAi line expression could not detect any traces of the gene expression in the vegetative tissues (leaf, flag leaf, stem, and glume), and only reduced expression could be detected in the reproductive tissues comparing to the wild-type, with no noticeable presence of the mRNA in the lemma and palea. Generally, there was a reduction in the expression of *BdMADS13* in the RNAi line.

4.4.1.2 *BdMADS3* and *BdMADS13* mRNA quantitation and comparison using real-time reverse transcription-polymerase chain reaction (RT-qPCR)

With the aim of quantifying the difference in the mRNA expression illustrated by the RT-PCR expression, a comprehensive qRT-PCR examination was performed on corresponding tissue samples comparing the mRNA quantity between wild type and RNAi lines. In this particular experiment, F3 generation tissues were used with germinating wild-type and

RNAi line at the same time and under the same growth conditions and collecting organs at equivalent stages. In this experiment, we aim to validate the RNAi response in the knockdown lines, measure the expression in the wild-type and demonstrate gene silencing after dsRNA ingestion.

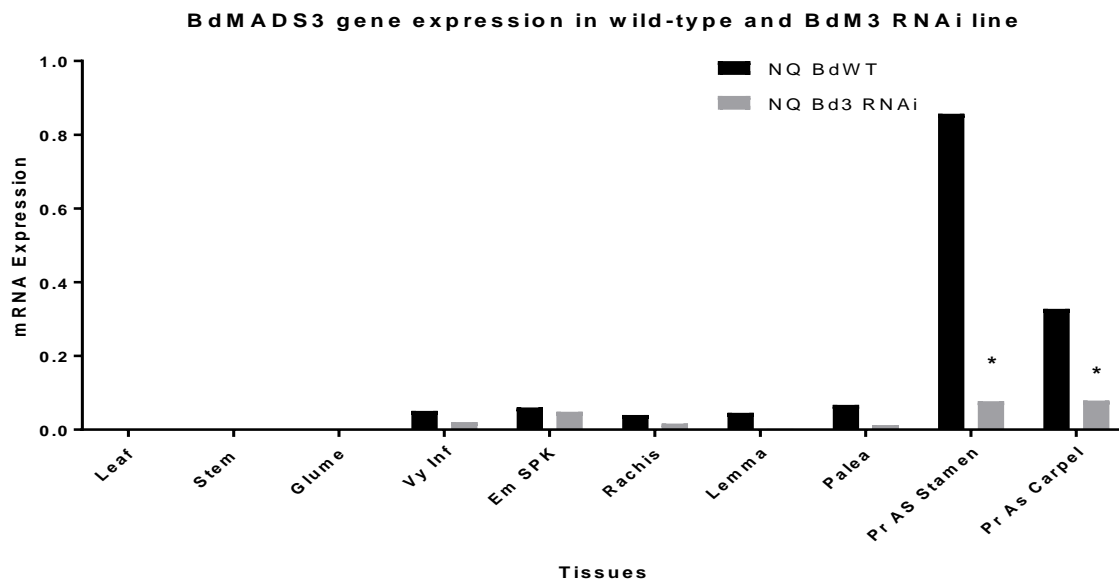


Figure 4-5: mRNA expression of *Bdmads3* RNAi and *Bd21* (WT) using qRT-PCR. Differential expression of *BdMADS3*, showing changed Expression in *Bdmads3* RNAi line compared to the wild-type. Using t test there was a significant difference in the expression of *BdMADS3* in the two lines, P value = 0.0028, t=1.549, df=9 (*denotes significant differences). NQ: normalised quantity.

Bdmads3 RNAi line showed a reduction in expression in all the tissues that have been tested. Nevertheless, stamen and carpel had a remarkable reduction of 0.75-fold change and 0.8-fold change, respectively (Figure 4-5). *BdMADS3* starts expressing at the beginning of the inflorescence stage with small quantities in the spikelet organs (rachis, lemma and palea), however, the expression peaks in the stamen and carpel at pre-anthesis stage. The transgenic line on the other hand, shows silencing of the gene and the difference is very obvious for both stamen and carpel tissues with around -0.8-fold change decrease.

Due to the nature of the *Bdmads3* RNAi line phenotype which has deformed shape of rachis, we could collect it and test for the presence of the mRNA as well, therefore, a great effort was made to collect the wild type rachis in order to quantify the expression, and it is illustrated in the qPCR section (Figure 2-5). Overall, *Bdmads3* mRNA was relatively suppressed, especially in the stamen and carpel.

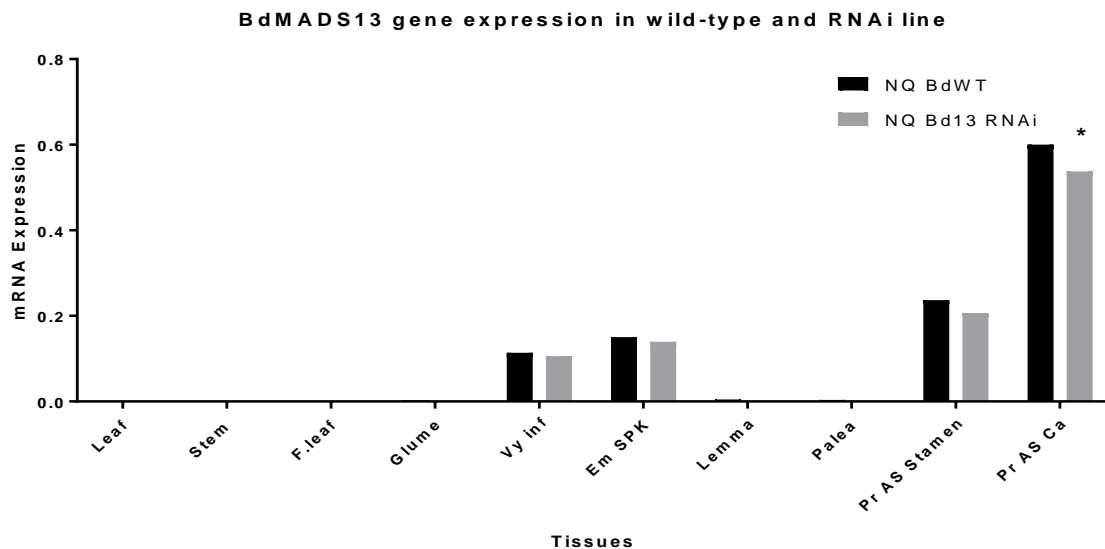


Figure 4-6: Differential expression of *BdMADS13* mRNA in the *Bdmads13-3 RNAi* and *Bd21* (WT) lines using qRT-PCR., showing changed Expression in *Bdmads13* RNAi line compared to the wild-type. Using t test there was a significant difference in the expression of *BdMADS13* in the two lines, P value <0.0001, t=1.983 df=9 (*denotes significant differences). NQ: normalised quantity.

To compare and contrast the expression level of *BdMADS13* mRNA between the wild-type (*Bd21*) and the transgenic line (Figure 4-6). The same stage has been chosen to collect the tissues from (pre-anthesis stage) which are 39 days after seedling for wild type and 59 days after seedling for the transgenic line as it presented a delay in the flowering stage (Figure 4-10) compare to *Bd21*.

We did not calculate the $2^{-\Delta\Delta C_t}$ of the Ct value and calibrate with very young inflorescence; instead, we calculated the quantity present in each tissue for both the housekeeping gene and the target gene (*BdMADS13*), then we normalised it to the housekeeping gene to have the quantity difference.

Figure 4-6 illustrates that there was no presence of the mRNA expression in the vegetative tissues, including leaf, stem, flag leaf and glume in both lines. The *BdMADS13* start expressing from early inflorescence to have its peak in the carpel at the pre-anthesis stage in the wild-type line, similar pattern was observed in the *Bdmads13-3* RNAi line with remarkable reduction in the expression in the very young inflorescence, emerging spike and Pre-anthesis stamen and the pre-anthesis carpel which present the highest difference between lines.

We demonstrated that *BdMADS3* and *BdMADS13* transcript was significantly reduced in the RNAi line, especially in the pre-anthesis stamen and carpel.

4.4.2 Floret and spikelet comparison

4.4.2.1 *Bdmads3* RNAi line floret development and spikelet morphology

Phenotypic alteration of floral organs could be observed in the *Bdmads3* RNAi line plant starting from the young stages where the leaves start to curl around itself in a spiral shape, and the plant was shorter compared to the wild-type. At the flowering phase, the spikelet rachis deforms into a zigzag shape giving the phenotype illustrated in (Figure 4-7 (c)).

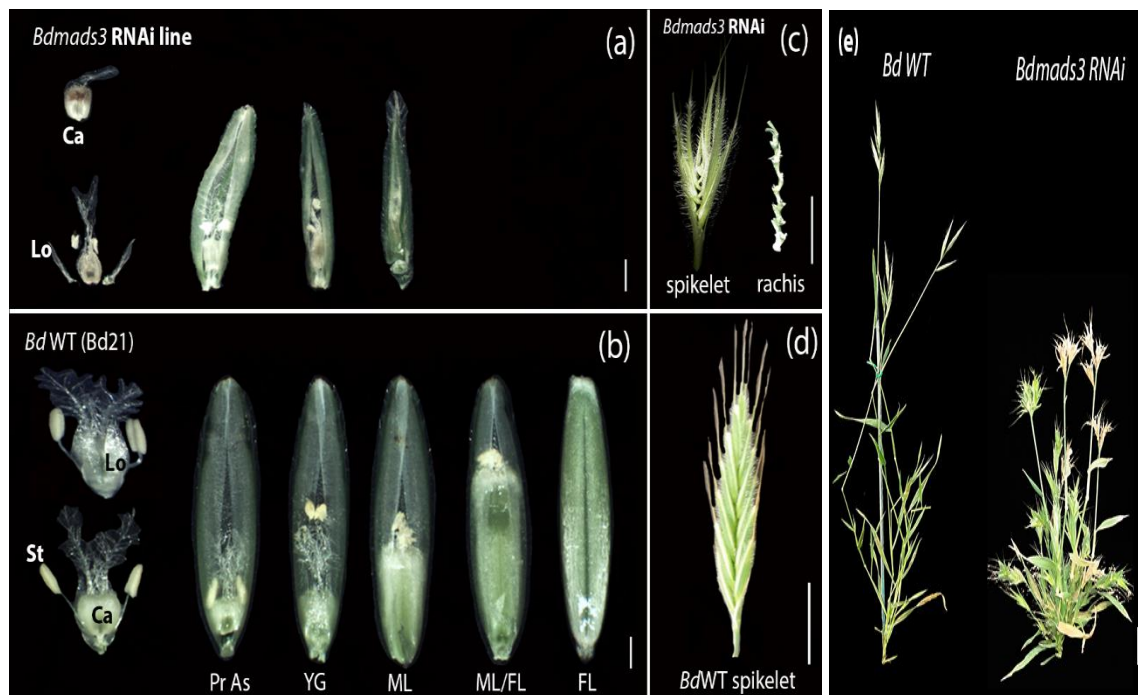


Figure 4-7: *Bdmads3* interference line phenotype in comparison with the wild type *Bd21* in different stages. (a) *Bdmads3* RNAi line floret and spikelet organs, (b) the Brachypodium wild-type floret and grain development, (d) *Bd* wild-type spikelet phenotype, (c) the *Bdmads3* RNAi spikelet phenotype, FL: full-length, ML: mid-length, YG: young grain, Pr As: Pre-anthesis, Ca: carpel, (Lo): lodicule, St: stamen, Scale bars: (c-d) 5 mm; (a-b) 1 mm.

The RNAi lines for *Bdmads3* showed a delay of 10 days in flowering time (Figure 4-10). The phenotypic observation displays a malformation of spikelet and an abnormal form of rachilla (zigzag shape), (Figure 4-7 (c)), which make the spikelet half the size of the wild-type. The glumes, lemma and palea present similar phenotype to the wild-type however they were rigid and bent towards the sides in some florets. The lodicules were long and dehydrated, around double the size of the wild-type (Figure 4-7 (b)). The carpel, stamen and stigma presented a normal phenotype at the pre-anthesis stage then start drying and shrinking at the anthesis stage. The anther did not present any dehiscence at the maturity stage to release the pollen grain.

The most remarkable observation was the absence of anther dehiscence and visible mature pollens which led to male sterility.

In rice *OsMADS3* was shown to be necessary for stamen specification during early flower development in rice using mutation *Osmads3-3* where almost all stamens in whorl three were transformed into lodicule-like organs. Some flowers showed mosaic organs of lodicules and stamens such that the anther-like organ was fused to the lodicule-like organ (Yamaguchi *et al.*, 2006).

A subsequent study (Hu *et al.*, 2011) revealed that *OsMADS3* is also crucial for late anther development and pollen formation by analysing *mads3-4* transgenic which displays abnormal late anther development and male sterile. It has been demonstrated that there was a remarkable dynamic change in the level of reactive oxygen species (ROS) homeostasis through the ROS scavenging protein MT-1-4b which led to abnormal late anther development (Qu *et al.*, 2014).

Our results are consistent with the latter study; it appears that the lack of visible pollen grain in the anthers came as a result of disruption of the function of *BdMADS3* which plays a major component in the development of 3rd whorl mainly male organs.

This phenotype resembles the flowers of transgenic *Arabidopsis* plants expressing antisense AG RNA (Mizukami & Ma, 1995). In this study, it has been shown that antisense AG RNA could lead to a different type of mutation affection mainly the floral meristem determinacy, which may require a higher level of AG activity and floral organ identity. In addition, our study on the grass model *Brachypodium* has been shown that the function of the reproductive organ could be affected by an alteration of the structure as well. different level of disruption can cause different phenotype.

4.4.2.2 *Bdmads13-3* RNAi line floret development and spikelet morphology

In order to elucidate the role of the *BdMADS13* gene in *Brachypodium distachyon*, a detailed phenotypic and cytological analysis of *Bdmads13-3* RNAi lines were performed. These lines (*BdMADS13-2*, *BdMADS13-3* and *BdMADS13-5*) displayed alterations phenotypic changes in flowering time, height, number of tillers and fertility.

There was a delay of 20 days in flowering time for the *Bdmads13 RNAi* lines (Figure 4-10). When the wild-type plant formed a complete spikelet with full structure of florets, the *Bdmads13 RNAi* lines were at booting stage, where the spike starts emerging out from the flag leaf. There was no noticeable difference in the very young inflorescence phenotype, however, as the spikelet started to differentiate its floral organs, the lemma and palea bend slightly away from the rachis axis when the spike is at the level of the flag leaf (Figure 4-8 (b)).

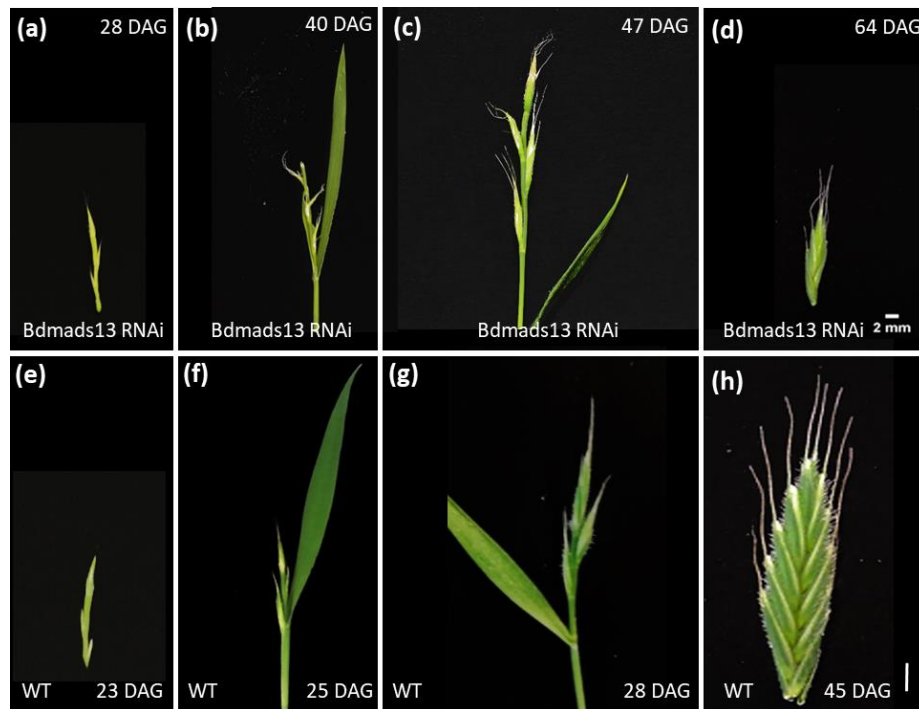


Figure 4-8: *Brachypodium* spikelet developmental stages and phenotypic alteration of *Bdmads13-3 RNAi* spike comparison. (a-d) *Bdmads13-3 RNAi* lines stages from bolting and heading to mature flower formation spikelet with only four degenerated carpels. (e-h) *Brachypodium* wild-type spikelet development stages to full-length grain, DAG: days after germination, Scale bar: 2mm.

The delayed spike of the *Bdmads13-3 RNAi* line starts heading at after germination (DAG). The visible young spikelet awns showed alterations and curve away from the main stem (Figure 4-8 (b)). The subsequent development stages appear normal yet slower than the wild-type. At the maturity stage, the *Bdmads13-3 RNAi* line spikelet presented a shorter spikelet with a small number of florets (about four to five) compared to 7-8 in the wild type (Figure 4-8 (d)). Moreover, after anthesis stages, there was no development of the carpel to mature grain (sterile spikelet).

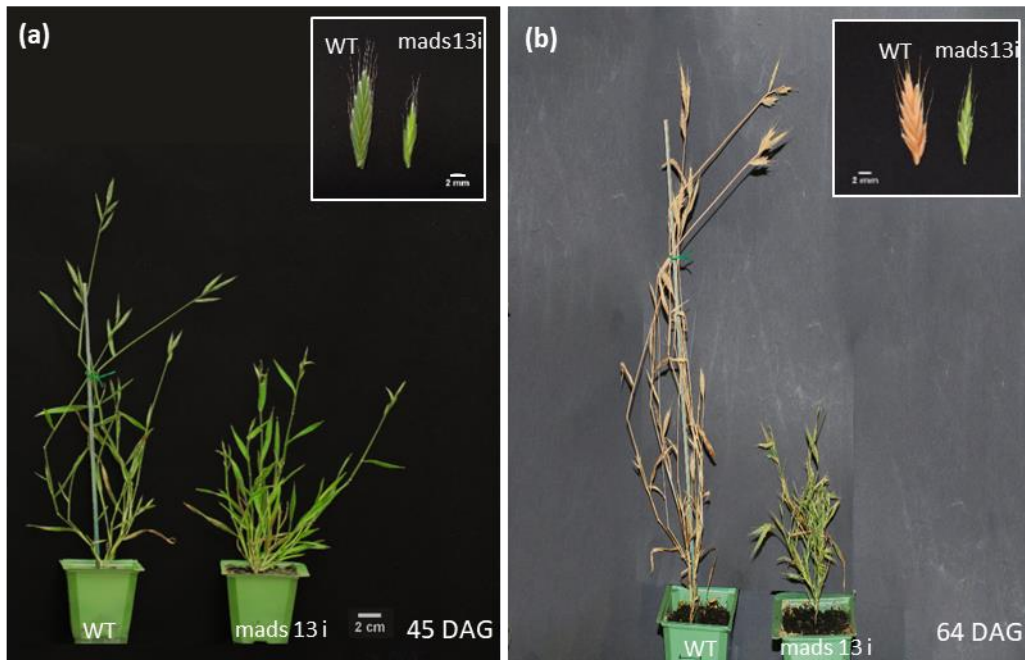


Figure 4-9: Phenotypic comparison of plants from both lines at two developmental stages. (a) *Bd* WT plant and *Bdmads13-3 RNAi* at 45 DAG with a close up of their spikelet (b) *Bd* WT Spikelet and *Bdmads13-3 RNAi* at 64 DAG a close up of their spikelet phenotype. DAG: days after germination, scale bar: 2 mm.

The height of the plant was dramatically reduced by half compared to the wild-type. At 45 DAG the plant had an increasing number of tillers with each developing three to four sterile spikelets. After 64 days after germination, the spikelet undergoes desiccation with undeveloped carpel or aborted grain.

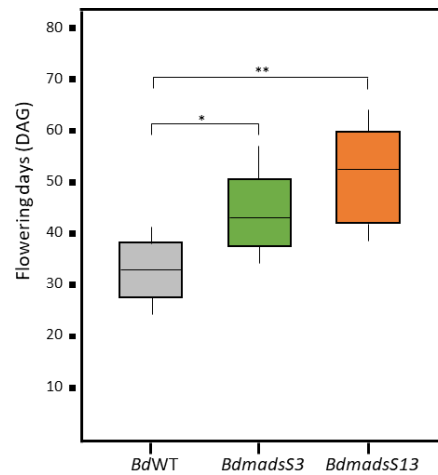


Figure 4-10: Flowering time period in *BdWT*, *Bdmads3-2* and *Bdmads13-3*.

There is a statistical significance in flowering time between the wild type and the RNAi lines. All lines had 48h vernalisation and grew under the same conditions. Flowering time interval start with the first appearance of the very young inflorescence and finish with anthesis, the RNAi lines interval finish with floret desiccation. (*denotes significant differences). DAG: days after germination. Groups were considered statistically different when $P \leq 0.05$.

4.4.3 Spikelet and floret histological analysis of the two RNAi lines

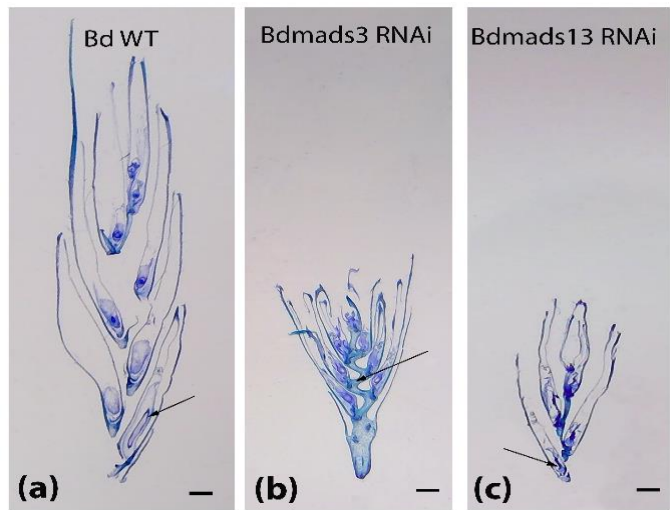


Figure 4-11: *Brachypodium* spikelet at post-anthesis stage under light microscope stained with toluidine blue. (a) *Bd* WT spikelet with developing grains, the arrow is pointing at mid-length stage caryopsis, (b) *Bdmads3-2* RNAi spikelet, the arrow in indicating the malformed rachilla, (c) *Bdmads13-3* RNAi spikelet, the arrow shows the degenerated carpel, scale bar 1 mm.

Figure 4-11 shows the different phenotypes presented by the three lines, *Brachypodium* wild-type, *Bdmads3-2* and *Bdmads13-3* RNAi lines respectively. While the wild type presents a developing grain arranged in alternate-distichous form, the *Bdmads3* RNAi line had a malformed rachis and a similar number of florets with aborted carpels (Figure 4-11 (b)). *Bdmads13-3* RNAi had a normal shaped rachis but a reduced number of florets to around 4 to five only. In addition, the carpel was not successful in reaching mature grain stage (Figure 4-11 (c)).

Carpel and grain development structure:

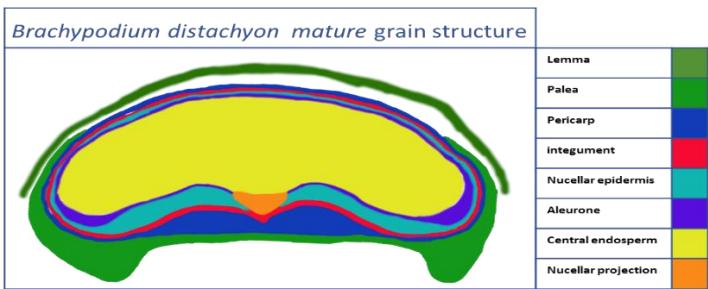


Figure 4-12: Illustration of *Brachypodium distachyon* mature grain transversal section structure.

To determine the cellular basis and morphological differentiation in grain profile between the *Brachypodium* wild-type and the RNAi lines, cytological analysis of carpel and grain development was performed.

In *Brachypodium* wild-type, the tri-cellular free mature pollens are released from the anther and are able to attach to the stigma on the carpel and form tubes to migrate to the ovule. In longitudinal sections it shows the fertilised caryopsis consisting of central nucellar tissue surrounded by a pericarp, divides and elongates rapidly and forms a syncytial endosperm and a circular central vacuole. The endosperm continues to cellularise to fill the cavity in the grain with smaller peripheral aleurone cells visible (Opanowicz *et al.*, 2011) (Figure 4-12).

A comparative analysis between the RNAi line and the wild type revealed remarkable differences in grain shape between the genotypes. Although the early stages are similar, with an unfertilised carpel containing an ovule enclosed by a pericarp tissue, the subsequent stages present noticeable differences. In the wild-type, after anthesis, the synchronous endosperm nuclei proliferate then migrate toward the chalazal region of the embryo sac and, as a result of enlargement of the central vacuole, are contoured at the periphery of the primary endosperm cell (Bennett *et al.*, 1975). The grain elongates and dorsoventrally flattens, and as small aleurone cells form around the periphery of the endosperm the lateral tissues including nucellar tissue become compressed to leave more area for grain filling with starchy endosperm (Figure 4-13 (i-m)).

The *Bdmads3* RNAi carpel appears to have difficulty in dividing and elongating after failing to be fertilised by the pollen. Figure 4-13 (i-h) shows that the differentiated carpel forms normally until the anthesis stage where a dark stained area starts to appear around the ovule. These pigmented strands correspond to the photosynthetic cells (Figure 4-13 (i-h)), these chlorophyll-containing tissues seem to extend to the stigma as well. In the subsequent stages, the carpel deforms and dries out.

Similarly, the *Bdmads13-3* RNAi carpel did not show any observable differences before anthesis, and the ovule was at the basal end surrounded by the pericarp tissues. However, in the subsequent stages, rather than being fertilised, the ovule shrinks and degenerates failing to form any grain. There was no embryo sac developed and it was replaced with abnormally proliferated nucellus in the RNAi lines. To investigate the cause of this further cytological analysis on altered floret organs were performed.

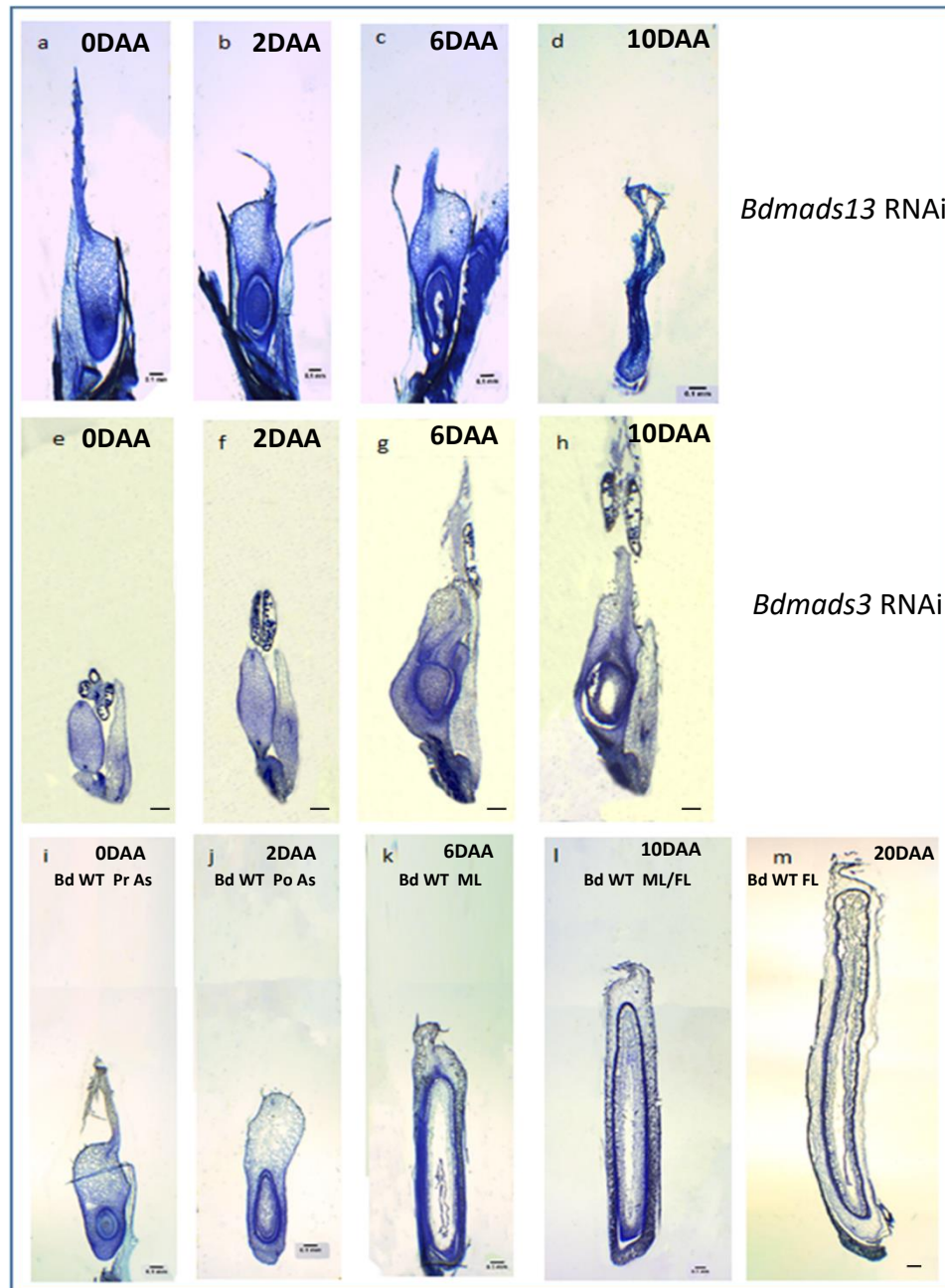


Figure 4-13: Different carpel and grain successive developmental stages compared with corresponding stages in *Bdmads13*, *Bdmads3* RNAi and in *Bd* WT line (longitudinal sections stained with toluidine blue). (a-d) *Bdmads13* RNAi line carpel in four stages. (e-h) *BdMADS3-3*, (i-m) *Bd21* (WT) grain development and filling stages, *Brachypodium distachyon* (Bd), Pre-anthesis (Pr AS), post-anthesis (Po-As), young grain (YG), mid-length grain (ML), full-length (FL), DAA: days after anthesis, Scale bar: 0.1 mm.

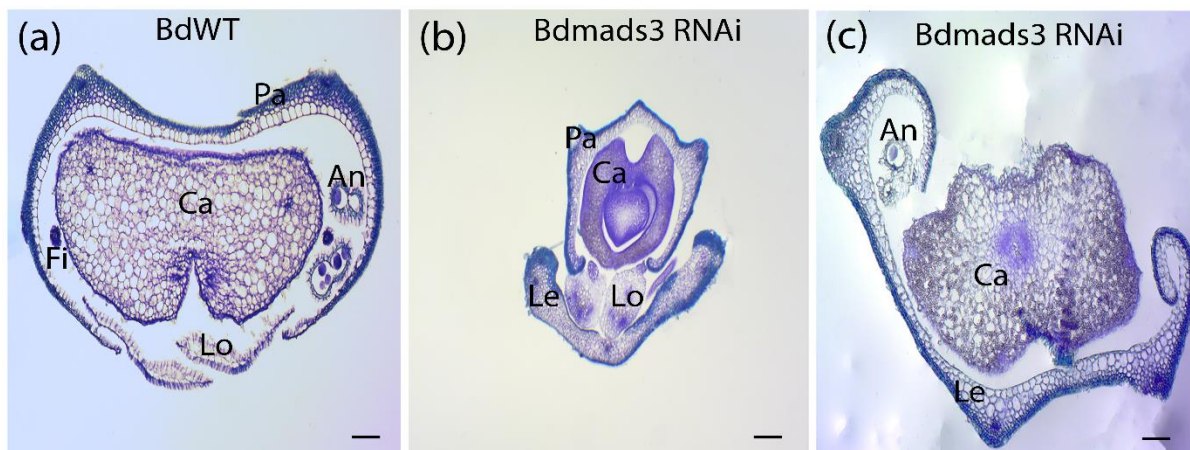


Figure 4-14: *Brachypodium* pre-anthesis floret transversal semi-thin sections observed with light microscopy stained with toluidine blue. (a) *Bd21* Wild-type apical section of the carpel, (b-c) *Bdmads3* RNAi, (b) apical section (c) basal section of the floret showing the inflated and thick-ended lemma with three vascular bundles, expended lodicules. (Le) Lemma, (Pa) palea, (Lo) lodicule, (An) anther, (Ca) carpel. Scale bar 0.1 mm.

The transverse sections reveal a very significant alteration in the first (lemma and palea) and second whorls (lodicules) in the *Bdmads3* RNAi line (Figure 4-14 (b)). At early stages the lemma is thick and short with round edges unlike the wild type which is thin and long enclosing the young grain. The peripheral edges are pointy and overlap with the palea to protect the young grain. In subsequent stages, the *Bdmads3* RNAi palea starts elongating containing not only two vascular bundles in the case of the wild type but three vascular bundles, one in the middle and two at the edge (lateral side). The carpel presents an apparent malformation at the apical side, failing to form a uniform round structure of the pericarp. In these sections, we could see a noticeable difference in the lodicule size as well, whereas, in the *Bdmads3* RNAi line revealed an increase in thickness and a round shape contrasting the wild type which is relatively flat at this stage.

To investigate the structure of the palea, a comparative histological analysis was performed on the altered organs in *Bdmads3* RNAi line and its corresponding organs in wild type, using semi-thin sections stained with Toluidine blue.

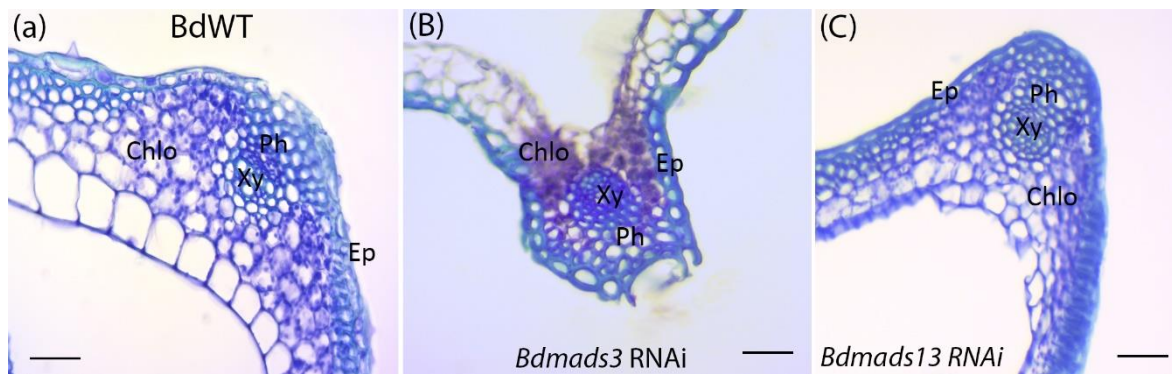


Figure 4-15: Transverse sections of *Brachypodium* palea vascular bundle observed with light microscopy stained with toluidine blue. (a) *Bdmads3-2* RNAi vascular bundle. (b) *Bd21* (WT) vascular bundle, (c) *Bdmads13-3* RNAi vascular bundle. Ep Epidermis. Co Cortex (chlorenchyma + sclerenchyma cells). Xy Xylem. Bu Bundle Sheath. Pi Pith. Scale bar 10µm.

The vascular bundle presents three rather than two in the palea. Although the section showed no difference in the number of the tissues forming the palea vascular bundle, the structure of some of these tissues were remarkably deformed in the *Bdmads3* RNAi. The epidermis presents a large cell along the adaxial side compare to the wild type (Figure 4-15 (b)); however, the xylem cells appeared to be condensed, and their sizes are uniform in the RNAi line. The bundle sheath cells transformed into a region of compactly arranged smaller cells surrounding a vascular bundle especially in the area closer to the xylem and phloem and then shrunk pulling the two side together and causing a twist in the palea. The pith seems to completely disappear in the area adjacent to the vascular bundle. There was no apparent difference between the wild type and *Bdmads13-3* RNAi vascular bundle (Figure 4-15 (c)).

The next organ that presented a malformation in the *Bdmsda3* RNAi line is the rachis. Therefore, a stained semi-thin section was observed with a light microscope in order to investigate and compare the structure carefully.

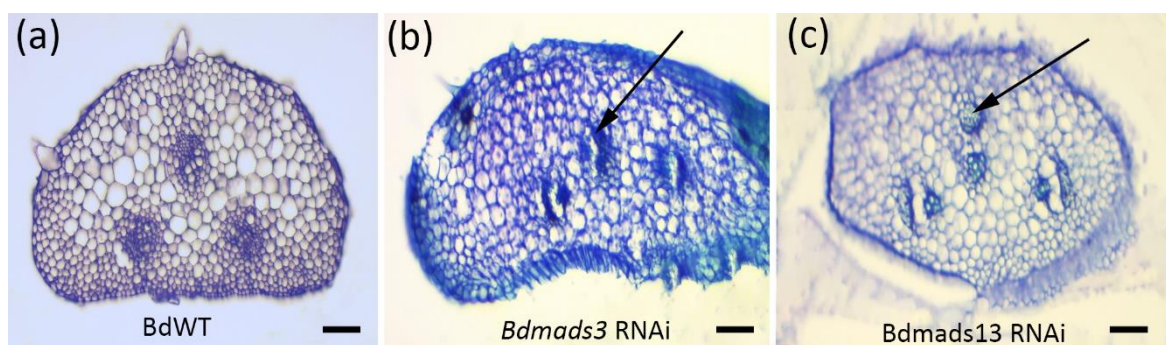


Figure 4-16: Transversal sections of *Brachypodium* ultrastructure of the rachis under light microscopy stained with toluidine blue at anthesis stage. (a) *Bd* WT, (b) *Bdmads3-2* RNAi line (*BdM3i*), (c) *Bdmads13-3*

RNAi line (*BdM13i*), the sections were taken from the internodes at the level of the pre-anthesis anthers.
Scale bar 50 μ m.

The *Brachypodium* wild type rachis presents three central vascular bundles surrounded by the cortex containing chlorenchyma and sclerenchyma cells and has a round shape (Figure 4-16 (a)). The *Bdmads3-2* and *Bdmads13-3* RNAi rachis structure show that there was a malformation in the vascular bundle and slight separation between condensed xylem and phloem by an air space, however, these results need to be further investigated with high-resolution ultra-thin sections.

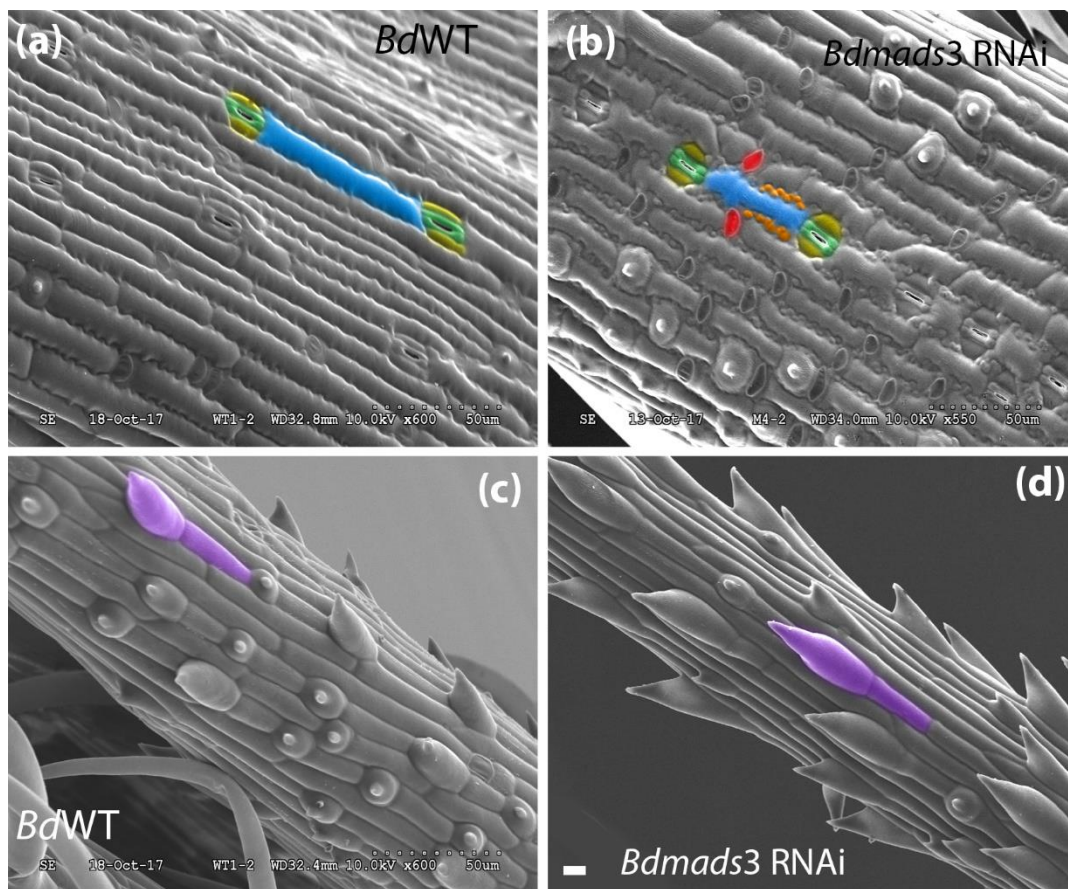


Figure 4-17: SEM of (a-b) the adaxial surface of the lemma and (d-e) awn of *Brachypodium* floret, (a) *BdWT* lemma, (b) *Bdmads3* RNAi line lemma, (c) *Bd* WT awn, (d) *Bdmads3* RNAi awn, green highlight: the guard cell of the stoma, yellow: subsidiary cell, blue: long epidermal cells, red: pitted siliceous papillae, organ: silica bodies (shrunken cells), purple: bicellular micro hairs. The zone analysed was the middle adaxial side.
Scale bar 50 μ m

Observing the lemma surface revealed that in wild type, the long cells were the dominant element of the lemma epidermal surfaces and varied in length, that is determined by the axial and transverse anticlinal walls surrounding it (Figure 4-17 (a)). The stomata were distributed uniformly along the lemma. On the other hand, the RNAi line showed an abundance of pitted siliceous papillae (Figure 4-17 (b)) (Terrell & Wergin, 1981) in the

lemma epidermis and had a differentiated anticlinal cell wall around the long epidermal cells with some shrunken silica bodies (Kumar *et al.*, 2017).

It is known that silicic acid is involved in photosynthesis, transpiration and many plant growth processes, and is selectively transported to the panicle of rice during its maturation, perhaps because of its increasing accumulation ability. Silica cells are among the first type of cells to be silicified in a tissue, sometimes even before emerging inflorescence. In the spikelet, silicic acid is concentrated and deposited in the inflorescence lemma and palea, which serve as a tough protecting shield to the developing grain after the walls of silica cells are lignified (Kumar *et al.*, 2017). There was no visible malformation of the awns in the RNAi lines as the structure of the cell, and the micro hair appears similar to the wild-type (Figure 4-17 (d-e)).

These results suggest that the abundance of these structure of the silica bodies cells in the RNAi line result in the rigid, abnormal lemma shape and cleistogamous flowers.

4.4.4 Lodicule morphology

Lodicules were the next floral tissue investigated. They are the two organs lying between the lemma and the ovary, which engorge with water at the basal cushions for the stigmas to emerge and stamen filaments to elongate in chasmogamous floret at anthesis stage.

In *Bdmads3* RNAi lines the lodicules seem to be dehydrated, thin and fail to enlarge. One plant in the *Bdmads3* RNAi line displayed a different phenotype for lodicules with a long shape and some trichomes at the end (Figure 4-18(a)). A similar phenotype was shown by (Kanget al., 1998) for rice *OsMADS3* transgenic line.

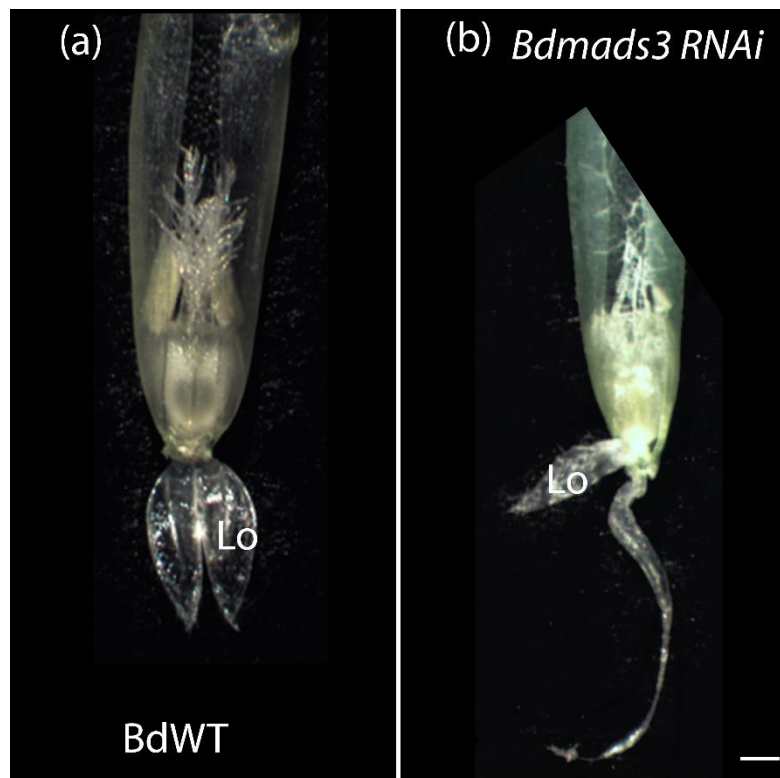


Figure 4-18: Lodicule morphology comparison between *Bd mads3* RNAi line and WT line at the pre-anthesis stage. (a) *Bd mads3* RNAi line floret presenting a malformed and elongated lodicule on the right, indicated by white arrow and shrunk lodicule on the left. (b) WT floret presenting full hydrated lodicules of similar size at the stage of pre-anthesis. Scale bar 1 mm.

A comparative lodicule observation was performed between the lines using fresh lodicule tissues at anthesis stage.

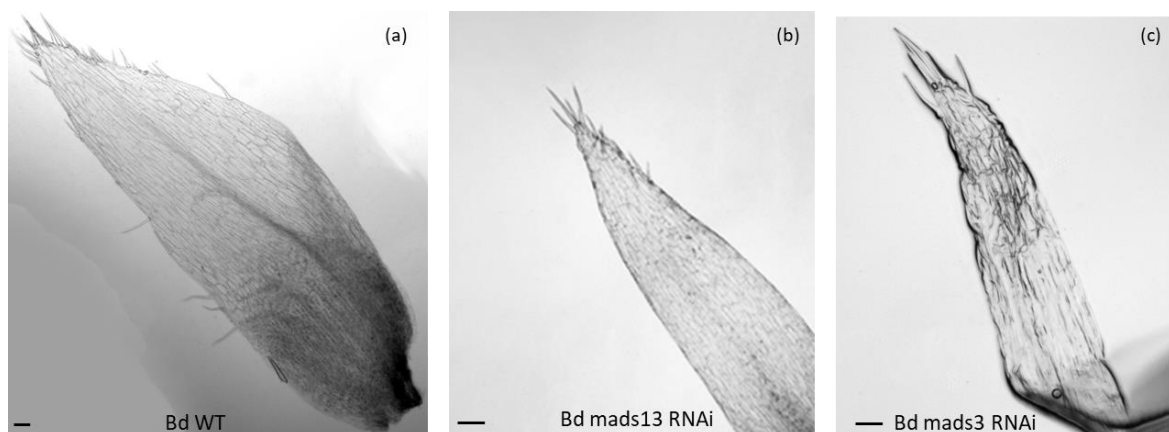


Figure 4-19: *Brachypodium* lodicule at anthesis stage. (a) *Brachypodium* WT lodicule, (b) *Bd mads13-3* RNAi line lodicule, (c) *Bd mads3-2* RNAi line. Scale bar 50µm.

In the wild type, at the base of each lodicule a flux of water associated with the accumulation of K^+ causes the distensible cells of the basal cushion to expand and swell facilitating the anthesis process to take place (Figure 4-19 (a)). In *Bd mads3* RNAi, both

Iodicules show no sign of hydration nor enlargement (Figure 4-19 (c)). The Iodicules of *Bdmads13-3* RNAi presented similar phenotype as the wild type there were hydrated at a mature stage and do not show an altered structure compared with that of the wild type (Figure 4-19 (b)).

4.4.5 Anther development and morphology

The stamens have distinct characteristics and highly conserved morphological structure, consisting of filament and anther with four locules. A comprehensive anther morphological and histological analysis was performed to further identify the role of *BdMADS3* and *BdMADS13* in anther development.

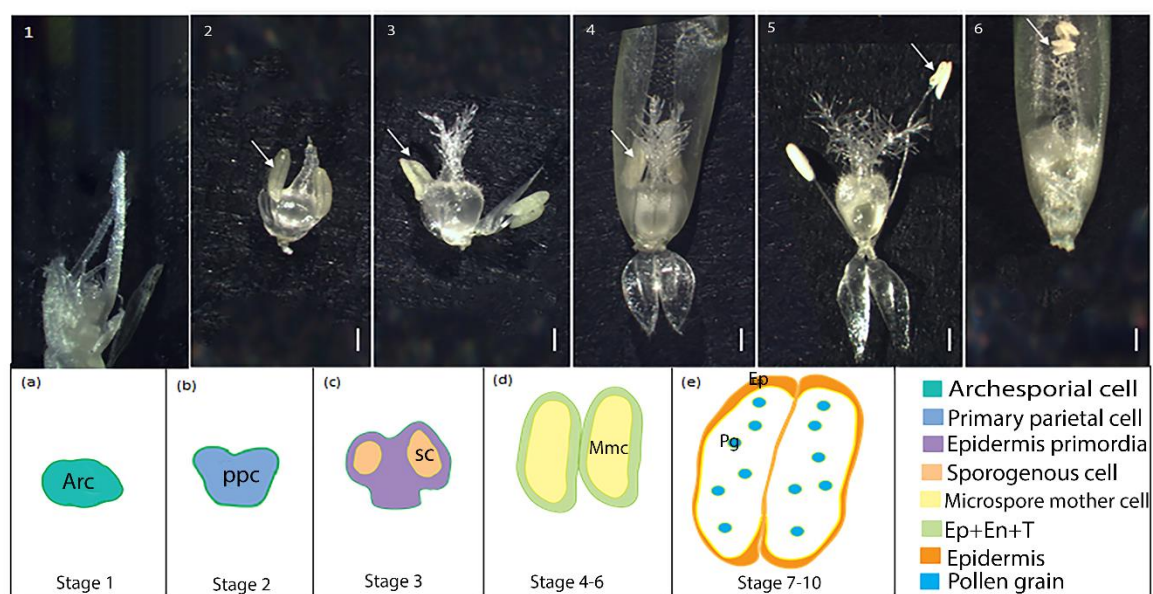


Figure 4-20: The anther developmental stages in *Brachypodium*. (1-6) anther morphology from meristem primordia to the post-anthesis stage. (a-e) anther anatomy schematic illustration. Ep: epidermis, En: endothecium, T: tapetum, Arc: archesporial cell, ppc: primary parietal cells, Sc: sporogenous cells, Mmc: Microspore mother cell. Arrow indicates anthers. Scale bar: 1 mm.

Male gametophyte development in the anther is an essential process in the angiosperm resulting in the formation of sperm cells in the pollen grains that are required for the fertilisation of the egg (Sharma, Singh *et al.*, 2015). Any defect in the development of pollen or the surrounding nutritive layer of the anther, the tapetum, can lead to male sterility and failure to set seed (Polowick, 1992).

In wild-type *Brachypodium* at stage 1 (Figure 4-20 (a)) the anther wall formation starts with archesporial cells which divide rapidly to generate distinct primary parietal cells and sporogenous cells surrounded by three anther wall layers; an epidermis, endothecium and

tapetum at stage 3 and 4 (Figure 4-20 (d)). Subsequently, a fourth layer appears between the tapetum and the endothecium called the middle layer. Then the sporogenous cells generate microspore mother cells and at this stage the tapetum become prominent. The mother cells undergo meiotic divisions resulting in the formation of tetrads and the tapetum becomes vacuolated at stage 5 and 6. The tetrads then release free haploid microspores by stage 7, which will become vacuolated bicellular pollen and the middle cell layer undergoes crushing events (Figure 4-20 (e)). The prominent tapetum layer starts to degenerate into cellular debris by the subsequent stages 8 and 9 when the microspores undergo mitosis I and II to become trinuclear pollen. At stage 10 (maturation) (Figure 4-20 (e)) the epidermis and endothecium layers collapse in the stomium area to release mature pollen grains for anthesis (Sharma, Singh *et al.*, 2015).

4.4.6 Anther morphology and structure analysis using scanning electron microscopy (SEM)

A comprehensive analysis was performed in order to follow the phenotypic and cytological development of the anther. Our investigation includes the anther structure and pollen grain viability. SEM was used to characterise the morphology of the developing *Bdmads3* RNAi anther.

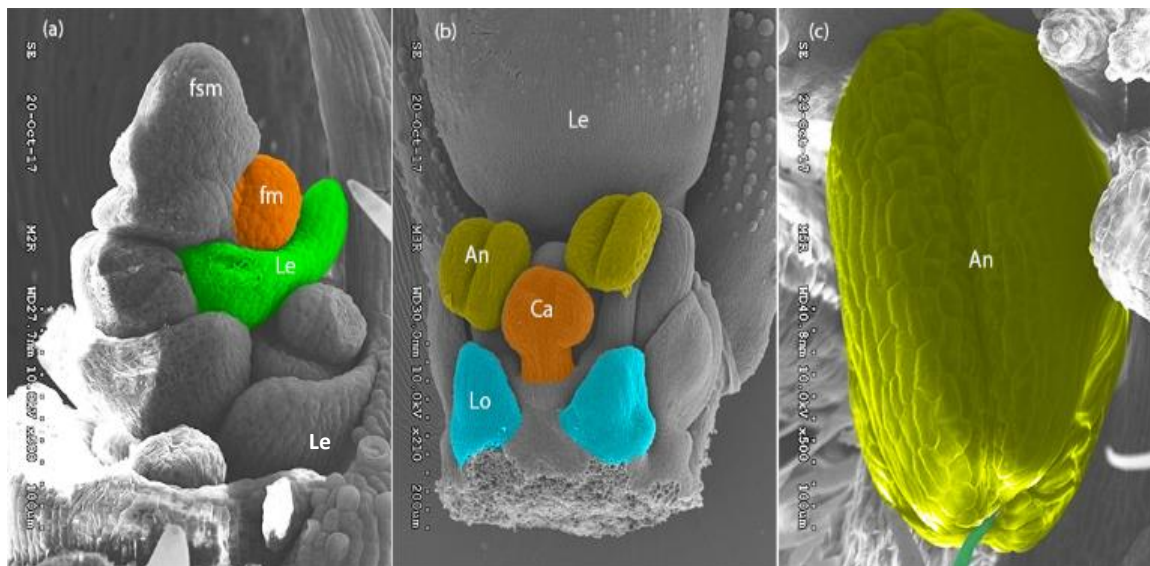


Figure 4-21: Scanning electron microscope of *Bdmads3* RNAi line developing anther, (a) stamen primordia (white arrow) at emerging spike stage (b) *Bd mads3* RNAi anther shows absence of dehiscence, (fsm) terminal spikelet meristem, (fm) floral meristem, (Le) lemma. (An) Anther, (Ca) Capel, (Lo) Lodicule. Scale bars (a, c) = 100 µm, (b) = 200 µm.

In the *Bdmads3* RNAi inflorescence, at the early stage of development (Figure 4-21 (a)), lateral floral meristems appeared similar to wild-type spikelets, and each branch had initiated axillary meristems which differentiate into developing floral organs. At

subsequent stages, the lodicule, anther and carpel primordia continue to elongate similarly to the wild-type and could be distinguished. However, the mature stage of the stamen reveals an absence of pollen dehiscence (Figure 4-21 (c)). These results suggest that *BdMADS3* controls wall dehiscence in the anther at maturity. This speculation is reasonable because in *Arabidopsis* members of *AGAMOUS* family, SHP1 and SHP2 control dehiscence in the *Arabidopsis* silique by an absence of lignified cells in the replum (Liljegren *et al.*, 2000).

4.4.7 Light microscopy analysis

Brachypodium anther development like major angiosperms is tetrasporangiate anther which has a four-layered wall, epidermis, endothecium, middle layer and tapetum, organised to contribute towards efficient pollen development (Bhandari, 1984).

To further analyse the alterations in the anther, a comprehensive phenotypic comparison was performed using light microscopy. A fresh anther at the pre-anthesis stage was compared to the RNAi lines (Figure 4-22), and structural organelle was then observed under light microscopy to determine the pollen organisation inside the locule (Figure 4-22).

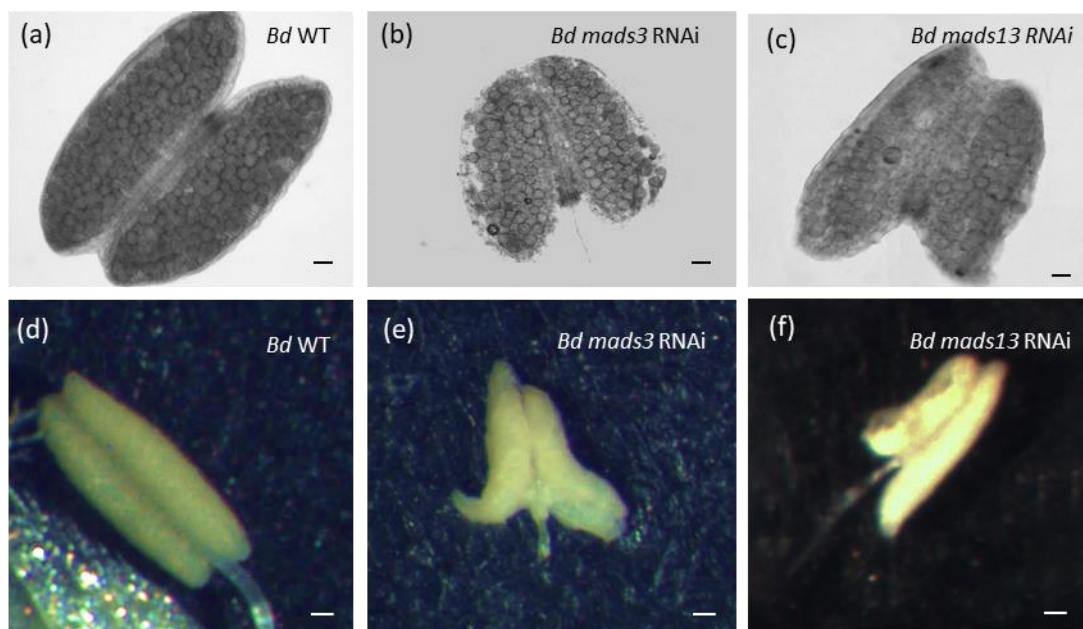


Figure 4-22: *Brachypodium distachyon* mature anther phenotype. (a-c) light microscopy of fresh anther, (d-f) phenotype of the anther, (a. d) *Bd* WT, (b. e) *Bd mads3-2* RNAi anther with malformed shape and small size, (c. f) *Bdmads13-3* RNAi, Scale bars: 100 μ m.

At stage 10 (Sharma, Singh *et al.*, 2015) (Figure 4-22 (a-d)), the bicellular pollen grains are filled with starch and lipids and ready for pollination. From the observations made on the

silenced line *Bdmads3 RNAi*, it appears that the anther at this stage is malformed and curls away from the connective channel (vascular bundle) with pointed ends for each lobe (Figure 4-22 (b-e)). The size is reduced longitudinally by about half in comparison with wild-type. Similarly, the *Bdmads13-3 RNAi* was deformed with unequal size of the anther lobes (Figure 4-22 (c-f)). The pollen grain showed no differences in the shape and organisation inside the anther for both RNAi lines.

4.4.7.1 Anther histological analysis and cytochemical staining

In angiosperms the development of the microspores and pollen maturation depends primarily on the tapetum cells which provide supplies for pollen wall formation and nutrients for pollen development. Microsporocytes produce a short life cell wall consisting of β -1,3-linked glucan (callose) between the primary cell wall and the plasma membrane. It is believed that its role is in protecting the developing microspores from the effect of the surrounding tissues and from premature swelling or bursting (Heslop-Harrison & Mackenzie, 1967). The degradation of the callose temporary wall facilitates the release of free microspores into the locular space (Waterkeyn, 1962).

During the late tetrad stage of microspore development, especially during exine synthesis, the tapetum cells release a β -1,3-glucanase (callase) into the locular space (Stieglitz, 1977) and the callose wall is degraded to free the microspores. The tapetum undergoes cellular degradation, considered to be a programmed cell death (PCD) event which is characterised by sequential elimination of the cellular structures at maturity stages of pollen development (Papini *et al.*, 1999; Wu & Cheung, 2000).

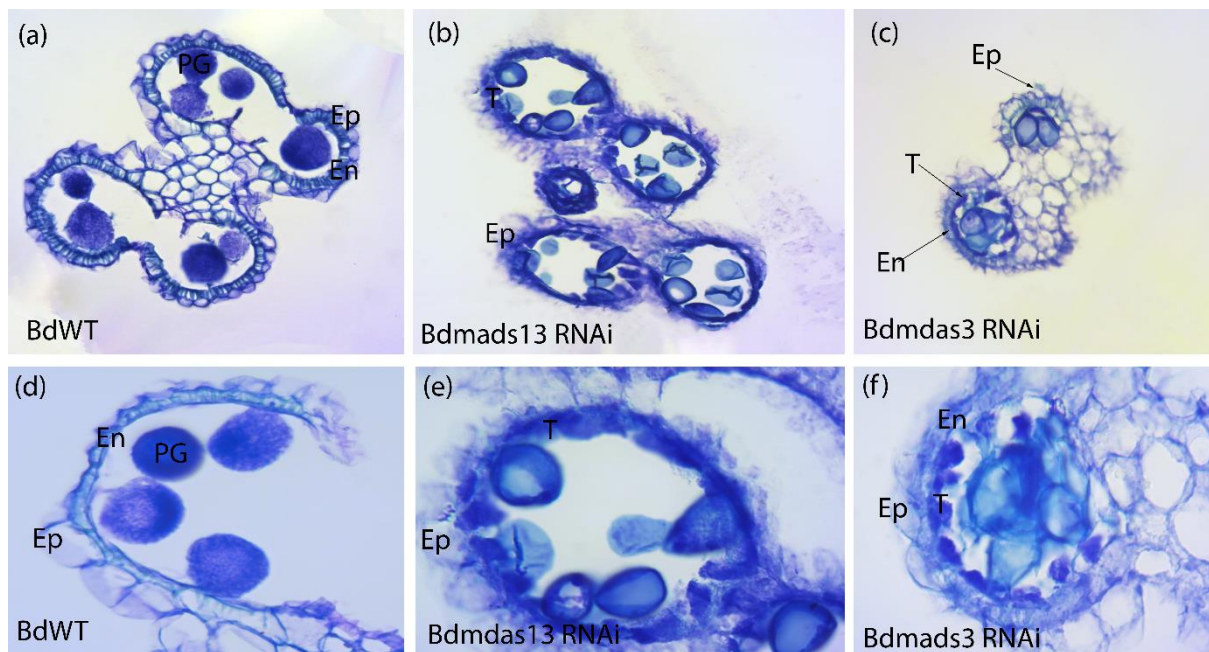


Figure 4-23: Light micrographs of longitudinal semi-thin sections of mature anther stained with toluidine blue. (a-d) *B. distachyon* 21 (Wild-type) anther wall showing the layers of the anther wall, epidermis, and endothecium. (b-e) *Bdmads13-3 RNAi* line anther wall presenting an absence of visible endothecium and irregular darkly stained tapetal cells. (c-f) *Bdmads3-2 RNAi* line anther wall with irregular darkly stained and shrunken tapetal cells EP Epidermis, En endothecium, PG pollen grain, scale bar: 10µm.

The comparison between anthers at maturity, stage 10 (pre-anthesis) of *B. distachyon* 21 (Wild-type) and both RNAi lines reveal that the structure of the anther wall is different (Figure 4-23). *B. distachyon* wild-type present a layer of tapetum in the anther wall on the side of the anther locule at early stages then starts shrinking by stage 6, degenerating and degrading by stage 8 (Figure 4-20). In subsequent stages 9 and 10, the tapetum layer disappears, and the epidermis collapses and endothecium degrade to release the pollen grains which contain lipids and nutrients (deeply stained) (Figure 4-23 (a-d)). However, the *mads3i* anther wall layers, including the tapetum, became disarranged and the degenerated tapetal cells contain dark cytoplasm with a large number of peroxisome-like organelles by stage 10 (Figure 4-23 (c-f)) that persist. The pollen grains were less condensed (lightly stained) comparing to the wild-type, which contains starch granules, indicating that they might be not fully mature.

On the other hand, *Bdmads13-3 RNAi* anther (Figure 4-23 (b-e)) presents an unstructured endothecium layer, replaced by a large arrangement of deeply stained tapetal cells. The pollen grains present a clear difference, while the wild type grains were dark and rich with starchy cytoplasm, *Bdmads13-3 RNAi* pollen appears to be shrunken and empty (Figure 4-23 (b-e)).

A similar phenotype was found in the effect of high-temperature stress on male and female reproduction in plants (Sage *et al.*, 2015). ROS especially H_2O_2 contribute to either cell wall loosening or cell wall stiffening (Cheeseman 2007, Bhattacharjee 2017). Therefore, we suggest that the loss of function of *BdMADS3* in the plant result in a cascade of reactions that led to high levels of reactive oxygen species (ROS) which induce a destruction in the cellular structure of the anther wall.

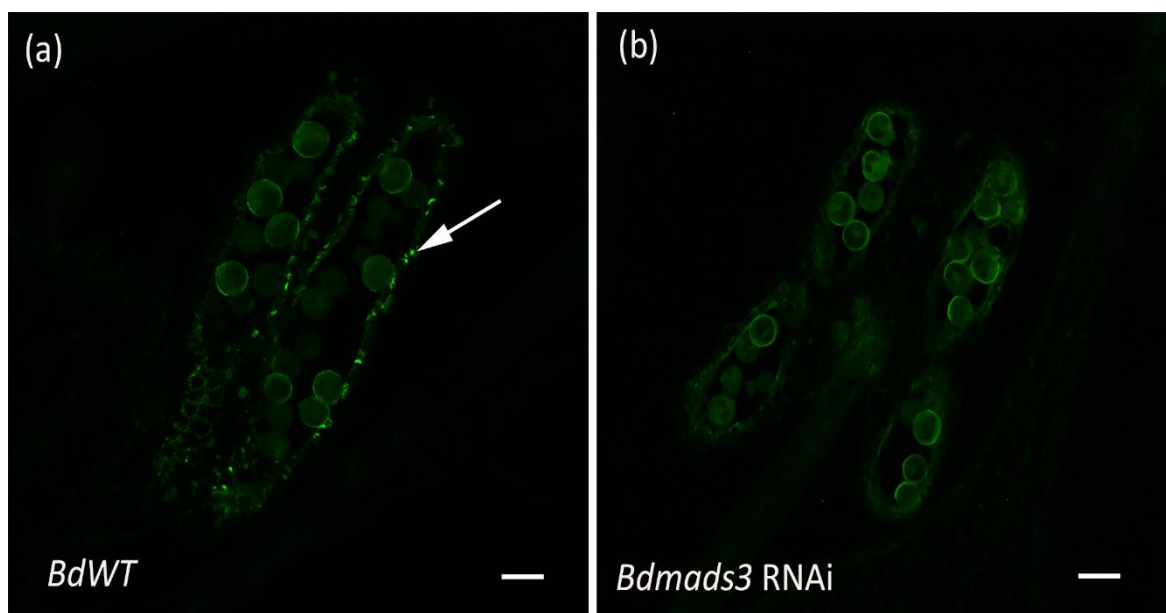


Figure 4-24: comparison of DNA fragmentation detected in *Brachypodium* anther at maturity stage between the wild-type and *Bdmds3* RNAi line at maturity stage. Scale bar 50 μ m.

The indehiscent anther of the *Bdmds3* RNAi did not show a significant nuclear DNA degradation around the anther wall (Figure 4-24 (b)). However, the wild-type at maturity stage degenerates the anther wall to disperse the pollen grain (Figure 4-24 (a)).

4.4.8 Pollen grains morphology

4.4.8.1 Pollen viability analysis using Alexander staining:

Alexander staining was performed to investigate the viability of the pollen. The malachite green stains the aborted pollen wall with green, while, viable pollen protoplasm is stained red-purple. Orange G has been used in plant histology as a background stain for nuclear dyes and in Alexander stain was found to improve the differentiation.

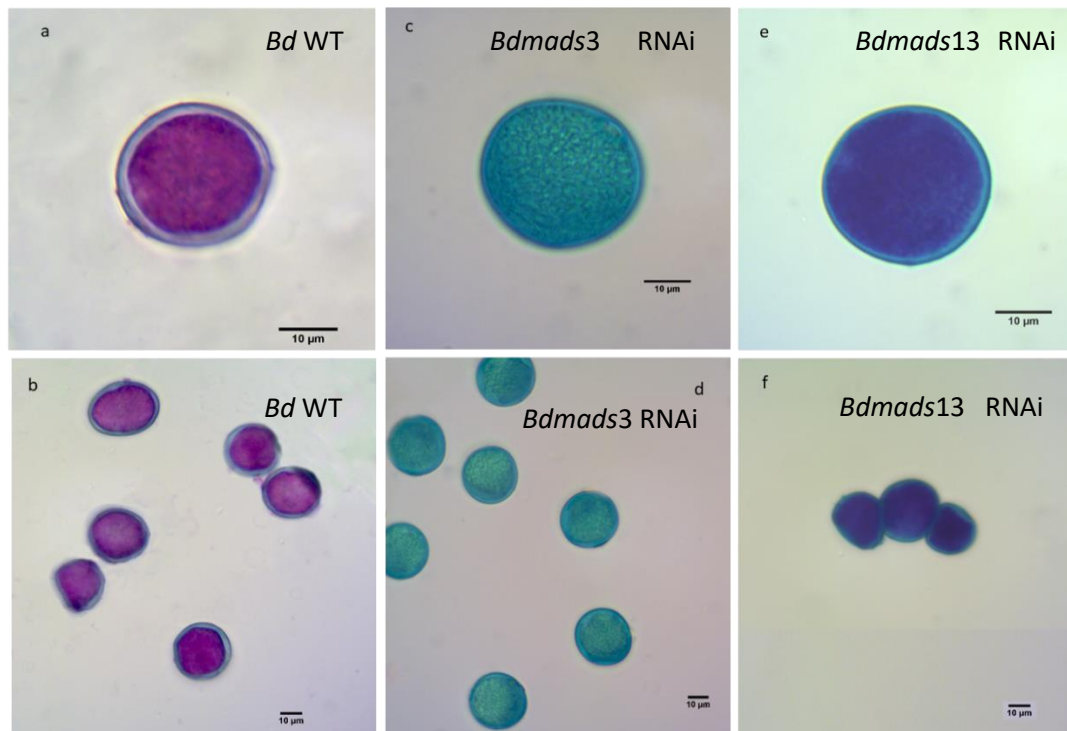


Figure 4-25: Mature pollen of *B. distachyon* stained with Alexander staining. (a. b) viable pollen of *Bd21* (WT) present well distributed protoplasm which becomes purple with acid fusion, (c. d) non-viable pollen from *Bdmads3-2* RNAi line showing irregular protoplasm and blue-green cellulosic wall, (e. f) deeply stained pollen from *Bdmads13-3* RNAi line showing dark blue cellulosic wall and protoplasm, scale bar 10 µm.

At mature pollen grain stage (Sharma, Singh *et al.*, 2015) Alexander staining demonstrates that the *Bdmads3* RNAi line pollen grains were aborted showing irregular protoplasm and blue-green cellulosic wall (Figure 4-25 (c-d)). However, the wild-type viable pollen present normal well-distributed protoplasm, which becomes purple with acid fusion red-purple (Figure 4-25 (a-b)). *Bdmads13-3* RNAi pollen grains were not clearly stained blue-green characterising the aborted pollen. This line showed a deeply stained dark blue cellulosic walls.

4.4.8.2 Pollen grains nucleic content and morphology analysis using fluorescence staining and phase-contrast microscopy

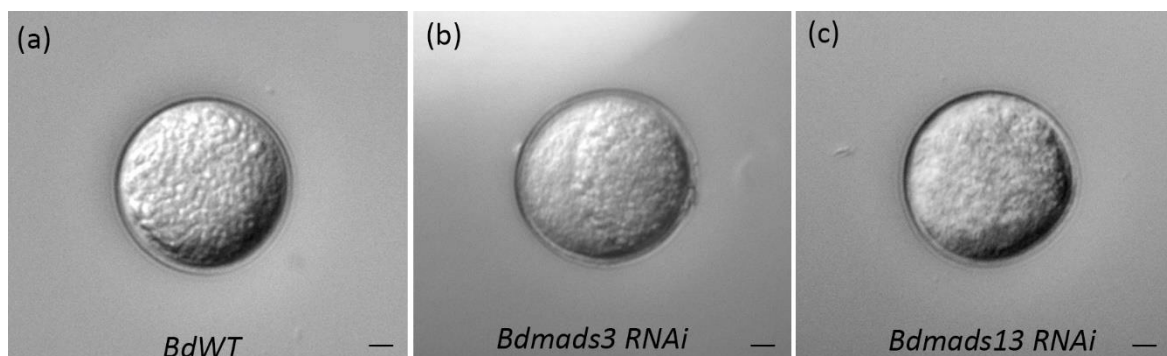


Figure 4-26 : Mature pollen grain morphology under phase-contrast microscope. (a) WT pollen grain, (b) *Bdmads3-2* RNAi, (c) *Bdmads13-3* RNAi. Scale bar: 10 µm.

Morphological analysis of mature pollen grain of wild type *B. distachyon* (Figure 4-26 (a)) showed that they are spherical and present a single operculate (annulate pore), the exine is uniform with tectum and tryphine (small granular shape).

Phase-contrast microscopy could not reveal a noticeable difference in RNAi lines pollen grain compared to the wild-type; however, the tectum and tryphine, which form the exine wall seem to be smaller (no measures were made to confirm) (Figure 4-26 (b-c)).

4.4.8.3 Nucleic acid organisation and male germ phenotype using fluorescent stain DAPI (4',6-diamidino-2-phenylindole)

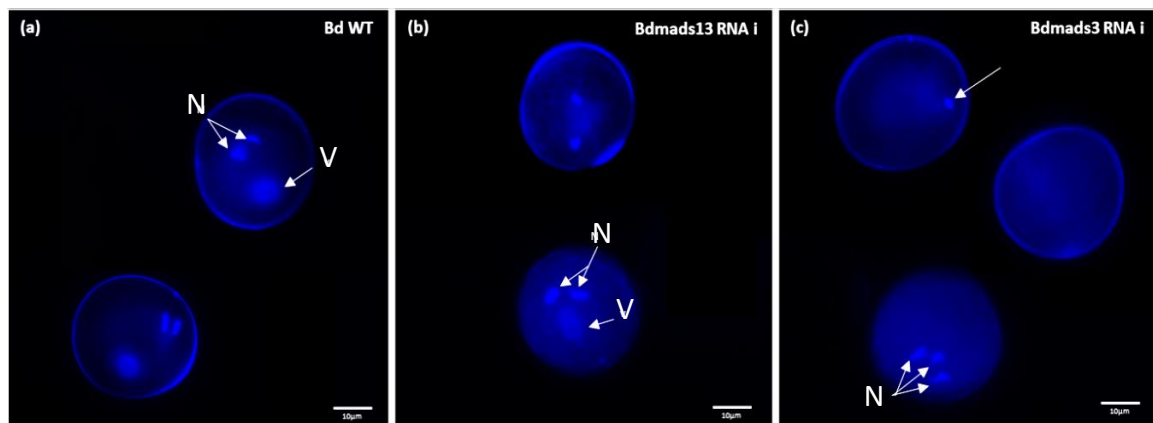


Figure 4-27: Fluorescent micrographs of fresh mature pollen grain stained with DAPI. (a) *B. distachyon* 21 (Wild-type) trinucleate mature pollen presenting one vegetative nucleus and two long spermatic nuclei (White arrow). (b) *Bdmads1-3* RNAi line mature pollen grain with three spermatic nuclei (White arrow), (c) *Bdmads3-2* RNAi line mature pollen grain with undivided nuclei, aborted nuclei, or malformed nuclei with trinucleate bright field, (v) vegetative nucleus, (N) long spermatic nuclei, Scale bar: 10µm

Fixed mature pollen grain was stained with DAPI revealed the nucleus organisation and structure. The wild type grain present lightly stained vegetative nucleus centred in between two long spermatic nuclei bright and deeply stained (Figure 4-27 (a)). On the other hand, the *Bdmads3* RNAi line present trinucleate pollen grain with equal size and equality stained with DAPI (Figure 4-27 (c)). There was some aborted pollen grain with no nuclei and some with undivided nuclei. There were no apparent differences in the pollen grain wall nor the size. On the other hand, *Bdmads13-3* RNAi pollen grain present similar phenotype as the wild type. DAPI staining reveals that the spermatic nuclei and vegetative nucleus were barely altered in this line. These results support the Alexander staining observation and suggest that there was a disruption that happened to the nuclei at the division level and formation in the *Bdmads3* RNAi line.

4.5 Discussion

The C lineage gene *BdMADS3* regulates late anther development and pollen formation.

In this study, we characterised RNAi lines of the C-lineage member *BdMADS3* and the D-lineage *BdMADS13* in *Brachypodium distachyon* by the loss-of-function approach. We further analysed the expression pattern of these genes through the life cycle of the plant. Our results showed that the loss-of-function of *BdMADS3*, result in a disruption of the anther wall formation and pollen function, while the loss-of-function of *BdMADS13* affected the anther wall layers structure.

It is known that *MADS-box* proteins apply regulatory effects on numerous metabolic processes (Riechmann *et al.*, 1996). However, genetic and molecular analyses have identified a number of ABCDE model *MADS* genes involved in plant anther development, including C class gene *AGAMOUS* (AG), B class genes *APETALA3* (AP3) and *PISTILLATA* (PI), and E class genes *SEPALLATA1-4* (*SEP1-4*), are all associated with anther development in *Arabidopsis* (Coen & Meyerowitz, 1991; Theißen, 2001; Xing *et al.*, 2011; Dreni *et al.*, 2013; Zhang & Yang, 2014; Qu *et al.*, 2014).

Additional *MADS-box* genes have been implicated in regulating plant anther development. In maize, *MADS2* was found to be essential for anther dehiscence and pollen maturation, and loss-of-function of *MADS2* resulted in abortion pollen grain and defective anther development (Schreiber *et al.*, 2004). In *Arabidopsis*, *SPOROCTELESS/NOZZLE* SPL/NZZ controls the development of anther walls and pollen mother cells, as the mutant fails to differentiate the primary sporogenous into pollen mother cells (Yang *et al.*, 1999). Moreover, *AGAMOUS* is essential for early stamen formation throughout the activation of the expression of SPL (Ito *et al.*, 2004). In the subsequent stages, *AGAMOUS* continues its expression in the anther to controls anther dehiscence by directly activating the expression of the *DEFECTIVE IN ANTER DEHISCENCE1* (DAD1) gene that encodes a jasmonic acid (JA) (Ito *et al.*, 2007). One of the *AGAMOUS* orthologs of in rice is *OsMADS58* found to act as an inhibitor of numerous photosynthetic genes by binding to the CArG-box motif in their promoter and subsequently affecting chloroplast differentiation in rice stamens (Chen *et al.*, 2015). Preceding transcriptomic analysis has shown that some ROS scavenging genes were significantly altered in *mads3* mutant anthers, which is the second member of the C-lineage in rice (Hu *et al.*, 2011). The stable state ROS level, that is important for normal

development processes, is determined by the equilibrium between ROS-producing and ROS-scavenging mechanisms (Hu *et al.*, 2011), members of ROS such as hydrogen peroxide and superoxide anion, play a critical roles in plant cell death (Bouchez *et al.*, 2007). (Hu *et al.*, 2011) also showed that during young microspore formation, there was a significant accumulation of the reactive oxygen species (ROS) in the wild-type rice.

In our study, flowers of *Bdmads3* RNAi plants display alterations in their rachis which was deformed into zigzag shape and also the stamens in the third whorls were altered with aborted pollen grains, resulting in complete male sterility. Furthermore, the cytological analysis showed that there was a defect in both *Bdmads3* and *Bdmads13* RNAi lines, in the anther wall structure and layers, especially the tapetum. The crucial role of the tapetum in pollen development is highlighted by the fact that male sterility in plants is often associated with tapetal abnormalities (Polowick, 1992). The tapetum is an innermost sporophytic layer within the anther wall, undergoing cell degeneration to supply nutrients for pollen development (Papini *et al.*, 1999; Varnier *et al.*, 2005; Li *et al.*, 2006). Tapetal cell differentiation and disintegration coincide with the postmeiotic development and start degrading at maturity stages. However, in *Bdmads3* and *Bdmads13* RNAi, there was a persistence of uneven layer of the tapetum cells in the anther wall. Similar cellular defects found to be related to oxidative stress caused by disrupted reactive oxygen species level and unbalanced plant stress and programmed cell death (Hu *et al.*, 2011). The high levels of reactive oxygen species (ROS) induce a distraction in the cellular structure (Tsanko *et al.*, 2006).

Gibberellins regulate several rice genes encoding transcription factors, such as GAMYB, and its mutation causes the tapetal cells to degenerate in slower rate (Kaneko *et al.*, 2004; Tsuji *et al.*, 2006; Aya *et al.*, 2009; Liu *et al.*, 2010), UNDEVELOPED TAPETUM1 (UDT1) that cause male sterility in its mutation (Jung *et al.*, 2005), and TDR (tapetum degeneration retardation) that is express in abundance in the tapetum and acts as a positive regulator of tapetal cells PCD in rice and microspore viability (Li *et al.*, 2006), all these genes have been reported to be associated with tapetal function and degeneration. In addition, previous research found some genes that play a crucial role in the post-meiotic tapetum and microspore development. For instance, the ABORTED MICROSPORE (AMS) gene encoding a basic helix-loop-helix (bHLH) TF regulates gene expression by binding to an E-box (CA-

NN-TG) and regulate tapetal cell development and microspores postmitotic formation (Sorensen *et al.*, 2003; Bouchard *et al.*, 1998). MALE STERILITY1 gene encodes a protein with a PHD finger found to be essential for proper tapetum function and normal microspore development (Wilson *et al.*, 2001). The Cys protease gene, member of A1 family (Os04g0670500) that is identified in chapter 6 and the protease inhibitor *OsC6* (Os11g37280) found to be a direct target for the TDR and the mutant line of *OsCP1* cause collapse pollen after release from the tetrad (Lee *et al.*, 2004).

To summarise the different aspects and organs that were disrupted by the loss-of-function of *BdMADS3* and *BdMADS13* genes, a table 4-1 was made to list and compare the effects on the plant.

Table 4-1: Phenotypic alterations in *Bdmads3* and *Bdmads13* RNAi lines.

characteristics	<i>Brachypodium</i> plant lines Phenotype	
	<i>Bdmads3</i> RNAi	<i>Bdmads13</i> RNAi
Blooming	Delay in flowering time	
Height	Short	
Leaf	Curly leaf	Normal
Number of floret in the spikelet	Appears normal	Less
Rachis	Vascular bundle with separate xylem and phloem	
Fertility	Sterile	
Lemma	Abnormal silica bodies and three vascular bundles	Appears normal
Palea	malformed	Appears normal
lodicule	Dry long shape	Appears normal
Anther	Curled lobes	
	Persisting tapetal cells	
	Indehiscent walls	Endothecium layer is absent
Pollen grain	Aborted	Appears normal

Our data also showed that *BdMADS3* has a critical role in regulating postmeiotic anther wall development at maturity stages. As well as playing an essential role in controlling the flowering time. Flowering time control is reported to be related to AGL15 in *Arabidopsis* through direct peroxidase regulation (Cosio *et al.*, 2017). The defect displayed by the *BdMADS3* exceeded the known array for a C-lineage gene that acts in the third and the fourth whorl; however, it also displayed an abnormal rachis, curly leaves, abnormal silica

bodies on the lemma with an extra vascular-bundle which are in the second whorl, and also, the anthers were indehiscent. It is known that in angiosperms, the dehydration of the endothecium force the opening of the stomium and connective cells surrounding the locules (Keijzer, 1987). These results suggest that the tapetal cells undergo PCD in very precise time during anther wall development in order to supply sufficient nutrition for the development of microspores and either premature or a delay of this process results in male sterility. In our case, perhaps the loss of function of *BdMADS3* in the anther cause the metallothionein to increase and prevent the accumulation of ROS in the anther wall in order to promote the tapetum PCD.

The *Bdmads13* RNAi line is sterile, contains viable pollen grain and shows disrupted anther walls development with less florets on the spikelet. Even though this gene has displayed a great sequence similarity to its ortholog in rice *OsMADS13* considered as a D-lineage gene (ovule identity gene), it appears that it also has an effect on the stamen wall formation, as the *Bdmads13* RNAi anther was malformed with curled lobes, persistent tapetal cells and its wall was missing the endothecium layer which is important for proper maturation of the pollen grains. Importantly, *BdMADS13* was sterile, although the carpel morphological structure looked similar to that of the wild type at pre-anthesis stages, the alexander and DAPI staining showed that the line has viable pollen, suggesting that *BdMADS13* is likely a female sterile.

These findings together provide insight into the role of C-lineage genes in regulating male reproductive development and demonstrate that they have an effect far beyond the third and fourth whorl and regulates more than the ROS homeostasis regulation.

Although our knowledge has significantly improved over recent years, we just started to discover the significant role of *AGAMOUS* genes. More is to learn about the interactions these *AGAMOUS* transcription factors form *in vivo*, the transcription factors and plant hormones that regulate them, and the network of genes that is downstream of these relevant homeotic regulators are still to be explored.

4.6 Conclusion

In conclusion, we have demonstrated the role of the floral homeotic C-class D-class genes *BdMADS3* and *BdMADS13* in regulating late anther development and pollen formation. This study illustrates the broad effect of the transcript reduction of these genes in plant stature, flowering time and floral whorls development. The downstream target of these genes is analysed in chapter 5 that sheds light on the molecular mechanisms underlying the genetic control of these genes and investigation the biological process involved.

Chapter 5

Identification of potential downstream targets for *BdMADS3*
and *BdMADS13*.

Chapter 5. Potential downstream targets for *BdMADS3* and *BdMADS13*

5.1 Abstract

Plant development is regulated by a robust system that controls gene expression in order to manufacture an appropriate level of protein or RNA in the right cells at a precise time. The *MADS-box* genes involved in floral morphogenesis are part of that system. Members of this family include C-class *BdMADS3* and the D-class *BdMADS13* genes that control the development of the third and the fourth whorl in flowers, respectively. However, how these genes control flower development remains largely unknown. In this study, we found that *BdMADS3* interacts with *BdMADS7* transcription factor to bind and activate the metallothionein (MT) gene promoter when using a transient gene expression assay in tobacco leaf. In addition, differential gene expression analysis revealed that the loss-of-function of *BdMADS13* affects the development of the plant severely and causes a significant interruption to the reactive oxygen species homeostasis by activating a high level of ROS-scavenging genes, resulting in complete sterility and desiccation of the plant at anthesis stage.

5.2 Introduction

Brachypodium is considered as a grass model system (Draper *et al.*, 2001) due to its close phylogenetic position to the economically valuable crops such as barley and wheat and also due to the availability of the full genome sequence in addition to the expanding accessibility of genetic resources include a large collection of TILLING lines and T-DNA mutants (<http://www.wheat-tilling.com/>).

The plant flower morphogenesis, including cell identity and differentiation, are determined by a complex interactional network of regulatory gene expressions that are established in ABCDE model for *Arabidopsis* and rice. According to this model in rice, *AGAMOUS* genes, C-class (*OsMADS3* and *OsMADS58*) specify the third, and the fourth whorl and the D-class genes (*OsMADS13* and *OsMADS21*) are more focused in determining the identity of the fourth whorl. These *MADS-box* genes interact or sometimes activate the expression of each other to sculpt a functional flower, including male and female organs.

In the dicot model plants *Arabidopsis*, *AGAMOUS* gene was found to play an essential role in the determinacy of the floral meristem as well as organogenesis in the third and fourth

whorl (Honma & Goto, 2001). At the early stage of inflorescence, *AGAMOUS* transcription factor activate two target genes, *KNUCKLES* and *CRABSCLAW*, which are crucial for meristem determinacy through repressing *WUS* (Alvarez & Smyth, 1999; Liu et al., 2011).

At stage 3, of the flower development, *AGAMOUS* gets involved in carpel identity and continues to express in order to promote stamen identity and microsporogenesis by controlling the expression of the *SPOROCTELESS/NOZZLE (SPL/NZZ)* gene required for promoting the formation of reproductive and parietal somatic cells. In the subsequent stages of flower development, the *AGAMOUS* gene is involved in dehiscence and filament elongation by controlling jasmonic acid level (Ishiguro et al., 2001; Hu et al., 2011).

In grass species, one of the orthologs of the *AGAMOUS* gene is *MADS13*. *AGAMOUS*'s regulatory genes upstream network is still unclear. However, we know that in rice, the *OsDAD1* (Os11g0146600) gene is a negative regulator of programmed cell death in the anthers (Wang et al., 2011). It encodes a chloroplastic phospholipase A1 which stimulates the accumulation of jasmonic acid that promotes water uptake and cell elongation during flowering time (Ishiguro et al., 2001). Biochemical and genetic evidence revealed that jasmonic acid plays an essential role in promoting the degradation of *OsJAZ1* via the *OsCO1b* proteasome in the process of spikelet development (Cai et al., 2014). The cascade continues, and it was found that *OsJAZ1* also interacts with *OsMYC2* to repress *OsMYC2*. Using yeast two-hybrid and luciferase assay, it was confirmed that *OsMYC2* activates *OsMADS1* by binding to G2-box on its promoter (Cai et al., 2014). *OsMADS1* which is a member in the *SEP* family, represses *OsMADS34* in order to promote the transition from spikelet to floral meristem and control floral organ specification (Khanday et al., 2013). This was confirmed by an *Osmads34* mutant in a separate study by (Ren et al., 2016). One of the genes that *OsMADS1* regulates is *FON4*, which controls the number of floral organs (Xiao et al., 2009). *OsMADS1* also physically interacts with the member of C and D-class *OsMADS3*, *OsMADS58*, and *OsMADS13* (Hu et al., 2015).

In rice, the D-lineage *OsMADS13* is expressed first just before the conversion of the floral meristem into the ovule primordium. This expression continues and increases in the ovule in three layers: integuments, nucellus tissues and the inner cell layer of the carpel wall (Lopez-Dee et al., 1999). It has been proposed that *OsMADS13* is a crucial regulator of ovule identity specification and floral meristem termination as its mutant is totally female sterile

because of the total arrest of the gametophyte development and the mutant showed a homeotic conversion of the ovule primordia into a new carpel primordium (Dreni *et al.*, 2007; Dreni *et al.*, 2013).

In *Brachypodium*, our study (chapter 3) revealed that *BdMADS13* mRNA first appeared in the floral organ primordia, however, unlike rice, *BdMADS13* expression is only in the inner cell layer of the carpel wall at pre-anthesis stage and increases in the outer integument after fertilisation. Our investigation provides more details about subsequent stages, where *BdMADS13* expressed in the developing endosperm then in the aleurone layer.

OsMADS3 identified in rice as one of the two members of C-class that are involved in floral meristem determinacy, stamen specification and also control late anther development at anthesis (Yamaguchi *et al.*, 2006). *OsMADS3* mRNA was detected at floral organ initiation, mainly in the third whorl that determines the stamens with a weak signal in the carpel as well. Subsequently, the expression of *OsMADS3* disappeared from the developing stamen primordia and maintains some expression in the fourth whorl just before carpel development initiates, then decreases rapidly once the ovule primordia initiated (Yamaguchi *et al.*, 2006).

In *Brachypodium*, the current study demonstrated that *BdMADS3* expression persists during anther development (chapter 3) (Figure 3.8), especially in the pollen grains at anthesis. We also illustrated the progress of the expression beyond these flowering stages during grain filling and maturity, where the *BdMADS3* mRNA accumulates in the endocarp at pre-anthesis then is expressed strongly in the nucellar lysate, outer integument and the aleurone in the subsequent stages.

CRISPR–Cas9 technology was used to disrupt (frameshift) *OsMADS3* in rice, and the resulting mutant line *Osmads3-fe1* displayed chimeric organs consisting of stamen and lodicule identities as a replacement of stamens and duplicated whorl 3, but no apparent changes in carpel identity in the fourth whorl which suggested that *OsMADS3* is likely to be involved in maintenance of the floral meristem (Yasui *et al.*, 2017). In addition, (Hu *et al.*, 2011) study also showed that their mutant line *Osmads3-4* had defective late anther development, with persisting tapetal cells and aborted microspores, resulting in male sterility.

The current investigation (chapter 4) revealed that in *Brachypodium* the loss-of-function of *BdMADS3* affects many aspects of the life cycle of the plant including flowering time, height, rachis shape, delay in the tapetum PCD, aborted microspores, which suggest that *MADS3* have a divergence of function from rice.

The upstream regulator of the C-class genes *OsMADS58* was analysed using ChIP-qPCR and found that *OsMADS6* directly binding to intron 2 of *OsMADS58* in rice (Li Hu *et al.*, 2011). On the other hand, metallothionein MT-1-4b was identified as one of the downstream targets of *OsMADS3* in rice, in the process of maintaining the ROS homeostasis during late anther development (Hu *et al.*, 2011). It is known that the high accumulation of reactive oxygen species (ROS) such as superoxide anion (O_2^-), hydroperoxide radicals (OH^\cdot), and hydrogen peroxide (H_2O_2) causes severe damage to the plant cells in stress events such as drought, salinity, flooding, heat, and cold (Miller & Mittler, 2006). Metallothioneins are one of the ROS-scavenging non-enzymatic systems that the plant develops to survive such oxidative stress by removing superoxide- and hydroxyl radical (Hu *et al.*, 2011; Gechev *et al.*, 2006).

ROS play an essential role in tapetum PCD during anther maturation. One of the genes that downregulates the ROS level in the anther is *OsMT2b* in rice. At early anther morphogenesis (ST6-ST8) *OsMT2b* keeps the ROS levels low in order to allow the differentiation of the anther wall layers; it is only when *OsDTC1* is produced and physically binds to the *OsMT2b* to block its action, the level of ROS rises to contribute to tapetum PCD (Yi *et al.*, 2016). At maturity stage of the anther (ST10) *OsMADS3* gene activates the expression of *OsMT1b-4* to reduce the level of ROS and maintain homeostasis (Hu *et al.*, 2011).

To our knowledge, no target was identified for *OsMADS13* or its ortholog in any other species.

In our study, we investigated the co-expression of C-class *MADS3* gene in four species (*Triticum aestivum*, *Arabidopsis thaliana*, *Hordeum vulgare* and *Oryza sativa*) and identified highly correlated genes that could be regulating the expression of *MADS3* or be potential downstream target genes. The results using transient assay showed that *BdMADS7* and *BdMADS3* may form a quaternary complex that binds to the metallothionein promoter and activates the gene in Tobacco leaf. Adding to the finding of (Hu *et al.*, 2011) who showed

that *mads3* has a high level of ROS and is low in ROS-scavenging gene expression in rice, suggesting some degree of functional conservation between the two species. We also investigated the D-class gene *BdMADS13* through analysing the transgenic line and identified a list of differentially expressed genes. The transgenic line *BdMADS13* appear to have a high level of some programmed cell death related genes such as cysteine endopeptidase which is senescence-specific protein and the desiccation-associated protein Ferritin-like BRADI_1g62680v3 (Zha *et al.*, 2013). In response to this deterioration, the plant produces a high level of ROS-scavenging genes, including metallothionein and peroxidase, to rescue the cell from oxidative stress.

On the other hand, *Bdmads13* RNAi line presented low expression of some vital genes that intervene in many pathways. For instance, there were genes that are involved in the endosperm development and reproduction, such as, the polycomb protein EED (BRADI_2g07820v3) and genes involved in pollen development (BRADI_1g62007v3). The reduction of the expression of these genes cause significant disruptions to the plant system and eventually complete sterility and desiccation of the plant at anthesis stage.

In this study we used a quantitative experimental approach to measure the differentially expressed genes between the *Bdmads13* RNAi line and the wild type line, in order to select the potential downstream target of the *BdMADS13* transcription factor or at least highlight the biological function that has been affected by the reduction of this gene expression.

5.3 Materials and methods

5.3.1 Plant materials

The lines previously described in the section 4.3.1. were used in this chapter, *Bd21* for the wild-type and *Bdmads13-3* RNAi.

Chapter 3 illustrates the expression of *BdMADS13* in *Brachypodium* which initiate at early floral development in the very young inflorescence then the expression peaks in the carpel at pre-anthesis stage, therefore these stages were selected for the RNA-Seq experiment.

5.3.2 Co-expression analysis and potential targets of *MADS3*

GENEVESTIGATOR is transcriptomes Meta-Analysis database that use high-throughput gene expression analysis retrieved from microarray technologies (Affymetrix, geneChip

data, RNA-seq etc) experiments in chosen environmental conditions, growth stages, or organs that are publicly available.

To identify the *BdMADS3* co-expressed genes, GENEVESTIGATOR® was used, which utilises curated RNA-seq data sets from *Triticum aestivum*, *Arabidopsis thaliana*, *Hordeum vulgare* and *Oryza sativa*. First, the RNA-seq of the desired organism data set was selected then submitted the gene of interest (*MADS3*) then calculate Person's correlation coefficient across the data set to select the co-expressed genes. The ID given by the software was with the old assembly names, therefore, a manual conversion was performed to individual gene names using EnsemblPlants archive (<http://plants.ensembl.org/index.html>). Then the candidate genes that co-expressed with *MADS3* in all selected species were selected to be tested in a luciferase reporter assay for downstream target or protein interaction.

Another approach was used to identify potential targets for *MADS3* from the observation of the transgenic line anther morphology the tapetum cell persisted in the mature anther stages. The same phenotype was reported in (Yi *et al.*, 2016) study in *Arabidopsis* caused by a mutation in MT2b (metallothionein) which play a key role in tapetum degeneration. Therefore its orthologue in *Brachypodium* was selected as a potential target for *MADS3*, which is BRADI_2g28254v3. In addition, (Cosio *et al.*, 2017) reported that *Arabidopsis* peroxidase is a direct target of one of the *MADS*-box transcription factor (*AGL15*) and participates in lignified tissue formation. The orthologue of the peroxidase in *Brachypodium* BRADI_2g12216v3 was selected for further investigations, Both genes (metallothionein and peroxidase) present at least one site for CArG-Box in their promoters.

5.3.3 cDNA and DNA promoters amplification and manipulation

5.3.3.1 Oligonucleotide primer design

Different oligonucleotide primer pairs were designed to amplify genomic promoters, cDNA, and cloned insert and vector sequences (appendix, table S1-S3). PrimerQuest® Tool (<https://eu.idtdna.com>) was used for designing primers for RT-PCR as well as qRT-PCR under the following criteria: the optimum annealing temperature (Tm) ~62°C, GC content ~50% the primer size range between 17 to 25nt with the required an amplicon size.

Adapters:

Table 5-1: Gateway® recombination cloning primers adapters sequences used for each vector.

Adapters name	Complete adaptors sequences (5'-3')	pDONR vector
attB4-F	GGGGACAACCTTTGTATAGAAAAGTTG	pDONRP4-P1r
attB1-R	GGGGACTGCTTTTTGTACAAACTTG	pDONRP4-P1r
attB1-F	GGGGACAAGTTTGTACAAAAAAGCAGGCT	pDONR221
attB2-R	GGGGACCACTTTGTACAAGAAAGCTGGGT	pDONR221

5.3.3.2 Reverse transcription polymerase chain reaction (RT-PCR) and PCR conditions

The cDNA and DNA was amplified using the appropriate primers and MyTaq Red master-mix (BIOLINE) the following amount: for total of 10µl reaction, 5µl from MyTap DNA Polymerase, 3µl from ddH₂O (DNase-free), 0.5µl from each primer (20µM) and finally 1µl from the template (cDNA/DNA) diluted 1:5. The conditions for the PCR were varied depending on the the specific size of the product, in accordance with the T_m and primers. The plate was preheated and the DNA was denatured 6 min at 95°C. The cycles were between 30 to 37 including 95°C for 30 seconds, 50-65°C annealing temperature for 30 seconds, 72°C, 30 seconds/kb for extension then 6 min 72°C final extension. The thermocyclers used for the PCR reactions were the TProfessional Basic Gradient/Trio Thermocycler (Biometra) and TProfessional TRIO Thermocycler ® (analytikjana). The low yield and different fragment size issues were resolved using a temperature gradient to select the optimal temperature for that specific reaction.

5.3.3.3 High-Fidelity DNA polymerase Q5 PCR

Due to the size of the promoter amplified which range between 1000bp to more than 2000bp we chose to use high-fidelity DNA polymerase Q5 to generate PCR products with accuracy.

Table 5-2: High-Fidelity DNA polymerase Q5 PCR reaction components and amounts required.

Components	Amount (25 µl)
5X Q5 Reaction Buffer	5 µl
10 mM dNTPs	0.5 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
cDNA	variable
Q5 High-Fidelity DNA Polymerase	0.25 µl
5X Q5 High GC Enhancer	5 µl

5.3.3.4 PCR product visualisation purification

Only 2 µl of the attB-PCR amplicons and 2 µl from gel Loading Dye (Purple 6X, NEB) was run on a gel to confirm the specificity of our primers and that there was no multiple bands has been amplified. The separation of nucleic acids by agarose gel electrophoresis was performed using molecular grade agarose powder (Bioline), which was melted in 1X TAE buffer (40 mM Tris base, 20 mM glacial acetic acid, 1 mM EDTA), cooled to ~50°C after which ethidium bromide was added at a final concentration of 0.2 µg/ml (0.5 nM). 1% w/v agarose as a concentration was adopted for all the amplicons. A molecular weight ladder 100bp/1Kb (NEB) was also loaded. The loaded gel was run at 100 v-500mA for ~30 minutes using Labnet Enduro E0303 Model 300V Power Supply. A UV transilluminator NuGenius (5 million pixel camera, lens of 8 – 48mm f/1.2 and a sensor of 1/2.5 inch) was used visualise the gel. The size(s) of DNA or RNA was determined using the molecular weight ladder (100bp/1Kb).

Following this step, the rest of the attB-PCR amplicons was purified using E.Z.N.A.® Cycle Pure Kit (Omega bio tek) according to the manufacturer protocol to get higher recovery of the purified amplicon.

5.3.3.5 Quantification of nucleic acids

NanoDrop™ 2000c UV-Vis spectrophotometer (Thermo Scientific) with its NanoDrop™ 2000/2000c software (Thermo Scientific) was used to quantify and measure the concentration of the RNA, cDNA and vectors. For the RNA the least concentration required was 70 ng/µl in order to synthesis the cDNA. The attB-PCR amplicons vary between 100 and 150 ng/µl as concentration; however, the purified vectors were with very high concentrations and required dilution to proceed with the subsequent steps. 1 µl of the blank solution (buffer solution used to elute the DNA, RNA or cDNA) was pipetted onto the pedestal, then the same amount for the samples as well. The software then calculated the concentration of RNA or cDNA based on the absorbance of 50nm for DNA and 40nm for RNA and the default extinction coefficient. Spectral measurements of the sample were taken and unloaded as excel sheets ready for the cDNA synthesis calculation.

5.3.4 Bacterial transformations

5.3.4.1 Preparation of chemically competent *E. coli* cells (DH5 α)

The DH5 α cells were prepared for chemical transformation of >50 ng plasmid DNA and One Shot[®] ccdB Survival[™] 2 T1R (Invitrogen) strains with transformation efficiencies ≥ 108 cfu/ μ g were used. In order to make the cell chemically competent, they were grown in 2 ml LB medium (without antibiotics) overnight with shaking (200 rpm) at 37°C. The OD₆₀₀ was adjusted to 0.25-0.3 by adding the overnight culture to 100 ml LB medium. The cell suspension was placed on ice for 15 minutes and then centrifuged in a 4°C for 10 minutes at 3,000 g. Pre-chilled filter sterilised 0.1 M calcium chloride (CaCl₂) was added to suspend the pellet and placed on ice for 30 minutes before centrifuging it again for 10 minutes at 4°C at 3,000g then the pellet was suspended in 6 ml of pre-chilled filter sterilised 0.1 M CaCl₂ containing 15% (v/v) glycerol and 25 μ l aliquots were pipetted into 1.5 ml centrifuge tubes, flash-frozen in liquid nitrogen and stored at -80°C.

5.3.4.2 Preparation of chemically competent *A. tumefaciens* cells

The competent *A. tumefaciens* cells (GV3101) were prepared as follows: the bacteria carrying the helper Ti plasmid pMP90RK was inoculated in 5 ml of LB medium containing rifampicin (50 μ g/ml) and gentamycin (50 μ g/ml) and grown overnight with shaking (200 rpm) at 28°C. 2 ml of the overnight culture was then added to 50 ml of LB medium (without antibiotics) and grown until the OD₆₀₀ reached 0.5-1.0. The cell suspension was chilled on ice for 15 minutes and then centrifuged at 3,000 g for 10 minutes at 4°C. The cell pellet was re-suspended in 1 ml of cold 20 mM Calcium chloride (CaCl₂). 50 μ l aliquots were pipetted into 1.5 ml micro-centrifuge tubes, flash-frozen in liquid nitrogen and stored at -80°C until required.

5.3.4.3 Transformation of competent *E. coli* cells (DH5 α)

For transferring the vector to the *E. coli* cells, a 50 μ l aliquots of chemically competent *E. coli* cells were taken from -80°C freezer and thawed on ice for 10 minutes. 2.5 μ l of Gateway[®] BP or LR recombination reaction or 0.1-1.0 μ g vector was added to the cells and the transformation reaction was placed on ice. After 30 minutes on ice, the reaction was heat-shocked at 42°C for 40-45 seconds and then returned to the ice for 2 minutes. 1 ml of LB medium was added to the cells, then placed in a 37°C shaker at 200 rpm for 1 hour. Two platings were performed one with 100 μ l the other plate with 200 μ l of transformation

reaction on LA medium supplemented with kanamycin-50µg/ml antibiotic. Plates were incubated overnight at 37°C.

5.3.4.4 Transformation of competent *A. tumefaciens* cells (GV3101)

Agrobacterium tumefaciens cells (GV3101) were transformed as following, 0.5 – 1 µg of plasmid DNA was added into a 50 µl of the *A. tumefaciens* aliquot after thawing them on ice for 5 min then pipet up and down to mix the vector before snap freeze the cells again in liquid nitrogen. The transformation mixture was then incubated at 37°C for 5 min. 400 µl of LB medium was added to the mixture and then incubated at 28°C with gentle shaking (200 rpm) shaker for 2-4 hours. The cells were then centrifuged for 30 seconds at 5000 rpm and the supernatant was discarded. Transformed cells were re-suspended with the LB medium and then spread on LB agar plate with (Rifampicin-50µg/ml, Gentamycin-100 µg/ml, spectinomycin-100 µg/ml) antibiotics. The plates were incubated at 28°C for 2 days. The colonies were obtained after 2-3 days and identified by colony PCR.

5.3.4.5 Bacterial cultures and antibiotics for selection

Two strains of *E. coli* were used including, One Shot® ccdB Survival™ 2 T1R (Invitrogen) for the pDONRP4P1r vector and DH5α for the rest of the vectors. Only one strain was used for *Agrobacterium* (GV3101). *Escherichia coli* (*E. coli*) and *Agrobacterium tumefaciens* (*A. tumefaciens*) were cultured in Luria Bertani Broth medium (LB Broth) supplied by the department of genetics consisting of 1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract and 1.5% (w/v) bacto-agar pH 7.

Table 5-3: Antibiotics concentration used for selection in *E. coli* and *A. tumefaciens*

Antibiotic	Supplier	Final concentration	
		<i>E.coli</i>	<i>A.tumefaciens</i>
Ampicillin sodium	Melford Laboratories	50 µg/ml 135 µM	50 µg/ml 135 µM
Carbenicillin disodium	Sigma-Aldrich	50 µg/ml 118 µM	50 µg/ml 118 µM
Chloramphenicol	Sigma-Aldrich	25 µg/ml 77 µM	25 µg/ml 77 µM
Gentamicin sulphate	Melford Laboratories	n.a	50 µg/ml 87 µM
Kanamycin monosulphate	Melford Laboratories	50 µg/ml 86 µM	50 µg/ml 86 µM
Rifampicin	Melford Laboratories	n.a	50 µg/ml 61 µM
Spectinomycin dihydrochloride	Melford Laboratories	100 µg/ml 202 µM	100 µg/ml 202 µM

The bacterial strains were stored in glycerol ready to use. A single bacterial colony was inoculated in 5 ml LB medium with suitable antibiotics and left to grow overnight at the required temperature. 700 µl aliquot of the overnight bacterial culture was transferred to Eppendorf tube and mixed with 300 µl sterilised 50% (v/v) glycerol. The cells were flash-

frozen in liquid nitrogen and stored at -80°C. The strains were recovered by streaking a small portion of the frozen culture on agar LB medium with appropriate selection.

5.3.5 cDNA and promoters manipulation, cloning, and sequencing

5.3.5.1 Gateway® recombination cloning PCR

The Gateway® BP cloning System (Invitrogen) was an efficient technique to clone our potential target gene promoters and for the selected transcription factors. The attB adapter primer was designed attached to the full-length attB sites (attB1-2 for TFs and attB4-1 for the promoters) to facilitate the manipulation of the region of interest then all the coding sequence (CDS) region of the transcript was amplified. However, for the potential target promoters 2kb upstream was amplified after confirming the CARG-box sites in them, resulting in an attB-PCR amplification product. These sites were added to attach to their complementary sequence in the Gateway® expression clones. The BP reaction was performed using the pDONR221 as donor vector (see appendix, Figure S8) which contains the resistant to kanamycin and the toxin-antitoxin system *ccdB* region enclosed with attP1-attP2 sites, to clone the transcription factors CDS and the green Renilla luciferase coding region.

The donor clone pDONR™ P4-P1r contains the attP4-1 site and resistance to kanamycin, was used to clone promoters amplified using attB4 att attB1 adapters.

The Gateway® BP recombination reaction

Table 5-4: The BP recombination reaction components and amounts.

BP reaction components	Amount
<i>attB</i> -PCR amplicon	125 ng
pDONR (221/P4-P1) vector	75 ng
BP Clonase™ II enzyme mix	1 µl
TE buffer	Up to 5 µl

This BP reaction step result in entry clone pENTRE-gene that carry our region of interest framed in between attL1 and attL2 sites and transformed in *E. coli* cells.

Gateway® LR recombination reaction to generate expression clones

Table 5-5: the LR recombination reaction components and amounts.

One-site LR recombination reaction	Amount
pDONR221(CDS/ Renilla)	125 ng

Destination vector (pB2GW7.0)	75 ng
LR Clonase™ II enzyme mix	1 µl
TE buffer	Up to 5 µl

The LR++ reaction used to clone two sites simultaneously, promoter region followed by the Renilla gene.

Table 5-6: Two-site LR recombination reaction components and amounts.

Two-site LR recombination reaction	Amount
pDONR221- Firefly	5 fmol
pDONRP4-P1r- promoter	5 fmol
Destination vector (pB7m24GW.3)	10 fmol
LR Clonase™ II Plus enzyme mix	1 µl

The mass (ng) of vector DNA required for the appropriate fmols was calculated using the following equation:

$$\text{Mass in ng} = (a \text{ fmol}) \times (N) \times (660 \text{ fg} / 1 \text{ fmol}) \times (1 \text{ ng} / 106 \text{ fg})$$

Where “a” is the number of fmol required and “N” is the length of vector DNA in bp.

5.3.5.2 Colony PCR (cPCR)

As a confirmation of cloning the right region in the right clone vector we performed colony PCR on 20 colonies generated from the transformation using designed primers; the forward was in the vector starting a few base pairs upstream of the insert, the reverse was the same used to amplify the PCR amplicon. Single colonies were picked up using a sterile pipette tip and dipped into 10 µl of PCR reaction mix. Cells remaining on the pipette tip were suspended onto appropriate antibiotic LA medium and grown in the required temperature.

5.3.5.3 Vector extraction and purification from bacteria

Isolation of the vector from *E. coli* cells was carried out using E.Z.N.A.® Plasmid DNA Mini Kit (OMEGA bio-tek) according to the manufacturer’s instructions. The culture was developed overnight from an isolated colony within the 5ml LB medium including the relevant antibiotics. A 1.5 ml of the overnight culture was transferred into an Eppendorf tube and centrifuged for 1 minute at 10000xg, the supernatant was discarded and the pellet was suspended and vortexed in (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 10 µg/ml RNase A) then 250 µl of the denaturation solution was added (0.2 M NaOH, 1% SDS) and incubated for 3 minutes at 55°C. Then neutralisation solution (3 M Guanidine-HCl, pH 4.8) was added and centrifuged at 13000xg for 10 minutes. The supernatant was transferred to

HiBind DNA Mini Column and centrifuged for 60 seconds. The column was flushed with 500 µl of 50% guanidinium chloride and 50% propan-2-ol then two washes with 700 µl of (10 mM Tris-HCl, pH 7.0, 50% Ethanol) before eluting the purified vector in 20-30 µl of DNase free water.

5.3.5.4 Sequencing of the region of interest

Purified vectors were sent to GATC Biotech (Eurofins Genomics) for sequencing the clones. In pDONR entry clones, M13 primer pairs forward and reverse were used to cover the full length of the inserted sequences. 80 - 100 ng/µl of the vector with a concentration of 5 µM (5 pmol/µl) from the primers in a total volume of 10µl was sent based on company requirements. Fasta format of the sequence was downloaded from the website (<https://www.eurofinsgenomics.eu>) and analysed with the FinchTV 1.5 software and compared to the reference *Brachypodium distachyon* v3.

5.3.6 Dual-luciferase transient expression assay

5.3.6.1 Transient transformation of *Nicotiana tabacum* leaf

The methodology described in the (Sparkes *et al.*, 2006) study was applied to transient transform the leaf of *Nicotiana tabacum* with *Agrobacterium tumefaciens* GV3101 (pMP90) carrying our transcription factor (TF) or promoter to test their interaction. Only one colony from the transformed *A. tumefaciens* was inoculated into 5ml of the LB medium with (Rifampicin-50µg/ml, Gentamycin-100 µg/ml, spectinomycin-100 µg/ml) antibiotics and this was subsequently incubated overnight at 28°C, along with continuous shaking of 200rpm. 1.5 ml from the overnight culture was centrifuged then washed twice with 1 ml of infiltration media (2mM Sodium Phosphate Dodecahydrate (Na₃PO₄·12H₂O), 0.1 mM acetosyringone, 50 mM 2-ethanesulfonic acid (MES) and 280mM D-glucose) in order to eliminate any trace to antibiotics. The density OD₆₀₀ of re-suspended cells was measured using Beckman Coulter, DU 730-UV/Vis spectrophotometer and the values were taken for each culture. The calculation was performed to achieve an OD₆₀₀ of 0.02 for the Control strains (p35-RenLUC) and 0.1 for the effector (p35-*BdMADS3*, p35*BdMADS7*), reporter (*BdMADS7*-FLUC, *Bd*-PEx-FLUC and *Bd*-MT-FLUC) and positive control (AT-DUO-GFP) strains. The four to six-week-old *Nicotiana tabacum* plants were kindly grown by June Saddington (Corporate Services, Gardens) and just prior to the infiltration they were placed under white fluorescent light for one hour to ensure that their stomata were fully opened.

Table 5-7: Optical density measured at a wavelength of 600 nm required for each strain and the volume taken from them to infiltrate the *Nicotiana tabacum* leaves.

	construct	Final DO ₆₀₀	Initial DO ₆₀₀	Volume
Control strains	p35S-Ren-Luc	0.02	1.7	47 µl
Effector strain	p35S-MADS3	0.1	1.6	250 µl
Effector strain	p35S-MADS7	0.1	1.2	333 µl
Reporter strain	Bd-MADS7-FLuc	0.1	1.8	222 µl
Reporter strain	Bd-Pex-FLuc	0.1	1.6	250 µl
Reporter strain	Bd-MT-FLuc	0.1	2.0	200 µl
Positive control protein	AT-DUO-GFP	0.1	1.7	235 µl
Positive control promoter	AT-DUO-FLuc	0.1	1.8	222 µl

Table 5-8: infiltration mixture containing the effector strains and the reporter strains.

	p35S::CDS	Volume	p35S::CDS	Volume	Promoter-FLuc	Volume	p35S::Ren-Luc	Volume	H ₂ O	total
1					Bd-MT-FLuc	200 µl	p35S-Ren-Luc	47 µl	3753 µl	4000 µl
2					Bd-Pex-FLuc	250 µl	p35S-Ren-Luc	47 µl	3803 µl	4000 µl
3					Bd-MADS7-FLuc	222 µl	p35S-Ren-Luc	47 µl	3775 µl	4000 µl
4	p35S-MADS7	333 µl	p35S-MADS3	250 µl	Bd-Pex-FLuc	250 µl	p35S-Ren-Luc	47 µl	3120 µl	4000 µl
5	p35S-MADS7	333 µl	p35S-MADS3	250 µl	Bd-MT-FLuc	200 µl	p35S-Ren-Luc	47 µl	3170 µl	4000 µl
6	p35S-MADS3	250 µl	-	-	Bd-MADS7-FLuc	222 µl	p35S-Ren-Luc	47 µl	3481 µl	4000 µl
7	p35S-MADS3	250 µl	-	-	Bd-Pex-FLuc	250 µl	p35S-Ren-Luc	47 µl	3453 µl	4000 µl
8	p35S-MADS3	250 µl	-	-	Bd-MT-FLuc	200 µl	p35S-Ren-Luc	47 µl	3503 µl	4000 µl
9	AT-DUO-GFP	235 µl	-	-	AT-DUO-FLuc	222 µl	p35S-Ren-Luc	47 µl	3496 µl	4000 µl

The 3rd and 4th large leaves, from the apical meristem of the plant, were chosen for infiltration, which was performed with 3 replicates in different leaves focusing on lamina of both sides of the midrib.

The mixture of the transformed *A. tumefaciens* were transferred to 1ml syringes and the abaxial side of the leaf prepped by pressing a 0.5cm² area with the solution to get rid of the cuticle. The end of the syringe was then placed to the rubbed regions then inject the *A. tumefaciens* mixture carefully to infiltrate into the plant with the aid of pressure by the finger. Infiltrated areas were marked and labelled. Gloves were sprayed with 70% (v/v) ethanol in between infiltrations to prevent cross-contamination. Plants were placed in a growth cabinet in 20°C in continues light and left for 2 days. Infiltrated regions were cut off and used in dual-luciferase assays.

5.3.6.2 The Dual-Luciferase® Reporter (DLR™) Assay System

Trans-activation and protein localisation in *Nicotiana tabacum* leaves cells were measured by the Dual-luciferase transient experiment assay in accordance with (Sherf *et al.*, 1996). Two days after the Agrobacterium infiltration and within the marked area, 7 mm were cut

using a cork borer. The $\sim 40\text{mm}^2$ circle samples taken from the infiltrated area were put into pre-chilled pestle and mortar for the of grinding for 30 seconds with 300 μl of 1x dilution of Passive Lysis Buffer (Promega) until the extracts seemed to be homogenous without any visible leaf pieces (the grinding was performed on ice). The mixture was transferred to an Eppendorf tube on ice then centrifuged at maximum speed at 4°C to precipitate all the cell debris and have a clear protein on the supernatant and kept on ice. For the luciferase assay, two reaction buffers were prepared according to (Dyer *et al.*, 2000); Renilla Luciferase and Firefly luciferase.

Table 5-9: Renilla Luciferase and Firefly luciferase buffers components. *FC: final concentration.

Firefly Luciferase Reaction Buffer		Renilla Luciferase Reaction Buffer	
Reagent FC*	Volume for 10 ml buffer	Reagent FC*	Volume for 10 ml buffer
25 mM Glycylglycine	1 ml (250 mM stock)	1.1 M NaCl	2.2 ml (5 M stock)
15 mM KPO ₄ (pH 7.6)	150 μl (1 M stock)	2.2 mM Na ₂ EDTA	100 μl (0.22 M stock)
4 mM EGTA	400 μl (100 mM stock)	0.22 M KPO ₄ (pH 5.1)	2.2 ml (1 M stock)
2 mM ATP	1 ml 20 mM stock)	0.44 mg/ml BSA	44 μl (100 mg/ml stock)
1 mM DTT	100 μl (100 mM stock)	1.43 μM Coelenterazine	14.3 μl (1 mM stock)
15 mM MgSO ₄	150 μl (1 M stock)		
0.1 mM CoA	100 μl (10 mM stock)		
75 μM Luciferin	1.5 ml (0.5 mM stock)		
dH ₂ O	5.6 ml	dH ₂ O	5.44 ml
pH adjust	1M KOH (8.0)	1M KH ₂ PO ₄	pH 5.0

25 μl from the supernatant of each sample was loaded into 96-well assay plate (one plate for each buffer) with two duplications on each plate (on ice) and two replicates from in-filtered *tabacum* leaf. Finally, the substrates Luciferase and Coelenterazine were added just prior to the reading which was performed with microplate reader FLUO star omega (BMG-LABTECH). The reading started by rinse twice the tubes with 2500 μl 70% ethanol followed by water then air to empty the canals before priming the microplate reader with Firefly Luciferase Reaction Buffer and start measuring the emission from each well using the Omega software. The second plate was treated the same but with the Renilla Luciferase Reaction Buffer for the second plate. The reads were downloaded from the Omega-data analysis software. A normalised dual of the luciferase actions (Fluc/Rluc) was determined

for each of the infiltrations through the process of divided the firefly luminescence action (Fluc) with that of the luminescence activity of Rluc.

5.3.7 RNA-seq plant material and growth

Both lines were planted in a phytotron under the following growth conditions: 51% humidity, 23 °C for 20 hours photoperiod, 4 hours darkness. Plants were watered by flooding under tray every other two days.

Eight libraries were prepared from the wild-type and the *Bdmads13* RNAi lines in two different developmental stages, very young inflorescence, where the expression starts and it believes that *BdMADS13* gene plays a key role in organ differentiation and specification, and from pre-anthesis carpel where *BdMADS13* is highly expressed. The total RNA was extracted using RNeasy Plant Mini Kit from QIAGEN. The quality of the RNA was assessed using NanoDrop™ 2000/c Spectrophotometers by Thermo Fisher Scientific and around 40 ng/μl was prepared for each sample. The RNA integrity number (RIN) was tested in NUCLEUS Genomics at the university of Leicester.

Table 5-10: Library preparation, quality and concentration sent for RNA-seq

Barcode	Sample name	Type	Source / Species	Sample Amount (μl)	Concentration (ng/μl)	RNA Integrity number RIN	Buffer
124329	Bd-WT-Vyinf-A	RNA	<i>Brachypodium distachyon</i>	26	38	8.7	water
124330	Bd-WT-Vyinf-B	RNA	<i>Brachypodium distachyon</i>	26	50	9	water
124331	Bd-WT-PrCa-A	RNA	<i>Brachypodium distachyon</i>	26	41	8.8	water
124332	Bd-WT-PrCa-B	RNA	<i>Brachypodium distachyon</i>	26	39	9.2	water
124333	Bd-M13i-Vyinf-A	RNA	<i>Brachypodium distachyon</i>	26	40	9	water
124334	Bd-M13i-Vyinf-B	RNA	<i>Brachypodium distachyon</i>	26	42	8	water
124335	Bd-M13i-PrCa-A	RNA	<i>Brachypodium distachyon</i>	26	45	8.7	water
124336	Bd-M13i-PrCa-B	RNA	<i>Brachypodium distachyon</i>	26	37	8.6	water

Eight purified total RNA samples were attached to the specific barcode and sent to GATC BIOTECH (A Eurofins Genomics Company) with the following requirements:

Standard-specific cDNA library synthesis using purification of poly-A containing mRNA molecules via poly-A selection, mRNA fragmentation, random primed cDNA synthesis (standard-specific), adapter ligation and adapter specific PCR amplification. The Sanger/Illumina 1.9 sequencing was performed using the following criteria; the run-type

was a single read, the read length was 50bp, the guaranteed was 30 million reads (+/-3%) per sample, and it was organised in FastQ files (sequence and quality scores).

Reads processing and analysis

FastQ files were received and downloaded for each sample with its corresponding file. Due to the size of the files, it has been downloaded in high performance computing platform called SPECTRE (Special Computational Teaching and Research Environment) high performance Computing Facility at the University of Leicester, which provides an interactive and batch Linux environment using no machine enterprise client for access.

The quality of the reads was assessed using FastQC (version 0.11.5)(Andrews & FastQC, 2015), which is a quality control tool for high throughput sequence data.

Conventional base calls were used in all samples with 0 sequences flagged as poor quality.

Table 5-11: samples files and their reads quality.

Sample name	File name NG-18508_	Total reads	Quality score	Read length (bp)	GC (%)
Bd-M13i-PrCa-A	Bd_M13i_PrCa_A_lib308252_6256_5_1.fastq	48610159	32-44	51	50
Bd-M13i-PrCa-B	Bd_M13i_PrCa_B_lib308253_6256_5_1_fastq	40902637	32-44	51	50
Bd-M13i-Vyinf-A	Bd_M13i_Vyinf_A_lib308250_6256_5_1_fastq	41544667	32-44	51	50
Bd-M13i-Vyinf-B	Bd_M13i_Vyinf_B_lib308251_6256_5_1_fastq	41255716	32-44	51	50
Bd-WT-PrCa-A	Bd_WT_PrCa_A_lib308248_6256_5_1.fastq	40543530	32-44	51	50
Bd-WT-PrCa-B	Bd_WT_PrCa_B_lib308249_6256_5_1_fastq	38312745	32-44	51	50
Bd-WT-Vyinf-A	Bd_WT_Vyinf_A_lib308246_6256_5_1_fastq	42164526	32-44	51	50
Bd-WT-Vyinf-B	Bd_WT_Vyinf_B_lib308247_6256_5_1.fastq	45080560	32-44	51	50

There was no need to use Trimmomatic (Bolger, Lohse and Usadel, 2014) to trim low-quality bases as all the samples showed high-quality score. STAR software package (2.5.2b) (Dobin *et al.*, 2013) was performed to generate reference index file for each sample this step is essential as STAR requires an index of the reference genome to allow fast access to the reference during the mapping stage. After, STAR aligns RNA-Seq reads to the latest version of the reference genome of *B. distachyon* accessions *Bd21* (*Bd21*, JGI v3.0 assembly, JGI v3.1 annotation) obtained from EnsemblPlants database (FASTA file) (Hunt *et al.*, 2018), as a result a list of SAM (Sequence Alignment Map) files generated from mapping all the reads to the reference genome. The output of the previous step (SAM files) was converted to BAM files which are simply a binary version, sorted and indexed using Samtools package (1.3.2) (Li & Durbin, 2009). The resultant bam aligned reads files were

assembled and quantified using featureCounts in Bioconductor Rsubread (Liao *et al.*, 2019) and a table of raw counts was produced for all the samples.

RNA-seq differential expression:

The raw counts' reads were normalised and analysed using edgeR v3.9 from Bioconductor package (Robinson, McCarthy and Smyth, 2010) (McCarthy, Chen and Smyth, 2012) by cpm function which computes log2 counts/million, in order to estimate differentially expressed genes. The list of DEGs was generated based on FDR-adjusted p-value threshold of < 0.05 and fold change >1 . The following step identified overlapping differentially expressed genes between the two stages, the very young inflorescence and the pre-anthesis carpel.

Gene identification:

BLAST2GO software (Götz *et al.*, 2008) was used to blast our DEGs sequences against the available databases to identify and annotate them. The cDNA sequences were loaded in Blast2GO v5, blast sequences using NCBI QBLAST (blastx-fast). Blast Expectation value = $1.0E-3$, Number of Blast Hit = 20, HSP Length Cutoff = 33, HSP-Hit Coverage = 0, then, InterPro, EMBL-EBI InterPro web-service was used to scan the sequences looking for families, domains, sites, repeats and other sequence features. Mapping, latest database Goa Version 2019.01 originates from the Gene Ontology Association and Uniprot's ID-Mapping. annotation, cutoff = 55, GO weight = 5, no filter by taxon, Blast filters, E-Value-Hit-Filter = $1.0E-6$, HSP-Hit Coverage = 0, Hit Filter = 500, Computational Analysis Evidence Codes was run with the default settings, orthology level = 2

Enrichment analysis (Fisher's Exact Test), filter value = 0.05, filter mode = FDR, Annotation = GO IDs, Biological Process, Molecular Function, Cellular Component.

PARTEK analysis

In order to validate the results generated from edgeR package we performed the same analysis using Partek Genome Suite 6.6 with the help of BBASH officer Dr Matt Blades. The workflow started by choosing RNA-Seq option then import the bam files, followed by an alignment against the new annotation of *Brachypodium*. Differentially expressed genes were identified from the data using analysis of variance, then applied a false discovery rate $FDR \leq 0.05$ cut-off to filter for significantly regulated genes. An attribute was created to compare samples, then cluster based on significant genes.

5.4 Results

5.4.1 Selecting the potential target genes for *BdMADS3* using co-expression analysis

5.4.1.1 *BdMADS3* co-expresses with genes involved in the ROS signalling regulation

The STRING database (<http://string-db.org>) uses meta-analysis approach to provide protein-protein Interaction networks as well as functional enrichment analysis that are annotated at various levels of detail in online resources by automated mining of scientific texts for interaction information, transferring this interactions from one model organism to the other and statistical enrichment information of any known biological function or pathway.

Using string analysis (Version: 11.0), we selected 13 genes in the network that co-express with *BdMADS3*, then using Davidson et al. 2012 expression read counts we calculated the K means clustering correlation coefficient that revealed some strong correlations between the expression of *BdMADS58* and *BdMADS13* with 0.869 (see table S6, appendix). This could indicate that members of C-class might interact with D-class. Moreover, there was a three mitogen-activated-protein kinase (BRADI2G15620, BRADI2G44350 and BRADI2G45010) highly correlated with *BdMADS3* with 0.9. These proteins (MAPK) are well known to intervene in the induction or activation of different transcription factors that regulate the ROS-scavenging and ROS-producing pathways (NADPH oxidase) in plant defence signalling (Mittler et al. 2004; Zhang and Klessig 2001; Jalmi and Sinha 2015).

5.4.1.2 *MADS7* co-express with *MAD3* in *Triticum aestivum*, *Arabidopsis thaliana*, *Oryza sativa* and *Hordeum vulgare*.

GENEVESTIGATOR® is a meta-analysis database that use high-throughput gene expression analysis retrieved from microarray technologies experiment (affymetrix geneChip data, RNA-seq etc) from chosen environmental conditions, growth stages, or organs that are publicly available. In

Using the GENEVESTIGATOR® platform which contains several RNA-Seq studies, we performed a co-expression analysis for *MADS3* in four species calculating the Pearson's correlation coefficient (Figure 5-1).

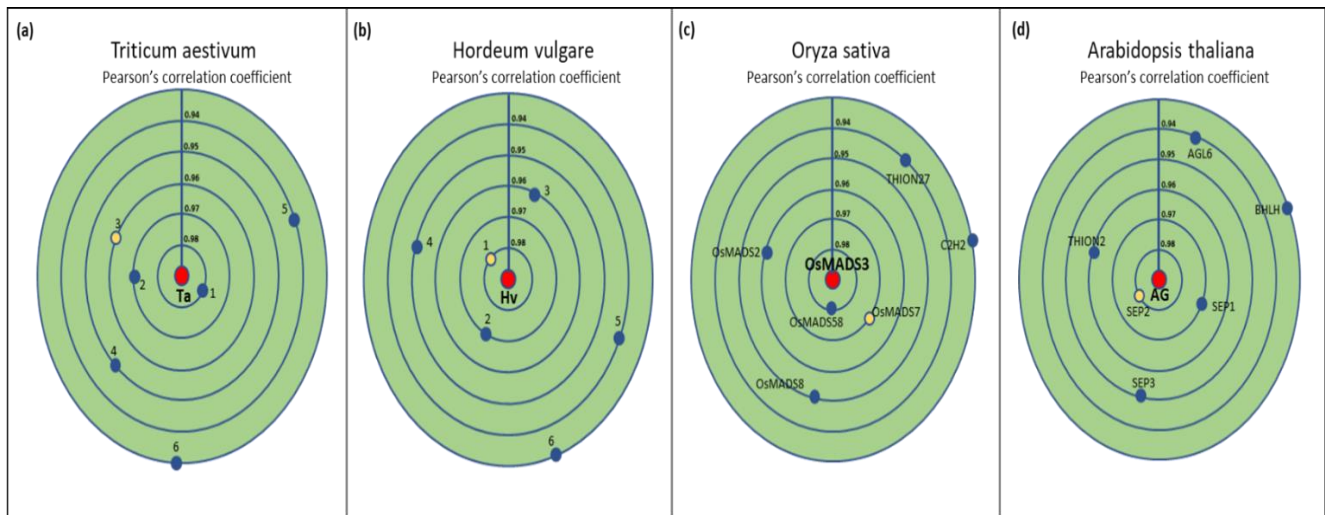


Figure 5-1: Co-expression analysis for *MADS3* in four species showing genes with correlation above 0.93, the central red circle is an orthologue of *MADS3*, and the yellow circle is an orthologue of *MADS7* in each species (*Arabidopsis*: *SEP2*, AT3G02310; rice: *OsMADS7*, Os08g0531700; Wheat: *TaMADS7*:

TraesCS7A02G260600; barley: *HvMADS7*, HORVU7Hr1G054220), the gene ID in barley and wheat was replaced by numbers (see appendix, figure S2 for detailed list), analysed using GENEVESTIGATOR® (Hruz *et al.*, 2008).

The analysis revealed that *BdMADS7* from E-class co-expresses with *BdMADS3* in all four species. In rice, the analysis illustrated that *OsMADS58*, *OsMADS2* and *OsMADS8* co-express with *OsMADS3* (Figure 5-1), which suggest that there is an interaction between different classes of *MADS-box* transcription factors to form quaternary complexes and perhaps activate genes from the same sub-family. The expression pattern of the thionins, and the *bHLH* gene appear to be close to the expression of *BdMADS3* in only rice and *Arabidopsis* (Figure 5-1 (c)(d)). It is known that thionins are one of the most essential proteins in plant defence against pathogens (Almaghrabi *et al.* 2019). Overall, only *MADS7* and its orthologs in the four species appears to co-express with *MADS3*, which indicates a strong interaction between them.

In wheat as well we performed a co-expression analysis using GENIE3 network which was created using 850 diverse RNA-seq samples from wheat and predicts transcription factor targets (Geurts, 2018). The analysis was accomplished by Dr Philippa Borrill following the method described in (Borrill *et al.*, 2019). The GENIE3 network construction calculates the connection between the transcription factor and their potential targets. In our analysis, *TaMADS3* expression pattern from 850 RNA-seq was predicted from the expression patterns of all the other genes (input genes) and this is taken as an indication of a putative regulatory link (Irrthum *et al.* 2010).

Analysis in GENIE3 indicated that *TaMADS3* (TraesCS3A02G314300, TraesCS3B02G157500, TraesCS3D02G140200) potential downstream targets were mitochondrial membrane translocase, Pollen Ole1 (Figure 5-2) (table S10, Appendix). *TaMADS58* (TraesCS1A02G125800, TraesCS1B02G144800, TraesCS1D02G127700) was tightly correlated with Photosystem II CP47, Photosystem I P700 chlorophyll, FBD-associated F-box protein and mitochondrial genes (table S11, Appendix). This suggest that C-class genes in wheat are related to genes involved in photosynthesis process. In the D-class genes, *TaMADS13* (TraesCS5A02G117500, TraesCS5B02G115100, TraesCS5D02G118200) was co-expressed with NPK1-related protein kinase 1, Tapetum determinant 1, Protein kinase-like protein, Peroxidase and Nucellin-like aspartic protease (table S12, Appendix). Its ortholog *TaMADS21* (TraesCS1A02G262700, TraesCS1B02G273300, TraesCS1D01G262700) was co expressing with Myb transcription factor, vacuolar-processing enzyme, Ras family protein (cellular signal transduction) and chloroplast ATP synthase (table S13, Appendix). These results associate the D-class gene in wheat to the PCD processes and the genes (Figure 2-12) are highly express during grain development. In order for the carpel to expand and start filling with endosperm some maternal tissues such as nucellus should degenerate and leave space for the new developing cells. This could explain the appearance of the vacuole-processing enzyme, peroxidase and nucellin-like aspartic protease that are directly involved in the PCD (Wang et al. 2009).

All 12 *MADS-box* genes are illustrated in figure 5-2. Higher numbers of the directed edge weight indicate a stronger connection; in this diagram, we could see that the closest genes were the one involved in the photosystem II CP47 located in the chloroplast. This could link our *MADS-box* genes to the two organelles; the chloroplast and the mitochondria since the ROS are primarily formed in chloroplasts, mitochondria, and peroxisomes (Mittler 2017).

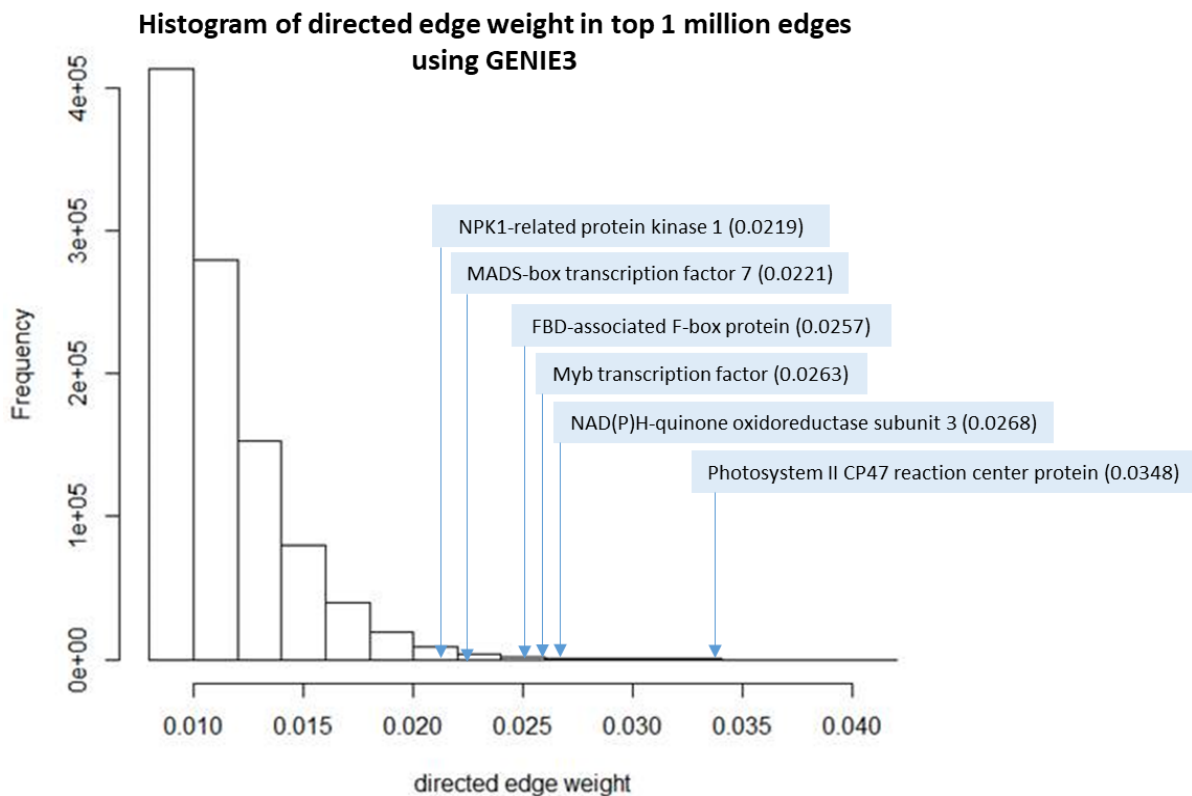


Figure 5-2: Directed edge weight in top 1 million edges *TaMADS3* co-expressed genes and their using GENIE3. Higher numbers indicate a stronger connection.

Furthermore, (Hu *et al.*, 2011) study in rice highlighted the role of the C-class gene *MADS3* in regulating late anther development and pollen formation by regulating ROS homeostasis in anthers through the ROS-scavenging protein metallothionein (Os03g0288000). Its ortholog in *Brachypodium* was selected *BdMT*: BRADI_2g28254v3. Another ROS-scavenging enzyme was linked to male sterility is peroxidase (Hu *et al.*, 2011), therefore, its ortholog in *Brachypodium* (*BdPex*: BRADI_2g12216v3) was selected to be tested as a potential target for *BdMADS3*.

Their promoters were scanned for the *MADS-box* domain binding motif GArG-box and the results are shown in the table below.

Gene name	Gene ID	CAR-G-Box motif	Start	End	Score
<i>BdMADS7</i>	BRADI_3g41260v3	ggaa CCAAAAATGG acgc	-847	-838	1
<i>BdMT</i> (metallothionein)	BRADI_2g28254v3	gata CCAAATAAGG tgtt	-324	-315	1
<i>BdPex</i> (peroxidase)	BRADI_2g12216v3	attg CCATAAATGG tccg	-1860	-1851	1

All three genes contained the CAR-G-Box motif in their promoter in different locations within the 1kb size.

5.4.2 *BdMADS3* interact with *BdMADS7* transcription factor to bind and activate metallothioneins (*MT*) gene promoter

To investigate the capacity of *BdMADS3* and *BdMADS7* to activate the promoter of metallothioneins, peroxidase and *BdMADS7*, transient transcriptional activation assays using the firefly luciferase (FLuc) reporter gene were performed. In these assays, the promoter of our three potential target genes (*BdMADS7*, *BdMT* and *BdPex*) was cloned upstream of the FLuc CDS to generate a reporter construct, while the full CDS of *BdMADS3* and *BdMADS7* was cloned downstream of the P35S promoter. These constructs were transferred to *Agrobacterium* then transiently transformed into tobacco leaf epidermal cells with and without a transcription factor (effector) by agroinfiltration. If the transcription factor binds to the promoter of the potential target genes the measurable level of FLuc increases and is then quantified using the light emission luciferase substrate and the difference in FLuc activity between reporter alone and reporter and effector is calculated to give a readout of the strength of transactivation. A construct containing the Renilla luciferase gene driven by 35S promoter was infiltrated as well in order to normalise for transformation efficiencies between biological replicates.

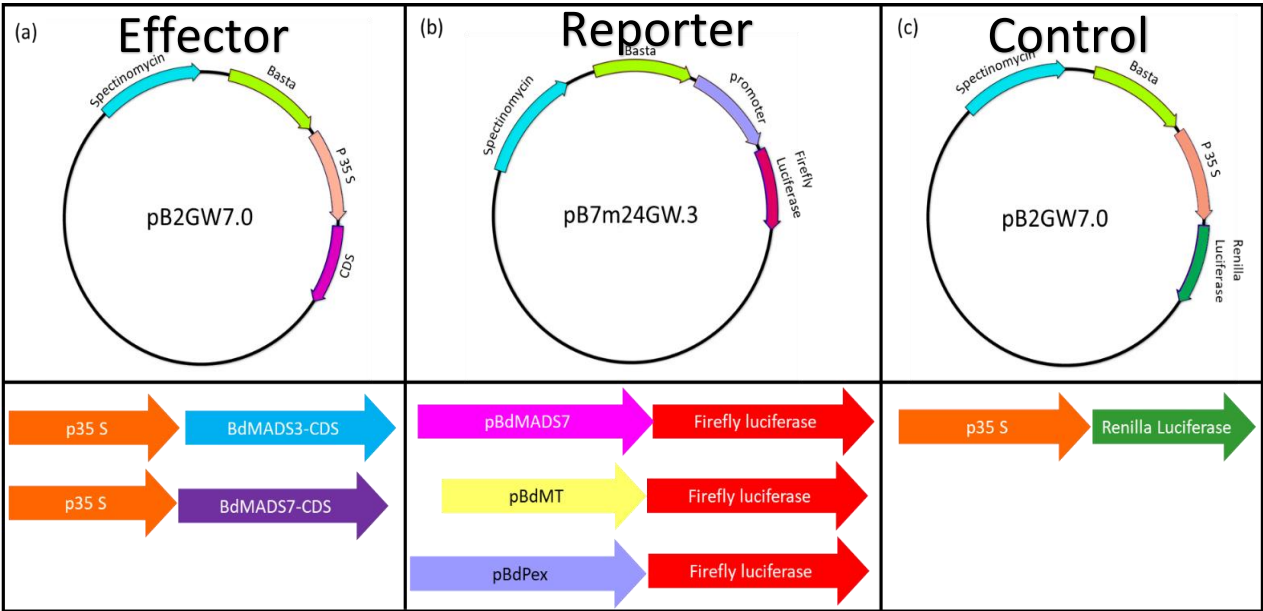


Figure 5-3: Plasmid and genes used to construct vectors.

A combination of constructs were prepared in order to test the activation of the three promoters (Figure 5-3). We used p35:S-*BdMADS3* with individual promoters and the same was performed with p35:S-*BdMADS7*. As well as using the two effectors constructs

combined to test if they will be any formation of a quartet or any of multimeric interactions.

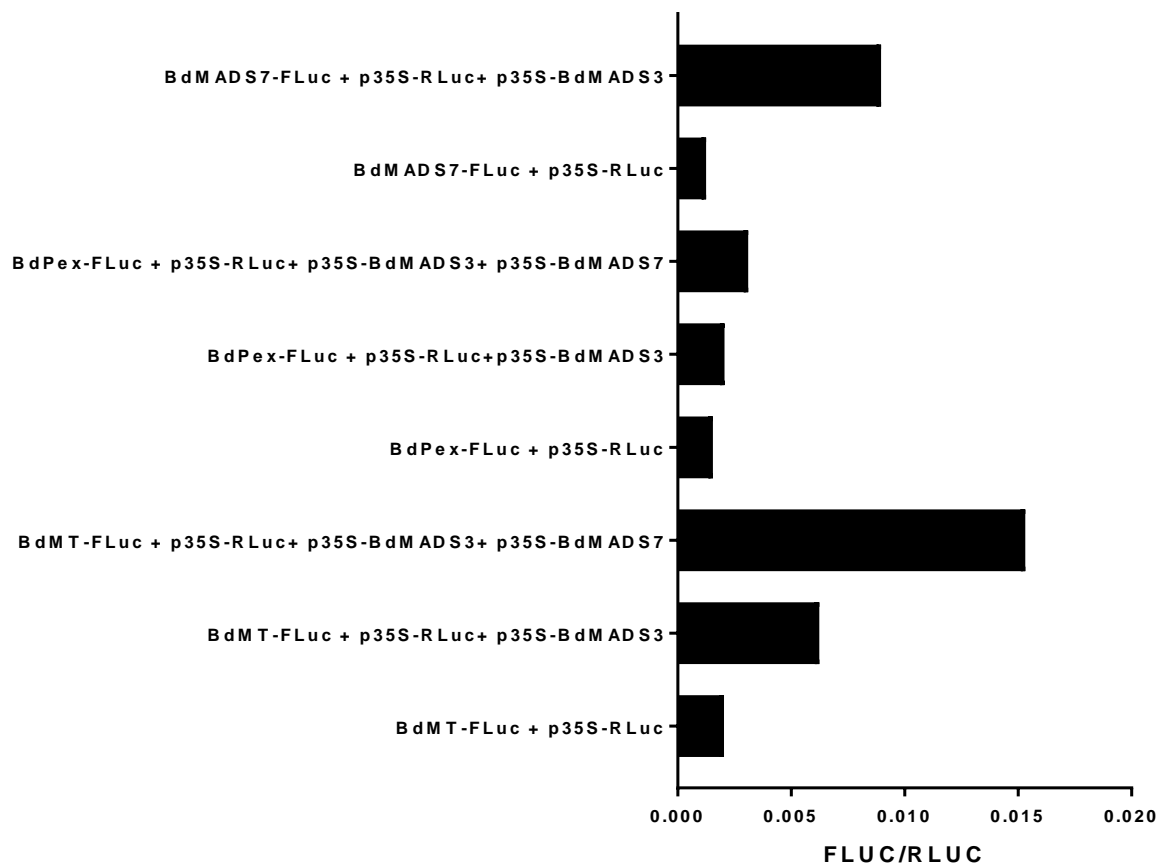


Figure 5-4: Transactivation assays using the luciferase reporter gene to report promoter activity in tobacco leaves. Performed only once with four well replicats.

Results from the above series of experiments (Figure 5-4) suggest that there was no activation of the peroxidase promoter by *BdMADS3* and *BdMADS7* transcription factors. This result is combined with the location of the CArG-box motif found in the peroxidase promoter at -1851bp that has been considered very far from the gene. However, the member of the C-lineage *BdMADS3* transcription factor could activate *BdMADS7* gene that is considered as E-class. Whereas, *BdMADS3* alone could not fully activate the promoter of the metallothionein gene (Figure 5-4).

It is only when we add *BdMADS3* together with *BdMADS7* that we could see that they are a transcriptional activator of the metallothionein gene (Figure 5-4).

Collectively these results propose that *BdMADS3* and *BdMADS7* form quaternary complexes that bind to the CArG-box motif of the metallothionein gene, which mediates ROS balance during oxidative stress.

5.4.3 Differential gene expression analysis using high-throughput RNA sequencing (RNA-seq)

Eight libraries were prepared from the wild-type and the *Bdmads13-3* RNAi lines in two different stages. The very young inflorescence; where the expression starts and from pre-anthesis carpel where *BdMADS13* is highly expressed (Figure 3-11). This analysis aims to identify the lists of genes that are down and up-regulated by *BdMADS13*.

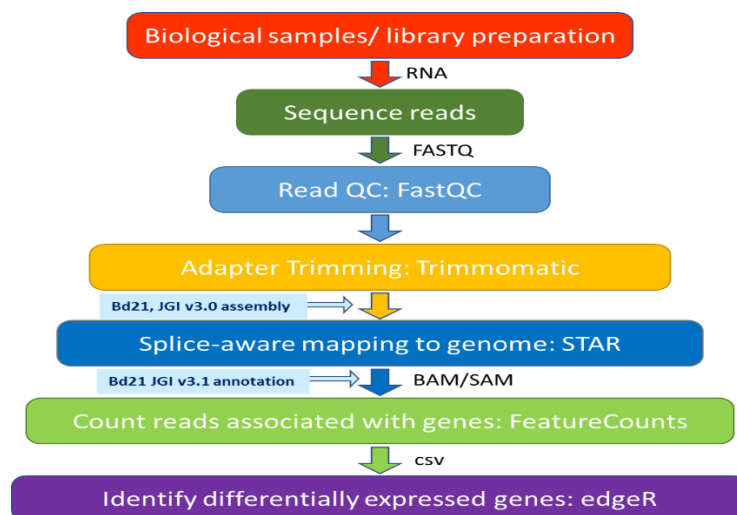


Figure 5-5: A workflow diagram of methods for identifying the normalised count of the transcriptome in the wild type and differentially expressed genes between the *Bdmas13* RNAi and the wild type.

The flowchart (Figure 5-5) illustrates the significant steps performed to generate the normalised read counts. The reads were aligned and annotated using the latest available *Brachypodium* assembly (V3.1). The abundance estimation of the transcript was calculated using featureCounts, then an CSV file of the expression count was generated using Rsubread.

5.4.3.1 Reads processing and analysis:

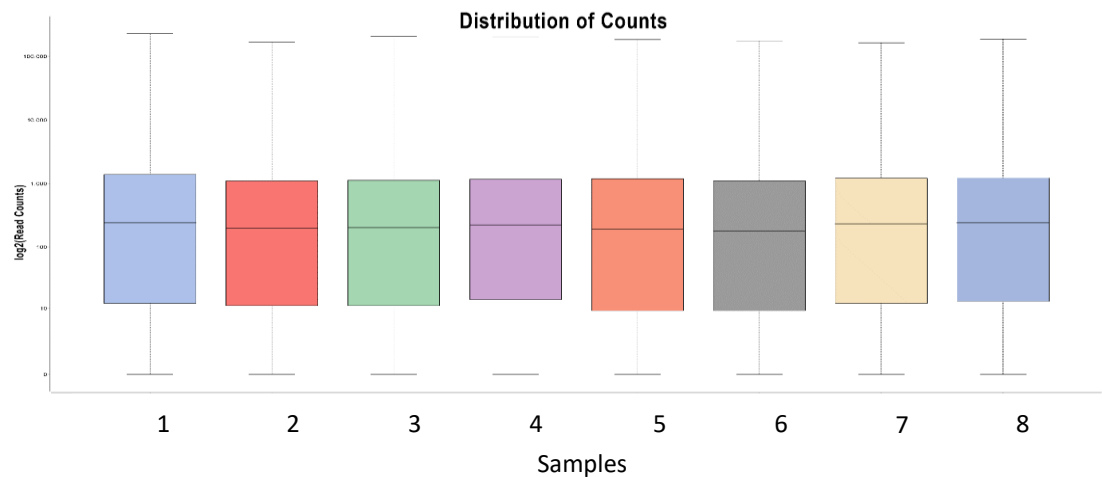


Figure 5-6: The read counts of the transcripts and their distribution among the eight samples. (1-2) replicates of *Bdmads13-3* RNAi PrCa, (3-4) replicates of *Bdmads13-3* RNAi Vy, (5-6) replicates of *BdWT* PrCa, (7-8) replicates of *BdWT* Vy.

The read counts vary from 32 to 48 million reads of 51 bp in length, with high-quality scores for all the RNA libraries (Figure 5-6).

5.4.3.2 The identification of Differentially Expressed Genes in RNA-seq Data:

The raw read counts were normalised and analysed using edgeR v3.9 to further estimate differentially expressed genes. The below Venn diagram (figure 5-7) illustrate the total differentiated genes in each tissues sample and the overlapping genes numbers.

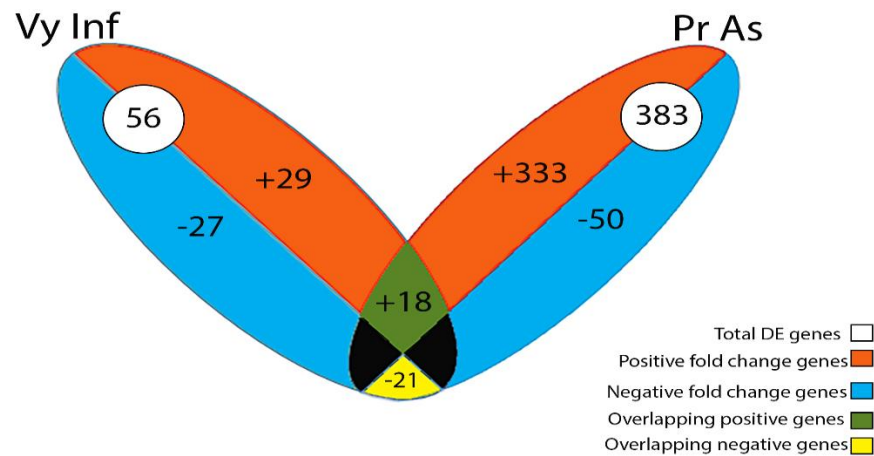


Figure 5-7: Venn diagram of differentially expressed genes in very young inflorescence (Vy Inf) and pre-anthesis carpel (Pr As). The sum of numbers in each big oval in white circle is the total number of differentially expressed genes in each sample, divided into up and down-regulated genes, and the overlapping areas are the number of common differentially expressed genes among the comparison groups, positive fold change genes are the genes that are up-regulated in *Bdmads13* RNAi line and the negative fold-change genes are down-regulated.

The statistical analysis revealed that a total of 56 genes were significantly differentially expressed: 27 genes downregulated and 29 upregulated in very young inflorescence sample compared to wild-type. The Pre-anthesis carpel sample presents a total of differentially expressed genes of 383, 333 from them are up-regulated, and 50 are downregulated in the *Bdmads13-3* RNAi line (see appendix table S14, S15).

5.4.3.3 The Expression of differentially expressed genes (DEGs)

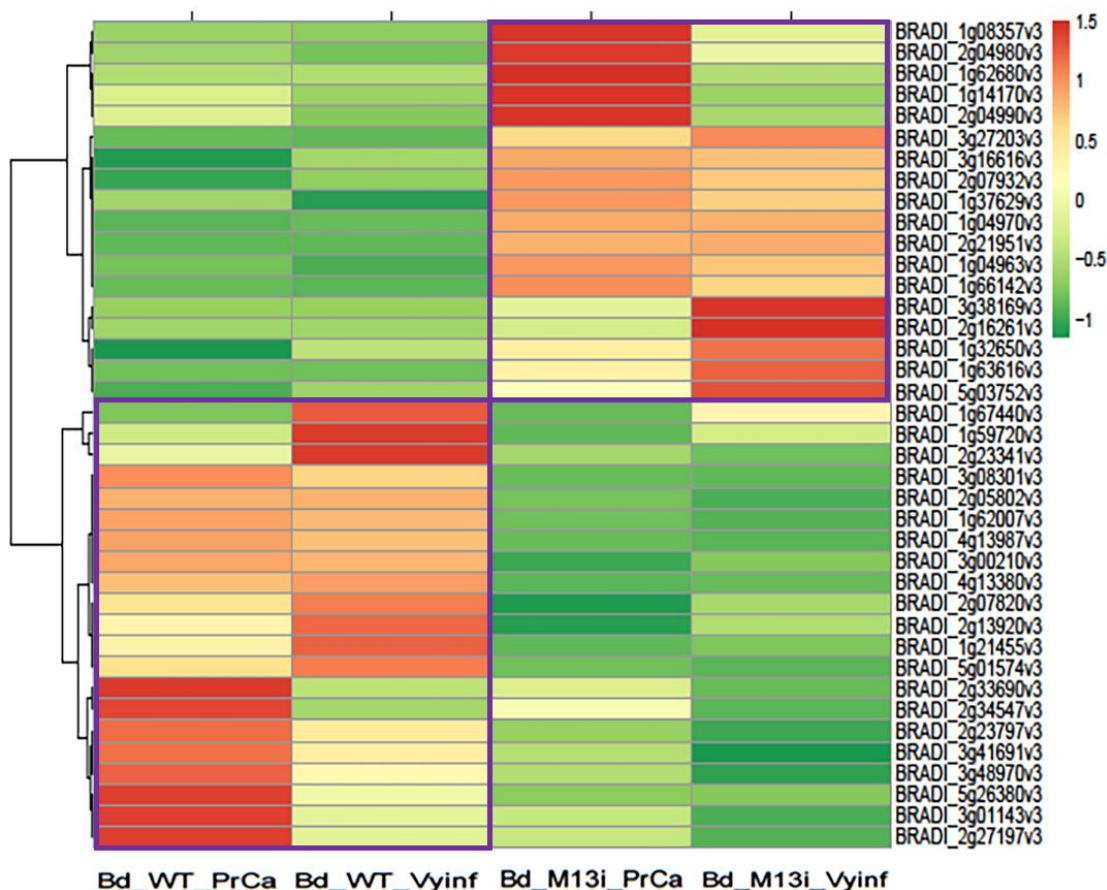


Figure 5-8: Heat-map of the 39 identified differentially expressed genes (DEGs) (18 up-regulated and 21 down-regulated) selected using edgeR and illustrated using clustvis (<https://biit.cs.ut.ee/clustvis>). The data was transformed, and hierarchical agglomerative clustering was applied the data as well as row scaling.

Differentially expressed genes were identified using edgeR and the overlapped genes between two stages, the very young inflorescence and the pre-anthesis carpel, were selected. In this 39 overlapping genes, 18 were up-regulated (Figure 5-8; right purple rectangle), and 21 were down-regulated (Figure 5-8; left purple rectangle) in the *Bd_mads13* RNAi line.

Hierarchical clustering and heat-map representation of the sample's expression levels, demonstrated a clear distinction between the wild-type and the *Bd_mads13* RNAi line

(Figure 5-8). In order to identify the function of these genes, we implemented a GO term analysis.

5.4.3.4 Functional annotation of differentially expressed genes:

BLAST analysis and GO term annotation was performed to comprehend the functions of these differentially expressed genes. GO term enrichment analysis was used to have a profound understanding of these genes and the biological processes that require them. BLAST2GO v5 (Götz *et al.*, 2008) uses an efficient strategy for analysing the representation of genes in different categories. First, it BLASTs the sequences to identify the description of differently expressed genes then GO term enrichment analysis annotate the sequences and categorised them into three main categories: biological process (BP), molecular function (FM) and InterProScan domain ID (IPS ID). The sequences were subsequently re-analysed and validated using Pfam groups, just to confirm there protein family.

5.4.3.5 *BdMADS13* putative up-regulated genes

The analysis was successful in annotating 14 genes from the up-regulated genes identified in the *Bdmads13* RNAi line. table 5-12 illustrates a wide range of molecular functions and biological processes that these genes are involved in.

Table 5-12: Gene Ontology (GO) of the *BdMADS13* putative up-regulated 18 genes represented in InterProScan IDs (IPS ID), molecular function (MF) and biological process (BP) annotation. Generated using BLAST2GO (Götz *et al.*, 2008).

InterProScan ID	Molecular function	Biological process
Chloramphenicol Acetyltransferase	Transferase activity, transferring acyl groups	DNA metabolic process
Transferees	Nucleic acid binding	Chromosome organization
Ribonuclease	Carbohydrate binding	Maltose metabolic process
Glucosyl hydrolase	Protein serine/threonine kinase activity	Proteolysis involved cellular protein catabol
Dormancy /auxin associated protein	Helicase activity	Protein phosphorylation
Papain-like cysteine endopeptidase	Cysteine type endopeptidase activity	RNA phosphodiester bond hydrolysis endonucleolytic
Zinc finger CCHC-type	RNA/DNA hybrid ribonuclease activity	Cellular response to DNA damage stimulus
Dilute domain	Alpha-1-4-glucosidase activity	
Glycoside hydrolase 31	Maltose alpha-glucosidase activity	
Peptidase C1A, papain	ATP binding	
Zinc finger, CCHC-type	Nucleotide binding	
Papain-like cysteine peptidase	Zinc ion binding	
Glycoside hydrolase		
Aspartic peptidase		

InterproScan showed that four genes represented in the up-regulated list were transferase (IPR003480); glutathione S-transferase T3 and chloramphenicol acetyltransferase domains, followed by three peptidases: peptidase C1A papain C-terminal, papain-like cysteine peptidase superfamily and aspartic peptidase domain superfamily. The zinc finger, CCHC-type superfamily was represented by two genes, and another three genes were for glycoside hydrolase family and glycosyl hydrolase. Ribonuclease H domain was illustrated by one gene which are endonuclease enzymes that catalyse the cleavage of RNA or DNA substrates via a hydrolytic process (table 5-13). In the molecular function category, the DEGs were illustrated by a high number of genes with transferase activity and nucleic acid binding, followed by serine/threonine kinase activity, helicases, glucosidases, endopeptidases and zinc-binding activities. In the biological process category, the genes were further classified in DNA metabolic processes, maltose metabolic processes, chromosome organisation, protein phosphorylation and cellular response to DNA damage stimulus (table 5-13).

Overall, the genes identified in table 5-13 were related to programmed cell death such as papain and aspartic peptidase and also DNA helicase enzyme which in this case represent a cellular response to DNA damage stimulus by unwinding the DNA double helix through breaking the hydrogen bonds down the centre of the strand. These reactions need energy that is provided in this list by a high expression of glycoside hydrolase and glycosyl hydrolase genes (table 5-13).

Table 5-13: GO term annotation of the overlap (Vyinf and PrAs samples) list of the genes that are up-regulated in *Bdmads13* RNAi. IEA (Inferred from Electronic Annotation).

Sequence Name	Sequence Description	Annotation GO Term	Enzyme Name	InterPro GO Term
BRADI_2g16261v 3	EPIDERMAL PATTERNING FACTOR-like protein 2			
BRADI_1g66142v 3	uncharacterized protein LOC104582872			
BRADI_5g03752v 3	helicase-like protein	nucleotide-binding DNA helicase activity DNA metabolic process cellular response to DNA damage stimulus	Nucleoside-triphosphate phosphatase	
BRADI_1g14170v 3	dormancy-associated protein 1			Dormancy/auxin associated protein
BRADI_1g32650v 3	alpha-glucosidase	alpha-1,4-glucosidase activity hydrolase activity, vacuole, cell wall carbohydrate-binding maltose alpha-glucosidase activity apoplast	Alpha-glucosidase	hydrolyzing O-glycosyl carbohydrate metabolic process; Glycoside hydrolase family 31; Glycosyl hydrolase
BRADI_2g21951v 3	protein CHROMATIN REMODELING 35	helicase activity ATP binding		
BRADI_1g63616v 3	hypothetical protein BRADI_1g63616v3	an integral component of membrane		
BRADI_3g27203v 3	hypothetical protein BRADI_3g27203v3			
BRADI_1g04963v 3	uncharacterized protein LOC100826907			Domain of unknown function DUF1618
BRADI_2g04990v 3	predicted protein	transferase activity, transferring acyl groups other than amino-acyl groups		Chloramphenicol acetyltransferase-like domain
BRADI_2g07932v 3	hypothetical protein BRADI_2g07932v3			
BRADI_2g04980v 3	Transferase family	transferase activity, transferring acyl groups other than amino-acyl groups		Transferase; Chloramphenicol acetyltransferase-like domain
BRADI_1g08357v 3	senescence-specific cysteine protease SAG39	cysteine-type endopeptidase activity extracellular space lysosome proteolysis involved in cellular protein catabolic process	Acting on peptide bonds (peptidases)	proteolysis/cysteine-type peptidase activity; Peptidase C1A, papain C-terminal; Papain-like cysteine endopeptidase
BRADI_1g04970v 3	retrotransposon protein, putative, unclassified	nucleic acid binding; RNA-DNA hybrid ribonuclease activity; RNA phosphodiester bond hydrolysis, endonucleolytic	Acting on ester bonds; Acting on ester bonds; Ribonuclease H	nucleic acid binding/RNA-DNA hybrid ribonuclease H activity
BRADI_3g16616v 3	uncharacterized protein LOC110437531	nucleic acid binding; zinc ion binding DNA integration		nucleic acid binding/zinc ion binding; Zinc finger, CCHC-type; Aspartic peptidase domain superfamily
BRADI_1g37629v 3	predicted protein	protein serine/threonine kinase activity ATP binding protein phosphorylation myosin complex	Transferring phosphorus-containing groups	Dilute domain
BRADI_1g62680v 3	Desiccation-related protein PCC13-62	Ferritin-like domain (Ferritin_2)		
BRADI_3g38169v 3	glutathione S-transferase T3-like			

The genes that *BdMADS13* could directly regulate should have at least one CARG-Box motif in its promoter. Therefore, a complete promoter motif scan was performed in the potential

targets genes (3Kb). The results from the overlapping genes that appear in the Very young inflorescence and Pre-anthesis carpel showed that only seven genes presented the *MADS-box* binding domain in one or two positions (table 5-14).

Two of the potential target genes that presented the CArG-box motif were genes with transferase activities (BRADI_2g04990v3 and BRADI_2g04980v3) (table 5-13 and table 5-14), transferring acyl groups and essential in many cell biological processes (Table 5-12). In addition, the protease gene papain presented two motifs (table 5-14) which suggest that *BdMADS13* might down-regulate this member of cysteine protease family C1 in order to keep the plant safe from protein degradation during flowering time. Moreover, one of the potential targets was Ribonuclease H domain (BRADI_1g04970v3) gene (table 5-14) that is involved in RNA-DNA hybrid cleavage activity and two genes that are responsible for catalysing the separation of duplex nucleic acids into single strands such as helicase protein (BRADI_5g03752v3) and chromatin remodelling (BRADI_2g21951v3)(Table 5-12 and Table 5-13). This supports the theory that *BdMADS13* is acting as a repressor of PCD genes in *Brachypodium*.

The genes that were up-regulated only in the Pre-anthesis carpel samples in the transgenic line gave more information about the processes that the plant is undergoing in the absence (transcript reduction) of *BdMADS13*. More than a 12 fold-change increase was observed in the *Bdmads13* RNAi line of the serine-type endopeptidase inhibitor activity (BRADI_2g39260v3) and between 8 to 10 fold-change increases in some RNA scavenging genes such as peroxidase (BRADI_2g11320v3), metallothionein (BRADI_4g09257v3) (appendix, table S14). Some peptidases also were increased in the plant by around 10 fold-change including cysteine-type endopeptidase and aspartic-type endopeptidase (appendix, table S14). Overall, these shows that the plant is under oxidative stress; therefore, it activates more ROS-scavenging and more endopeptidase inhibitors genes to maintain the plant homeostasis and prevent cell damage.

Table 5-14: Overlapping gene list (PrAs+VyIn) that are up-regulated with their fold change values generated from edgeR for differentially expressed genes. 3kb promoter was extracted from ensemble plant and submitted in RSAT (<http://rsat.sb-roscoff.fr/>) for CarG-Box (CC(A/T)₄NNGG) (Riechmann *et al.*, 1996) localisation.

	Seq ID	Log FC	Log CPM	P Value	FDR	CarG-Box motif	Start	End	Score
1	BRADI_1g62680v3	9.512403567	6.434794415	4.42E-22	5.71E-19	-	-3000	-1	0
2	BRADI_1g63616v3	7.816217569	2.706472272	6.27E-51	1.30E-46	-	-3000	-1	0
3	BRADI_3g38169v3	7.719833111	0.218531275	6.98E-09	1.40E-06	-	-3000	-1	0
4	BRADI_2g16261v3	7.549793292	0.091918857	8.17E-09	1.61E-06	-	-3000	-1	0
5	BRADI_2g21951v3	6.736053884	3.591572389	7.35E-30	1.90E-26	-	-3000	-1	0
6	BRADI_3g27203v3	4.881613193	3.228294815	6.56E-09	1.33E-06	-	-3000	-1	0
7	BRADI_1g04970v3	4.724440792	0.897394715	2.59E-13	1.21E-10	tgaaCCATATTGGtgct	-2261	-2252	1
8	BRADI_1g08357v3	4.04830638	0.758580034	1.93E-06	0.000165425	aaatCCTTAAATGGgta	-2439	-2430	2
						tgatCCAATTGCGGcgac	-795	-786	
9	BRADI_3g16616v3	3.985987254	5.651676208	1.77E-32	5.23E-29	gcacCCATTACGGGgata	-2946	-2937	1
10	BRADI_2g07932v3	3.441550356	4.422819433	1.20E-14	6.39E-12	-	-3000	-1	0
11	BRADI_5g03752v3	2.799059538	2.663455972	1.13E-12	4.58E-10	-	-3000	-1	0
12	BRADI_1g32650v3	2.426497747	4.13791673	8.86E-18	7.65E-15	-	-3000	-1	0
13	BRADI_1g14170v3	2.332225645	6.254373047	0.00000459	0.000346736	-	-3000	-1	0
14	BRADI_1g04963v3	2.319255174	0.192607344	0.0000866	0.003880142	tcccCCTTTTGAGGtaat	-806	-797	2
						tactCCATAACTGGaaca	-2524	-2515	
15	BRADI_1g37629v3	2.065328	3.255739721	7.24E-05	0.003405879	-	-3000	-1	0
16	BRADI_1g66142v3	1.823556384	3.3041866	1.54E-06	0.000137733	gaagCCATTAATGGaact	-2077	-2068	2
						cgtgCCTAATTCGGttgt	-270	-261	
17	BRADI_2g04990v3	1.702330161	6.255320699	3.14E-15	2.03E-12	tacgCCATTATTGGtgac	-1327	-1318	1
18	BRADI_2g04980v3	1.690271662	4.292213769	5.86E-11	1.84E-08	gtcaCCAATAATGGcgta	-2795	-2786	1

KEGG pathway analysis showed that alpha-glucosidase (BRADI_1g32650v3) gene is involved in two pathways, the first is galactose metabolism the second is starch and sucrose metabolism (see appendix, figure S5). This finding does not necessarily mean that *BdMADS13* suppresses the expression of these energy-producing genes, but perhaps the process of saving the plant from oxidative stress need to produce more energy to maintain the homeostasis. Theoretically an increase in alpha-glucosidase activity will cause an accumulation of glucose in the cells. This speculation is reasonable, however it needs experimental validation.

5.4.3.6 *BdMADS13* putative down-regulated genes

Table 5-15: Gene Ontology (GO) of the *BdMADS13* putative down-regulated 21 genes represented in InterProScan IDs (IPS ID), molecular function (MF) and biological process (BP) annotation. Generated using BLAST2GO (Götz *et al.*, 2008).

InterProScan ID	Molecular function	Biological process
Development/cell death	ATP binding	Chromatin silencing
The fantastic four family	Zinc ion binding	Cell differentiation
Integrase, catalytic core	DNA binding	Histone lysine methylation
Reverse transcriptase, RNA-dependent DNA polymerase	Cytosine C-5 DNA demethylase activity	Base-excision repair
Ribonuclease H superfamily	Heterocyclic compound binding	Oxidation-reduction process
Retrotransposon	Endonuclease activity	Translation initiation
F-box-Like domain	Organic cyclic compound	Cellular macromolecule metabolic process
Thioredoxin domain	Protein disulfide isomerase	Cell redox homeostasis
Senescence regulator	Ion binding	DNA demethylation
Thioredoxin-like	ADP binding	Negative regulation of transcription
Oxoglutarate/ion-dependent dioxygenase	RNA-directed DNA polymerase activity	Primary metabolic process
Leucine-rich repeat domain	Transferase activity	Response to oxidative stress
Myb-like transcription	Helicase activity	Maintenance of seed dormancy
Non-heam dioxygenase N-terminal	Nucleic acid binding	Endosperm development
P-loop containing nucleoside triphosphate hydrolase	Histone-lysine N-methyl transferase activity	Post-embryonic plant organ development
Isopenicillin N synthase-like	Protein binding	Response to endoplasmic reticulum stress
Ribonuclease H-like	4 ion, 4 sulfur cluster binding	RNA-dependent DNA biosynthetic

The table 5-15 illustrates the molecular function and the biological process, together with the domains contained in, the putative up-regulated genes of *BdMADS13*. The InterProScan analysis showed that there was a developmental/cell death domain, this gene is differentially expressed in stress conditions and developmental events in *Arabidopsis*, however, its molecular function is not clear (de Camargos *et al.*, 2019). The scan highlighted the presence of the chloroplastic fantastic four (FAF)-like domain, which is implicated in regulating shoot meristem size (Wahl *et al.*, 2010). A member of this family, FAF3, was also found to be highly inducible by cytokinin in *Arabidopsis* (Bhargava *et al.*, 2013). A polycomb protein that has an essential role in epigenetic maintenance by silencing inappropriate expression by methylation (Schuettengruber *et al.*, 2007), and also involved in the controlling the reproduction process by affecting the transition to a flowering phase in the plant (Guitton & Berger, 2004). Moreover, this Polycomb (PcG) group proteins were also found to play an essential role in timing endosperm development in *Arabidopsis* (Ingouff *et al.*, 2005). Some of *BdMADS13* RNAi line down-regulated genes were the F-box protein

BRADI_2g27197v3 that is an ortholog of the rice *Os F-box82* (Os02g0294100) which has QTL trait locus related to plant height (qtaro_375; study: 2003_pha_308) and spikelet fertility (qtaro_11_study: 2004_pha_22) according to Q-TARO (QTL Annotation Rice database). Therefore, a reduction in the expression of this gene would cause the plant to be short and sterile as in our phenotypic analysis in chapter 4 (Figure 4-5). In addition, there were two *MYB* genes (BRADI_2g23341v3) and (BRADI_2g13920v3) that show a reduction in their expression. It was reported that one of the *MYB* genes, *OsAID1* (Os06g0181300), is involved in anther development in rice, as its mutant showed male sterility due to the absence of starch accumulation of in pollen grains and an indehiscent anther due to the absence of PCD in the septum and stomium (Qian-Hao *et al.*, 2004). These results together with the phenotype observed in chapter 4 (Figure 4-18) suggest that there is some degree of functional conservation in these genes between *Brachypodium* and rice.

Table 5-16: Go term annotation of the overlap (Vyinf and PrAs samples) list of the genes that are downregulated by *BdMADS13*, IEA: Inferred from Electronic Annotation, IDA: inferred by direct assay, IBA: Inferred from Biological aspect of Ancestor, IMP: inferred by mutant phenotype,

Sequence Name	Sequence Description	Mapping GO Term	Annotation GO Term	InterPro GO Term
BRADI_5g26380v3	FAMILY NOT NAMED (PTHR33065)			
BRADI_5g01574v3	receptor like protein kinase		protein serine/threonine kinase	Protein kinase domain
BRADI_4g13987v3	Disease resistance RPP13-like	Protein kinase-like domain superfamily	cytoplasm vesicle-mediated transport exocytosis ion binding	ADP binding, Leucine-rich repeat domain superfamily; NB-ARC; P-loop containing nucleoside triphosphate hydrolase
BRADI_4g13380v3	Domain of unknown function DUF1618			
BRADI_3g48970v3	probable inactive histone-lysine N-methyltransferase		histone-lysine N-methyltransferase zinc ion binding chromosome organization	F-box-like domain superfamily/ Pre-SET domain/ histone-lysine N-methyltransferase activity/nucleus/zinc ion binding Histone-lysine N-methyltransferase
BRADI_3g41691v3	1-aminocyclopropane-1-carboxylate oxidase		oxidoreductase activity metal ion binding	oxidoreductase activity/oxidation-reduction process/ Oxoglutarate/iron-dependent dioxygenase; Isopenicillin N synthase-like; Non-haem dioxygenase N-terminal domain
BRADI_3g08301v3	ATP-dependent RNA helicase drs1		helicase activity translation factor RNA binding ion binding	Nucleoside-triphosphate phosphatase Domain of unknown function DUF4283
BRADI_3g01143v3	Helicase protein MOM1	DNA binding	biosynthetic process cellular nitrogen metal ion binding HDA1 complex	Non-haem dioxygenase N-terminal domain; Chromo-like domain
BRADI_3g00210v3	protein disulfide isomerase-like 1-4		mitochondrion vacuolar membrane endoplasmic reticulum protein folding response to oxidative stress/ plastid protein disulfide homeostatic process	Thioredoxin domain ; cell redox homeostasis. Protein disulfide-isomerase
BRADI_2g34547v3	DCD (Development and Cell Death) domain protein		Kelch motif	Development/cell death domain
BRADI_2g33690v3	predicted protein		retrotransposon nucleocapsid-IEA	Development/cell death domain
BRADI_2g27197v3	F-box protein		integral component regulation of transcription	Senescence regulator \$40/ F-box domain
BRADI_2g23797v3	A/G-SPECIFIC ADENINE GLYCOSYLASE/ENDONUCLEASE III	DNA repair-IEA DNA demethylation-IDA base-excision repair-IEA	molecular function DNA metabolic response to stress	HhH-GPD domain
BRADI_2g23341v3	transcription factor MYB3R-2	DNA binding-IEA integral component	DNA binding nucleus	Myb-like transcription factor

BRADI_2g13920v3	hypothetical protein BRADI_2g13920v3		integral cellular component of membrane	Myb-like transcription factor
BRADI_2g07820v3	POLYCOMB PROTEIN EED	PcG protein complex-IDA histone methyltransferase activity-IBA methylation-IMP maintenance of seed dormancy-IMP endosperm development- chromatin silencing complex-	reproduction methyltransferase maturation cell differentiation cellular nitrogen chromosome organization	WD domain, G-beta repeat Histone-lysine N-methyltransferase
BRADI_2g05802v3	F-box/LRR-repeat protein			Senescence regulator S40
BRADI_1g67440v3	protein FAF-like, chloroplastic		Fantastic Four meristem regulator	The fantastic four family
BRADI_1g62007v3	retrotransposon protein	histone acetyltransferase complex-IEA histone acetylation-IEA phosphorylation-IEA kinase activity-IEA transferring hexosyl groups zinc ion binding-IEA retrotransposon nucleocapsid-IEA pollen development-IEA protein serine/threonine	ion binding Alpha-N- acetylglucosaminidase (NAGLU) tim-barrel domain	Reverse transcriptase, RNA-dependent DNA polymerase/ Ribonuclease H superfamily/ Retrotransposon Ty1/ Integrase, catalytic core/ Ribonuclease H-like superfamily/ retrotransposon nucleocapsid
BRADI_1g59720v3	predicted protein		Senescence regulator	Senescence regulator S40
BRADI_1g21455v3	Protein of unknown function DUF3615	nucleic acid phosphodiester bond hydrolysis-IEA RNA-directed DNA polymerase activity-IEA endonuclease activity-IEA biosynthetic process-IEA transferase activity-IEA	nuclease activity DNA metabolic biosynthetic process Nucleotidyl- transferase activity zinc ion binding	RNA-directed DNA polymerase

In the overlap list (very young inflorescence and pre-anthesis carpel) of the down-regulated genes, the CARG-box motif scan revealed that only nine genes presented the binding site in their promoters, including the fantastic four meristem regulator (BRADI_1g67440v3) with FAF domain that regulates size and shoot meristem and two of the MYB-like transcription factor (BRADI_2g13920v3) and (BRADI_2g23341v3) that play an important roles in plant growth and development, especially in auxin pathways and anthocyanin biosynthesis (Nguyen & Lee, 2016). Another antioxidant, thioredoxin domain (BRADI_3g00210v3), showed negative fold-change and also had one CARG-box binding site (redox signaling). Some genes have even four motifs such as DCD (Development and Cell Death) domain protein (BRADI_2g34547v3) (table 5-16) and Senescence regulator S40 protein that contain an F-box (BRADI_2g05802v3). The F-box proteins in plants play essential roles in controlling many biological processes, however, one of the F-box genes, *OsFbox217* (Os04g0464966), in rice was found to be essential for male meiotic progression; a mutation in this gene displayed a formation of disrupted telomere, impaired pairing and synapsis of homologous chromosomes, and aborted meiocytes (He *et al.*, 2016). Another F-box (*OsFbox509*) (Os10g0125300) in rice was found to regulate the tapetum cells PCD in the rice anther by the control of tapetum degeneration retardation (TDR) gene; the transgenic line of this F-box protein showed abnormal tapetal degeneration and aborted microspores development (Li *et al.*, 2015). These findings support the observation made

for the *Bdmads13* RNAi line anther in chapter 4 (figure 4-18) where the tapetum cells were persisting at anther maturity stage. Furthermore, these results revealed that the absence or reduction in the expression of *BdMADS13* causes the reduction of transport, molecular transduction, transcription regulation and enzyme regulation functions which are essential in maintaining homeostasis in the plant.

The results illustrated in table 5-12 and 5-16 was generated using BLAST2GO software (Götz *et al.*, 2008) that is linked to many databases, including InterProScan. These findings were re-analysed using Pfam database and confirm all the annotations given by InterProScan.

Table 5-17: Overlap gene list (PrAs+VyIn) that are down-regulated fold change and the position of CarG-box binding motif in their promoters.

	SeqID	logFC	logCPM	PValue	FDR	matching_seq	Start	End	Score
1	BRADI_5g26380v3	-6.057598837	0.746313409	1.06333E-13	5.24194E-11	aaaaCCAAATTGccta	-2509	-2500	1
2	BRADI_3g08301v3	-5.342553079	3.467748055	7.27115E-43	3.76373E-39	-	-3000	-1	0
3	BRADI_5g01574v3	-4.170035926	4.224638899	9.10257E-05	0.004044391	-	-3000	-1	0
4	BRADI_4g13987v3	-3.413021287	4.404291245	3.65438E-35	1.26106E-31	-	-3000	-1	0
5	BRADI_1g21455v3	-2.900976188	2.029447187	0.000699069	0.01990953	tcagCCATTTcAGGgaga	-2547	-2538	2
						taagCCTTTTCGGGtacc	-724	-715	
6	BRADI_2g07820v3	-2.229805743	0.521528416	3.06733E-07	3.73583E-05	-	-3000	-1	0
7	BRADI_2g23797v3	-2.095197192	6.610122511	4.03698E-15	2.5329E-12	-	-3000	-1	0
8	BRADI_3g00210v3	-1.9185904	5.979250241	2.19751E-12	8.27264E-10	ctgtCCTATTTGGGgaact	-1884	-1875	1
9	BRADI_3g48970v3	-1.811167276	2.026455175	2.69399E-06	0.000224012	-	-3000	-1	0
10	BRADI_4g13380v3	-1.775881948	2.216065007	4.37023E-06	0.00033644	gactCCATTTTCGGGttgc	-863	-854	1
11	BRADI_2g27197v3	-1.688043569	0.616373884	2.71959E-06	0.000224524	-	-3000	-1	0
12	BRADI_3g41691v3	-1.687653901	2.158300672	0.000544344	0.016405605	-	-3000	-1	0
13	BRADI_2g33690v3	-1.586524996	3.204445245	1.32722E-07	1.79608E-05	-	-3000	-1	0
14	BRADI_3g01143v3	-1.563595058	4.893601223	1.78997E-08	3.16764E-06	-	-3000	-1	0
15	BRADI_1g59720v3	-1.537070417	1.585983604	0.000896424	0.023826015	-	-3000	-1	0
16	BRADI_1g62007v3	-1.501916191	5.410670525	1.67617E-15	1.11952E-12	-	-3000	-1	0
17	BRADI_2g23341v3	-1.454474303	1.923434886	0.000206125	0.007703647	agggCCATTTTGGGgggt	-959	-950	1
18	BRADI_2g13920v3	-1.34359812	3.410733002	1.28925E-07	1.75618E-05	tatgCCTTATTGGtatg	-2461	-2452	1
19	BRADI_1g67440v3	-1.332130029	5.647382553	2.81076E-08	4.71102E-06	gccgCCTTTTGTGGcaac	-2056	-2047	1
20	BRADI_2g05802v3	-1.083582881	3.693774984	0.000640324	0.018778914	ttttCCAATATGGGggca	-2248	-2239	3
						agatCCTTTTTTGGgaga	-873	-864	
						ttgtCCAATTTTGGgagtt	-492	-483	
21	BRADI_2g34547v3	-1.061848852	5.308356697	6.57756E-08	9.79772E-06	gaatCCTAAACGGccat	-1922	-1913	4
						ggatCCAATTGCGGcggc	-2197	-2188	
						ctcaCCTATATTGGcggg	-423	-414	
						atttCCAATTTTGGcccg	-255	-246	

Additional analysis was performed on the genes that were down-regulated in only pre-anthesis carpel sample (Table S15, appendix). The results included genes: RNA helicase containing the DEAD-box (Bradi_3g08301v3) that is involved in various aspects of RNA metabolism and unwinding of nucleic acids showed a fold change of around -5 in the transgenic line *Bdmads13* RNAi (table 5-16). It was reported that the nucleolar DEAD-box RNA helicase TOGR1 regulates thermos-tolerant growth in rice by maintaining vital rRNA homeostasis (Wang *et al.*, 2016). In the KEGG analysis that was performed in our study, we found that this particular gene list intervenes in thiamine metabolism as well as the purine metabolism (ec:3.6.1.15 - phosphatase) (Figure 5-9).

In addition, there was approximatively 4 negative fold-change in GDSL esterase/lipase LTL1 (Bradi_3g28100v3) (table S15, appendix). This family plays important roles in plant defence and growth, regulation of plant cell wall components, and is essential in many developmental processes (Lai *et al.*, 2017). There was a 3 fold reduction in Zinc-finger (BRADI_4g30523v3) and YABBY transcription factor (BRADI_3g50050v3) (table S15, appendix) This particular YABBY gene ortholog in rice (OS04G0536300) was related to plant height at 35 days after germination (qtaro_655) (QTARO database). There was also -2 fold change in two myeloblastosis-like family transcription factors MYB (BRADI_4g38965v3) and MYB/SANT-like (Bradi_1g12900v3) together with the bHLH (BRADI_3g38890v3) transcription factor. Proteins of these two transcription factor families can interact as homo/heterodimers and form MYB/bHLH complexes to regulate distinct cellular processes, such as cell wall synthesis, cell death and hormone signalling (Pireyre & Burow, 2015). The ortholog of *bHLH* (BRADI_3g38890v3) in rice was found to maintain iron homeostasis and is also involved in the heading timing for the plant (qtaro_1007) (QTARO database). Also MYB TF was found in figure 5-2 to be co-expressed with *MADS3* in *Triticum aestivum* and *bHLH* with *MADS3* in *Arabidopsis*. A recent study in rice revealed that *OsBHLH142* directly regulates the expression of *EAT1* and *TDR20* required for tapetum development and degeneration and it also affects the deposition of lignin and cell-wall degeneration-related genes required for the anther dehiscence (Ranjan *et al.*, 2017). In addition, there was a -2.4 reduction in WUSCHEL-RELATED HOMEODOMAIN 3 (BRADI_2g37650v3). Its ortholog in rice *OsWOX3A* is involved in organ development in rice, lateral-axis outgrowth and vascular patterning in leaves, lemma and palea morphogenesis

in spikelets, and development of tillers and lateral roots (Cho *et al.*, 2013). This demonstrates the conserved role of the *WOX3* in *Brachypodium* and supports the phenotype observed in the *Bdmads13* RNAi with more tillers (the roots was not analysed in this study).

Note: Differential gene expression analysis was re-analysed using another software (Partek Genome Suite 6.6) and presented similar results confirming our findings (see appendix, figure S6-7).

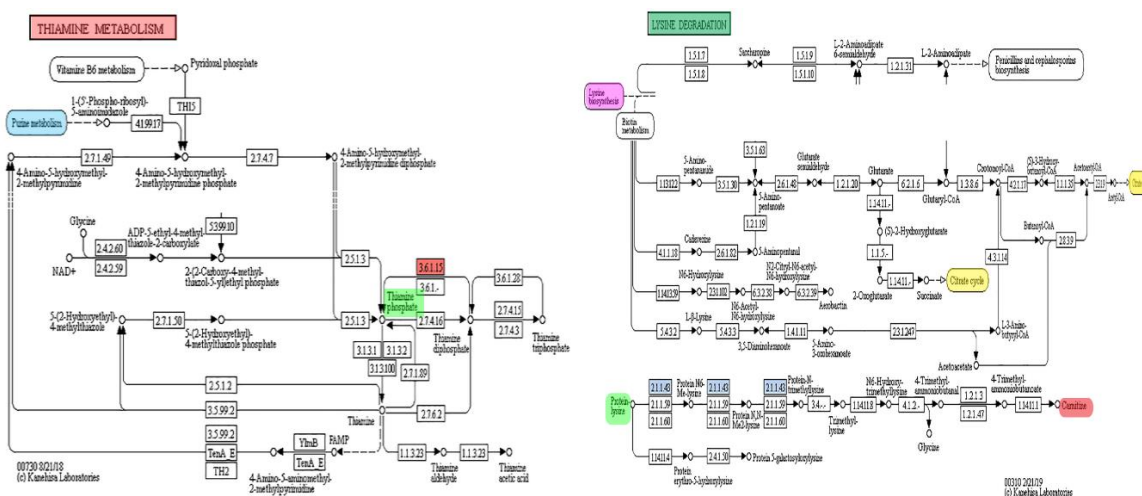


Figure 5-9: KEGG pathway analysis of the down-regulated genes list showing that 2 genes (BRADI_3g48970v3 and BRADI_2g07820v3) are involved in lysine degradation, and one gene (BRADI_3g08301v3) in purine and thiamine metabolism.

Among the genes that show low expression in the *Bdmads13* RNAi line is BRADI_3g08301v3 involved in thiamine metabolism (figure 5-9). It is believed that in the plant thiamine diphosphate is an important enzymatic cofactor in major metabolic pathways acting as a signalling molecule in response to changing environments such as biotic and abiotic stress (Tunc-Ozdemir *et al.*, 2009). Moreover, there were two genes (BRADI_3g48970v3 and BRADI_2g07820v3) involved in lysine degradation. In maize, lysine catabolism occurs in the developing seeds, mainly in the developing endosperm for storage protein synthesis (Arruda *et al.*, 1982). Therefore, transgenic plant might have a disruption of the endosperm development as they have a low level of the gene that participates in this pathway (plants died at pollination stage).

5.4.3.7 RNA-seq validation using RT-qPCR

Three genes were selected from the putative down-regulated genes, and three from up-regulated genes in order to validate the RNA-seq analysis.

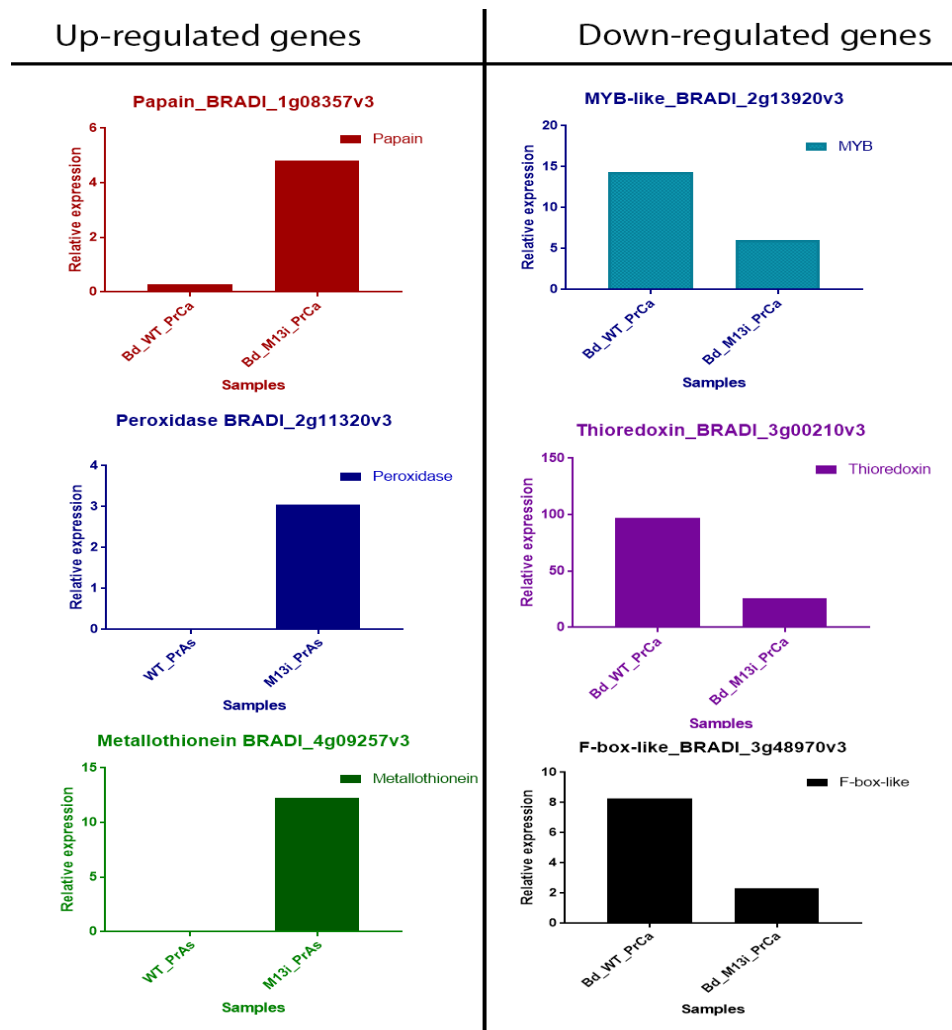


Figure 5-10: Quantitative RT-PCR validation of differentially expressed genes that showed a significant expression differences between the wild-type and *Bdmads13* RNAi line.

The *Bdmads13* RNAi line appears to have excessive production of ROS scavenging enzymes (Figure 5-10) such as peroxidase and non-enzymatic systems such as metallothionein (maintain metal homeostasis) and glutathione (BRADI_3g38169v3) (Table 5-13). These antioxidants can be seen in plants under environmental stress (heat, UV, salinity and flood) to prevent the oxidative stress that can result in significant damage to cell structures and ultimately, cell death. The metallothioneins have been reported to play a role in various pathways and cellular processes, including the regulation of cell growth and proliferation (Grennan, 2011), while, the glutathione peroxidases are also involved in secondary metabolite metabolism and respiration (Ozyigit *et al.*, 2016).

Although the plant antioxidant surviving mechanism is activating many scavenging enzymes, it appears that some of the programmed cell death proteases are very high such as papain (BRADI_1g08357v3) and aspartic peptidase (table 5-13).

5.5 Discussion

Male Sterile germplasms provide an essential resource in research to understand more about reproductive development and function in plants (Perez-Prat & van Lookeren Campagne, Michiel M, 2002). The transgenic line for the *BdMADS13* gene presented completely sterile plants that are not able to successfully surpass the anthesis stage. In order to investigate potential down-stream target genes of C and D-Class genes, a co-expression analysis was performed to select the candidate genes, and then transient expression analysis was applied to confirm the interaction and activation after scanning for binding motifs in their promoters. Furthermore, a functional examination of *BdMADS13* was explored applying RNA-seq analysis to the RNAi line.

Our finding showed that *MADS7* strongly co-expresses with *MADS3* in four species (*Triticum aestivum*, *Arabidopsis thaliana*, *Hordeum vulgare* and *Oryza sativa*) using multiple platforms such as GENEVESTIGATOR and GENI3. This suggests that *MADS7* transcription factor either interacts with *MADS3* to form dimeric or quaternary complex to activate downstream promoters or that one of them activates the other. The transient expression analysis results showed that there was a significant activation of the metallothionein promoter in the presence of both *MADS3* and *MADS7* transcription factors. These findings suggest that they interact with each other to bind to the CARG-box motif and activate the metallothionein gene.

Regarding the member of the D-class *BdMADS13*, the differentially expressed genes gave a comprehensive understanding of what occurred to the plant in the case of reduction of the *BdMADS13* transcript. Our findings indicated that many biological processes were affected in the plant. The genes that were highly expressed in the transgenic line were mainly programmed cells death related genes such as cysteine endopeptidase which is senescence-specific protein and the desiccation-associated protein Ferritin-like BRADI_1g62680v3 (Zha *et al.*, 2013) (table 5-12); also, there were genes that are dormancy-associated (BRADI_1g14170v3). In addition, the deterioration of the plant

antioxidant coping mechanism appears to be very active with high expression of ROS-scavenging genes such as peroxidase and metallothionein (figure 5-12); these rescue genes were accompanied with energy-producing genes including alpha-glucosidase and glycosyl hydrolase (figure 5-12).

On the other hand, *Bdmads13* RNAi line presented low expression of some vital genes that intervene in many pathways. For instance, the chloroplastic fantastic four protein (BRADI_1g67440v3) containing FAF-like domain which plays an essential role in determining and regulating the shoot meristem size in *Arabidopsis thaliana* (Wahl *et al.*, 2010), another chloroplastic enzyme that was low in RNAi line, is thioredoxin (Table 5-15 and Table 5-16) which is considered essential for light-mediation and photoautotrophic growth (an essential factor in light-mediated activation of Calvin cycle enzymes) (Jeon, 1996); this is highly correlated with the spatial expression in chapter 3 of *BdMADS13* in the chlorenchymal endocarp cells (Figure 3-8). In addition, there were genes that are involved in the endosperm development and reproduction, such as, the polycomb protein EED (BRADI_2g07820v3) and genes involved in pollen development (BRADI_1g62007v3) (table 5-16). There was a low expression of at least two F-box proteins (BRADI_2g27197v3) and (BRADI_3g48970v3). F-box proteins in plant play important roles in controlling many biological processes, however these specific genes were analysed using KEGG analysis showed that the F-box gene (BRADI_3g48970v3) together with polycomb protein EED (BRADI_2g07820v3) are involved in the lysine degradation pathway (ec:2.1.1.43 - N-methyltransferase) (figure 5-9) which is an essential process in the developing endosperm for storage protein synthesis (Arruda *et al.*, 1982). In rice, for instance, the F-box (OsFbox509) (Os10g0125300) was found to regulate the tapetum cells PCD in the anther which affects microspores development (Li *et al.*, 2015). Also, one of the F-box protein found to act as a transcriptional co-factor to regulate floral development (Chae *et al.*, 2008). The reduction of the expression of these genes cause significant disruptions to the plant system and eventually complete sterility and death.

We conclude that the loss of function of *BdMADS13* cause numerous enzymatic and non-enzymatic ROS-scavenging genes to increase. This could be explained by that the plant activates its army of antioxidants in order to prevent any accumulation of reactive oxygen species in the cells; this affects the tapetum PCD which need to be degraded to nourish the

developing pollen grain for successful maturation, in the *Bdmads13* RNAi the tapetum persist in the anther causing sterility to the plant (chapter 4, figure 4-18). Metallothionein was one of the ROS-scavenging that was very high in the pre-anthesis of the silenced line, these metal-responsive elements, can bind up to seven atoms zinc or cadmium, resulting in low levels of these metals. This can cause a cell dysfunction in some eukaryotic organisms if produced in abundance in the cell (Mehrian-Shai *et al.*, 2015).

Another scenario is that *BdMADS13* directly down-regulates some enzymatic and ROS-scavenging genes and the loss of its function activates the gene that increases ROS level. Therefore, the imbalance between the ROS and their scavenging molecules cause considerable damage to the plant and this can be seen clearly by the high level of the proteases such as papain and aspartic peptidase.

5.6 Conclusion

Although the transgenic line has a high level of ROS scavenging genes such as peroxidase, glutathione and metallothioneins in order to prevent the redox stress, there is a high level of desiccation genes that promote the programmed cell death such as ferritin like and papain from C1 protease family. It appears that the coping mechanisms of the plant to survive did not prevent the fate of desiccation and senescence in the plant. Future efforts are needed to completely unravel the roles that *BdMADS13* plays in the plant; perhaps, a complete silencing would answer many questions.

Chapter 6

Brachypodium grain development and programmed cell death process.

Chapter 6. *Brachypodium* grain development and programmed cell death process

Note: This chapter was accomplished in collaboration with my colleague Charles Solomon and was written as a manuscript submitted for publication and in bioRxiv. The chapter is included in both theses. We designed the study, performed all experiment analyses and wrote the manuscript.

6.1 Abstract

At early stages of the grain development, maternal tissues such as the nucellus, the pericarp, and the nucellar projections undergo a progressive degeneration by programmed cell death (PCD), which facilitate the remobilisation of their cellular components to provide nutrient for the new tissues such as the embryo and the endosperm. During grain maturation, the PCD occurs in the endosperm as well and determines the lifetime of cells; therefore, affect growth cereal yield. However, the features and mechanisms of PCD in the developing starchy endosperm in the *Poaceae* remain unclear. In the present study, we provide a detailed investigation of PCD progress in developing grain of a wild model species, *Brachypodium distachyon*, focusing on the spatial and temporal progress of PCD. Furthermore, we investigate the expression pattern of the major protease families involves in this process. *Brachypodium* nucellus degenerates by PCD in a centrifugal pattern following anthesis although at a slower rate compared to cereals. Similarly, the mesocarp with rapid degradation in the lateral sides. At the molecular level, we investigate RNA-Seq expression profile of *Brachypodium* protease genes belonging to aspartic, papain, vacuolar processing enzymes and meta-caspase families, in vegetative and reproductive tissues. We identified protease genes from these families that are highly expressed in grain tissues and may be part of the molecular machinery for PCD execution in *Brachypodium* grains.

6.2 Introduction

Programmed cell death PCD is a highly organised process that eliminates certain kind of cell in very accurate plant growth time in order to maintain the equilibrium and assure a normal development. It has been studied in numerous plant developmental processes

including gametogenesis and seed development and shown to be an essential process of cereal grain development and germination (Daneva *et al.*, 2016; Domínguez & Cejudo, 2014).

Different types of PCD play crucial roles in vegetative and gametophytes development of the plant. At anthesis, the pollen grains land on the carpel stigma then develop tube to migrate to the ovule to fertilise it with two polar nuclei, this event results in triploid primary endosperm proliferation and degeneration of the synergids inside the embryo sac at the basal side of the grain and degradation of the antipodal cells at the apical side (An & You, 2004). Nuclear materials from the degenerating antipodal cells nourish the nuclear divisions in the growing syncytium surrounding the characteristic central vacuole (Engell, 1994). The inner cell layers nucellus nuclei start fragmentation shortly after fertilisation and gradually spread to the outer layers of the nucellus. As the grain develop and mature, the PCD progress and expand to all maternal tissues, which provide growth space and nourishment for the emerging endosperm until reaching the starchy endosperm at late stages of development. It was reported that in wheat gradient nucellus degeneration from the internal towards the external layers was observed. In fact, the initiation of PCD coincided with endosperm cellularisation. All what remains at maturity are, nucellar epidermis which enclose the aleurone and endosperm, nucellar lysate and nucellar projection which is at the chalazal region and acts as transport route for nutrient supply to the endosperm (Evers & Reed, 1988; Freeman & Palmer, 1984; Morrison *et al.*, 1978; Wang *et al.*, 1994).

The nucellar epidermis in barley and wheat, which is made up of a well-defined rectangular cell, continues to persist even later in development and actually retains its cellular integrity during grain development (Domínguez & Cejudo, 1998; Norstog, 1974). However, in maize the nucellar layers PCD was extremely rapid after fertilisation and coincide with endosperm expansion which must cause significant Osmotic stress to the adjacent nucellar layer triggering their PCD (Kladnik *et al.*, 2004). In rice, the nucellar epidermis located next to the dorsal vascular bundle, transport nutrients to the endosperm and embryo, start gradually degenerate in the first 5 days after pollination (DAP) and eventually collapse by 21 DAP (ELLIS & CHAFFEY, 1987; Oparka & Gates, 1981; Wu *et al.*, 2016). These studies showed the PCD in the nucellar epidermis in domesticated cereal. However, our investigation is

performed on a wild grass model, providing a very detailed and timed nucleic degradation in the all grain development stage.

The cereal pericarp is typically made up of three distinct layers, epicarp, which is the outermost layer, several layers of parenchymal mesocarp, which is the middle layer, and the chlorenchymal endocarp, which is the inner layer surrounding the ovary. This maternal protective tissue undergo the PCD shortly after anthesis starting from the endocarp closer to the integument to progress gradually towards the outer layers, as the grain endosperm cellularise these layers degenerate and collapse leaving only the cuticle that adjacent to the outer integument (Domínguez & Cejudo, 2014; Freeman & Palmer, 1984; Wu *et al.*, 2016).

Although the precise molecular mechanisms underlying PCD presses are not fully elucidated, a number of different protease gene families have been linked with PCD of specific tissues in developing cereal grains. The expression of different proteolytic enzymes has been considered as markers of cell degeneration, these genes could be grouped into four families.

One of the families related to the PCD is the aspartic protease that belongs to Clan AA Family A1 and called pepsin-like (aspartic protease) (<https://www.ebi.ac.uk/merops/>) (v12.1). In barley, an analysis of the genes that are specifically expressed in the maternal tissues revealed that an aspartic protease-like protein nucellin (HORVU4Hr1G088530) (IBSC_v2 assembly) is specifically expressed in the nucellus during nucellus degradation (Chen & Foolad, 1997). Its orthologues in rice were identified as Aspartic proteinase *OsAsp1* (Os11g0184800), *OsAsp2* (Os11g0183900) and *OsAsp3* (Os12g0177500) (Bi *et al.*, 2005). Another know aspartic protease is Oryzasin *Os05g0567100*, which found to be expressed during Seed ripening and germination (Asakura *et al.*, 1995) and *OsS5* *Os01g0140100* that plays a key role in regulating of the productivity and compatibility in rice (Chen *et al.*, 2008).

The following group is papain Peptidase family C1 belonging to the CA clan (Yin & Xue, 2012; Domínguez & Cejudo, 1998). It contains lysosomal enzymes including cathepsin B, L, S, K, H and dipeptidyl-peptidase I. There are gene from this family that are expressed during germination phase such as, Cathepsin-like enzymes encoded by *BdCathB* (BRADI_1g09729v3) in *Brachypodium*, *HvcathB* (HORVU4Hr1G010390) in barley, and their

orthologue in wheat TaCathB (TraesCS4A02G252800/ TraesCS4B02G062500/ TraesCS4D02G061400), OsCathB (Os05g0310500) in rice, and ATCathB3 (AT4G016100) in *Arabidopsis* (Cejudo *et al.*, 1992; Cercós *et al.*, 1999; Iglesias-Fernández *et al.*, 2014; González-Calle *et al.*, 2014). The PCD is essential during germination to degrade the nutrient storage to sustain embryo growth and development until reaching an autotrophic phase. These orthologues are highly induced in germinating seeds in the aleurone layers and its expression is regulated by transcription factors (TFs) encoded by genes *BdGamyb* and *BdDof24* and their orthologs in the corresponding species (González-Calle *et al.*, 2014). In addition to the EPB gene sub-family which consists of two very similar members, EPB-1 (HORVU3Hr1G091920) and EPB-2 (HORVU3Hr1G091790), are highly induced by gibberellins secreted by the embryo and express in the aleurone layers during germination, these genes are essential for the degradation of seed endosperm storage proteins in barley (Cercós *et al.*, 1999).

Another known member in C1 family is the Aleurain (HORVU5Hr1G061770) which is a barley thiol protease similar to cathepsin H (Rogers *et al.*, 1985) is expressed in the nucellar projections and the modified aleurone cells. The expression of this thiol protease gene overlaps with PCD in nucellus, nucellar epidermis and nucellar projection of developing wheat grains (Domínguez & Cejudo, 1998).

Two papain family members in rice, CP1 (Os04g0670500) and Cys-EP, are necessary for the degeneration of the anther tapetum at anthesis and the endosperm for successful germination (Schmid *et al.*, 1999; Lee *et al.*, 2004). In addition, the cysteine proteases Os02g48450 /Os02g0715000 (IRGSP-1.0), found to be essential for the regulation of nucellar projection degradation and early seed development and also is a target of OsMADS29 (Yin & Xue, 2012). Moreover, three oryzains found in rice including, Oryzain a (Os04g0650000), Oryzain b (Os04g0670200), Oryzain c (Os09g0442300), are expressed only in seeds during germination time (Watanabe *et al.*, 1991). This family C1, has been divided into sub-clades based on conserved motifs that they share, including, B. Like that have a similar sequence as the human B-cathepsin, F. Like that, share the conserved motive ERFNAQ, H. Like which contain the conserved motif EREFNIN, and six types of L. like (Martinez & Diaz, 2008).

The third family that has been linked to the PCD is the vacuolar processing enzyme (VPEs) alternatively called legumains from C13 family CD clan (Linnestad *et al.*, 1998; Julián *et al.*, 2013; Radchuk *et al.*, 2010; Sreenivasulu *et al.*, 2006). In barley, the expression of *HvVPE1* gene was detected in the maturing endosperm, this expression coincided with the degeneration of nuclei in starchy endosperm and transfer cells (Tran *et al.*, 2014). The nucellain *HvVPE2a* (HORVU2Hr1G092080), a barley homologue of VPE (Linnestad *et al.*, 1998), is found to be specifically expressed in the early endosperm fraction, but localised exclusively in the nucellus of the developing barley grain and together with closely similar *HvVPE2b* and *HvVPE2d* might be involved in nucellar PCD. The *HvVPE3* gene expression is detected in pericarp and endosperm. Its mRNA levels in pericarp peak between 4 and 10 DAA, while in endosperm they increase from 4 days after anthesis until maturation. In contrast, *HvVPE4* which strongly cells specific for pericarp parenchyma with a maximum between 8 and 10 days after anthesis and potentially plays a role in PCD in this tissue (Radchuk *et al.*, 2010). *OsVPE1* is a member of the C13 family and orthologues of barley nucellain was significantly down-regulated in *OsMADS29* knock-down transgenic seeds which may indicate direct regulation of *OsVPE1* by *OsMADS29* in rice (Yang *et al.*, 2012).

The fourth family is the Peptidase family C14 contains cytosolic endopeptidases called metacaspase (Vercammen *et al.*, 2004a; Tsiatsiani *et al.*, 2011; Bostancioglu *et al.*, 2018). These enzymes are well known by their unique domain P20 and P10 and are divided into three groups based on its positions in the sequence, Type I, type I* and type II. Type I contain *AtMC1* and *AtMC2* genes which antagonistically control PCD as *AtMC1* positively and *AtMC2* negatively regulate the cell death in *Arabidopsis* (Coll *et al.*, 2010). *HvMC1* and *HvMC2* expressions were high in the embryonic stage and *HvMC3* found to be high in the leaf. In addition, the analysis suggested that *HvMC4* and *HvMC5* have opposite expressions during excess-B (micronutrient boron) driven PCD and seem to have an antagonistic effect in barley (Bostancioglu *et al.*, 2018)

The type II Metacaspases *Atmc4* and *Atmc9* of *Arabidopsis thaliana* cleave substrates after Arginine and Lysine (Vercammen *et al.*, 2004). In addition, *AtMC4* found to accelerate the cell death under biotic and abiotic stresses (Watanabe & Lam, 2011). There is also *AtMC8* gene that is strongly upregulated during oxidative stress (Tsiatsiani *et al.*, 2011). Whereas, in barley, *HvMC6* and 8 were found to be the most abundant transcripts in control root

(Bostancioglu *et al.*, 2018). In *Brachypodium*, there was an identification of 10 genes (Fagundes *et al.*, 2015).

Brachypodium distachyon has become established model system for comparative grain evolution and development relevant to tropical and temperate cereals (Francin-Allami *et al.*, 2019; Guillon *et al.*, 2011; Hands & Drea, 2012; Kourmpetli & Drea, 2014; Opanowicz *et al.*, 2011; Trafford *et al.*, 2013). Although it expresses the genes required for starch synthesis, its grains have less starch endosperm cells (Trafford *et al.*, 2013). The PCD process has been described to play a key role in developing endosperm by degeneration the maternal cells around it. A detail of programmed cell death events, patterns, and progression in developing *Brachypodium* grains has not yet been reported.

Therefore, in this study, it has been suggested that the pattern and progression of PCD in *Brachypodium distachyon* grains may be different from cereals. We undertook a systematic histochemical and molecular study of PCD in developing *Brachypodium distachyon* grains by following the progress of the nucleic degradation throughout several stages of grain development and using *BdMADS29* gene as a marker to localise its expression using *in situ* habitation. In addition, we investigate the expression profiles of *B. distachyon* genes belonging to protease families; aspartic protease A1, papain protease C1, vacuolar processing enzyme C13 and meta-caspase C14 using RNA-seq dataset and validate their expression by RT-qPCR developing *B. distachyon* grains tissues.

Our results indicate that the rate of nucellar PCD is slow in *Brachypodium* compared to cereals. On the other hand, the degeneration of the mesocarp cells proceeds more rapidly in *Brachypodium* compared to cereals. Gene expression analysis suggests conserved roles of *Brachypodium* orthologues of proteases previously known to be involved in PCD and also provides new proteases candidate genes that may be part of the *Brachypodium* grain PCD molecular network.

6.3 Materials and Methods

6.3.1 Plant materials

The inbred line *Bd21* Seeds of *Brachypodium distachyon* were obtained from Dr Sinéad's lab. The seeds were de-husked and grown as described (Draper *et al.*, 2001), seeds were vernalised on moist filter paper in a 9mmØ Petri dish at 5°C for 48h in order to break the

dormancy and stimulate flowering, then moved to room temperature for 48h. The germinated seedlings were then sown in a soil mixture of two-part of the soil, one vermiculite and one sand supplied by Levington® then transferred to a phytotron (controlled environment) under the following conditions: photosynthetic photon flux density (PPFD) of 400 $\mu\text{mol m}^{-2} \text{s}^{-1}$, 20 h light/4 h darkness, with a day/ night temperature of 24°C under 52% relative humidity. Plants were watered by flooding under tray every other 2 days.

6.3.2 TUNEL assay

In order to detect cells that undergo extensive DNA degradation a Terminal deoxynucleotidyl transferase (TdT) dUTP Nick-End Labeling (TUNEL) assay was performed. After sectioning the samples according to the previously described method in the section 3.3.4.1 (Samples collection and preservation). The tissues on the slides were deparaffinised for 10 minutes in two histo-clear baths then rehydrated through a graded series of ethanol (100%, 80% then 50%), followed double distilled water. The samples were treated with 30 min of proteinase K at 37°C then by 5 minutes in two baths of PBS (phosphate buffered saline (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4). Two controls were used; positive (adding DNase recombinant) and negative (only the label solution) in this experiment due to the different nature of the tissues treated, one was for the flowering stage, and the other was for the grain development stages. The *in situ* Cell Death Detection kit (Roche Diagnostics, Germany) was used to label DNA breaks using the TdT enzyme, just prior to the experiment the TUNEL reaction mixture was prepared with 1 part of the enzyme mix to 9 from the labelling solution then applying 70-100 μl of the TUNEL mixture to each slide followed by incubation in dark very humid chamber for one hour at 37°C. The slides were rinsed three times with PBS then observed with a drop of PBS under fluorescence microscope Nikon eclips 80i (Nikon, Japan), using a fluorescence excitation LED Light-source (Precis Excite, Nikon & Cool-LED) in addition to Prior Scientific ProScan II with Joystick (H30V4) which allows control a stage, focusing motor and shutters. The pictures were taken by Nikon DS-QiMC camera (Y-QT, Japan) in PNG format, and visualised using NIS-elements software (BR Ver4.13.05 32bit edition), the filter used was GFP with an excitation peak 47/40 a range of 450-490.

6.3.3 Phylogenetic analysis of selected protease Family genes

Brachypodium Protein of the protease families A1, C1, C13 and C14 were retrieved using INTERPRO (Mitchell *et al.*, 2018) and PANTHER (Thomas *et al.*, 2003) accession number specific for each domain (A1: pepsin; IPR001461, PTHR13683), (C1: papain; IPR000668, PTHR12411), (C13: legumain; IPR001096), (C14: caspase; PTHR31773, PTHR31810) from the peptidase database MEROPS 12.1 (Rawlings *et al.*, 2011). These numbers were used to blast the available Databases using BioMart tool (Smedley *et al.*, 2015), then the protein sequences were uploaded into Geneious R10 (<https://www.geneious.com>) to select only unique longest version of the protein sequence. The alignment was performed using MAFFT v7.308 (Kato & Standley, 2013) under the following parameters: protein model: WAG+G+F and JTT+G+F, BLOSUM62 as scoring matrix, Gap open penalty of 1.53 and an offset value of 0.123. The same approach was used to obtain protease families from wheat, *Arabidopsis*, barley and rice.

Protein sequences of C13 and C14 of *Arabidopsis*, barley and rice were gathered from published reports on these proteins for each species (Bostancioglu *et al.*, 2018; Julián *et al.*, 2013; Rocha *et al.*, 2013; Vercammen *et al.*, 2004; Wang *et al.*, 2018; Radchuk *et al.*, 2010).

Table 6-1: Gene names and IDs for the *HvVPE* sub-family identified in previous studies and in our work.

Gene name	GenBank	Gene ID (IBSC_v2)	UniProt	Ref
<i>HvVPE1</i>	FR696360	HORVU6Hr1G060990	E5AXU3	(Radchuk <i>et al.</i> , 2010)
<i>HvVPE2a</i>	FR696361	HORVU2Hr1G092080	E5AXU4	
<i>HvVPE2b</i>	FR696362		E5AXU5	
<i>HvVPE2c</i>	FR696363		E5AXU6	
<i>HvVPE2d</i>	FR696364	HORVU2Hr1G091880	E5AXU7	
<i>HvVPE2e</i> (b)		HORVU2Hr1G092090		our study
<i>HvVPE3</i>	FR696365	HORVU3Hr1G048520	E5AXU8	(Radchuk <i>et al.</i> , 2010)
<i>HvVPE4</i>	FR696366	HORVU5Hr1G066250	E5AXU9	
<i>HvVPE5a</i>		HORVU1Hr1G012420		our study
<i>HvVPE5b</i>		HORVU6Hr1G039550		

6.3.4 RNA-seq data processing and expression analysis of protease genes

RNA-Seq data of seven samples from *Brachypodium* grain generated by (Drea *et al.*, unpublished) and available on (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-7607 was used. Data quality was checked using FastQC. Reads were mapped to

Brachypodium distachyon v3.0.dna.toplevel.fa downloaded from Phytozome v.12 genome with STAR v2.5.2b (Dobin *et al.*, 2013), then, mapped reads were counted with featureCount function from Rsubread v 1.22.2 package (Liao *et al.*, 2019). Raw read counts were normalised with DEseq2 v1.6.3 (Love *et al.*, 2014). The expression profile of protease gene families in the RNA-Seq data was visualised using ClustVis (<https://biit.cs.ut.ee/clustvis/>), the unit variance scaling was applied to rows, and Rows were clustered using correlation distance and average linkage.

6.3.5 Reverse transcription polymerase chain reaction (RT-PCR)

Oligonucleotide primers were manually designed to amplify *BdMADS29* cDNA products using SnapGene Viewer 2.8.3. The primers were designed to amplify the region of 3' UTR, with an amplicon of 377bp and annealing temperature (T_m°) of 65°C (F: 5' AACACTCTCCTGTGCCGCAT3', Reverse: 5' CCTCCACCGTGACCTTCTTA 3'). Primers were obtained from SIGMA and were diluted (1:10) to 10µm concentration for all reactions. Reactions were performed using Dream Taq Green PCR Master Mix (Thermo Scientific). The RT-PCR reaction contains (5µl of Dream Taq Green PCR Master Mix (2X), 0.5µl of 5µM forward/reverse primer, 3µl dH₂O and 1µl of Templet cDNA 1µg). PCR conditions were with the following cycles: 95°C 6 minutes for denaturation, then 30 cycles of 95°C 30 seconds, the annealing temperature 65°C for 30 seconds, 72°C for 1 minute followed by 72°C for 6 minutes for a final extension. Reaction conditions were varied in some cases to according to product size, primer T_m° and individual gene copy numbers.

6.3.6 Quantitative RT-PCR

Primers for *BdA1*, *BdC1*, *BdC13* and *BdC14* were designed using Integrated DNA Technologies with an amplicon size of around 100bp to anneal efficiently in 65°C and 50% GC content (Table S3). Quantitative RT-PCR reactions were carried out using SYBR Green (SensiMix SYBR Low-ROX Kit) (BIOLINE) on Applied Biosystems™ 7500 Real-Time PCR Systems with 7500 Software v2.0.6. on 96well MicroAmp® Fast Optical 96-Well Reaction Plate with the following reaction mix components (5 µl of 2x SensiMix™ SYBR® Low-ROX, 0.1 µl of 25 µM Forward/reverse Primers, 3.8 µl of DNase-free H₂O, 1 µl of cDNA template). The thermal cycling conditions started by polymerase activation at 95°C for 10 min followed by 40 cycles of 95°C for 15 seconds then the annealing temperature of 65°C for 15 seconds flowed by 72°C extension for 15 seconds. Samples were normalised using

BdACT7 expression, and relative expression levels were determined using the $2(-\Delta\Delta Ct)$ analysis method described in (Livak & Schmittgen, 2001).

6.3.7 mRNA *In situ* Hybridisation.

The primers used are listed in Table S2 and Table S3 following the protocol described in (Drea et al., 2005) and described in sections from 3.3.4.1 to 3.3.4.6.

6.3.7.1 Samples collection and preservation

For mRNA *In situ* Hybridisation experiments, different tissue samples and developing grains stages were freshly collected and fixed in formalin–acetic acid–alcohol (FAA) (3.7% formaldehyde, 5% acetic acid, 50% ethanol). Then, vacuum infiltrated under the pressure of 400 mbar (KNF LABOPORT diaphragm lab vacuum) for 5 minutes then left overnight at 5°C. Three washes with 70% ethanol was performed before the samples goes through an increasing concentration of ethanol for dehydration then gradually removing the ethanol using an increasing concentration of the clearing agent Histo-Clear (National Diagnostics) starting from 25% for approximatively 30 min each bath, so the samples can infiltrate the paraffin wax at 60°C (Formula 'R' Paraffin, supplied by Leica Biosystems) and prepared for the embedding process. All steps were performed manually using HISTOSETTE® Tissue Processing/Embedding Cassettes.

6.3.7.2 Tissues embedding and sectioning.

The paraffin wax infiltrated samples were then transferred to the Tissue KEDEE embedding console (KD-BMIV tissue embedding centre) for embedding in paraffin blocks. Wax sections of 8µm thickness were cut on a mechanical Microtome (KEDEE, KD2258), and organised on permanent positive charged slides (Superfrost® plus, Thermo scientific) and left overnight on a slide warmer (Thermo scientific hot plate) at 40°C.

6.3.7.3 Probe Transcription and Labelling.

The probe transcription performed using gene-specific fragments (*BdMADS29*) which has been amplified from cDNA using T7 reverse (GAATTGTAATACGACTCACTATAGGG) appended primers It was amplified using DreamTaq PCR Master Mix 2X (Thermo Scientific) then purified using E.Z.N.A.® Cycle Pure Kit (Omega bio tek) according to the manufacturer protocol to get higher recovery of purified amplicon. The product was used to generate digoxigenin-labelled RNA probes.

The transcription was performed in-vitro in 20µl reactions for 2 hours at 37°C in the following reaction: 5 µl from 0.1% diethylpyrocarbonate (DEPC), 2 µl of dNTPs (10 µM), 2 µl from digoxigenin-11-UTP 200 nmol (57µl) (Roche), 1 µl of RiboLock RNase Inhibitor (40 U/µL) (Thermo fisher Scientific-2500U), 2 µl from T7 RNA polymerase 50U/µl (New England Biolabs), 2 µl of 10x RNA polymerase reaction buffer, finally 6 µl from *BdMADS3/58/13/21* Probe template (Purified DNA). The transcription reactions were precipitated in 4 M Lithium chloride (LiCl) and 3 volumes of absolute ethanol (60µl) overnight at -20°C. Reactions were spin for 10 minutes at 11,400 xg at 4°C and pellets resuspended in 20µl TE (100 mM Tris, 10 mM EDTA) buffer.

See section 3.3.4.4 for slide pre-treatment, development and mounting.

6.4 Results

6.4.1 Phylogeny cysteine protease C1, vacuolar processing enzymes (VPE) C13 and meta-caspase C14 proteases families in wild and domesticated species

Comparative proteomic analyses could offer a valuable insight into the conservation and evolution of the proteases linked to the PCD in the monocots and dicots species, which could help to clarify questions concerning the function of these members of families. Therefore, we have performed a phylogenetic analysis of these gene families in five species, including *Triticum aestivum*, *Hordeum vulgare*, *Oryza sativa*, *Brachypodium distachyon* and *Arabidopsis thaliana*.

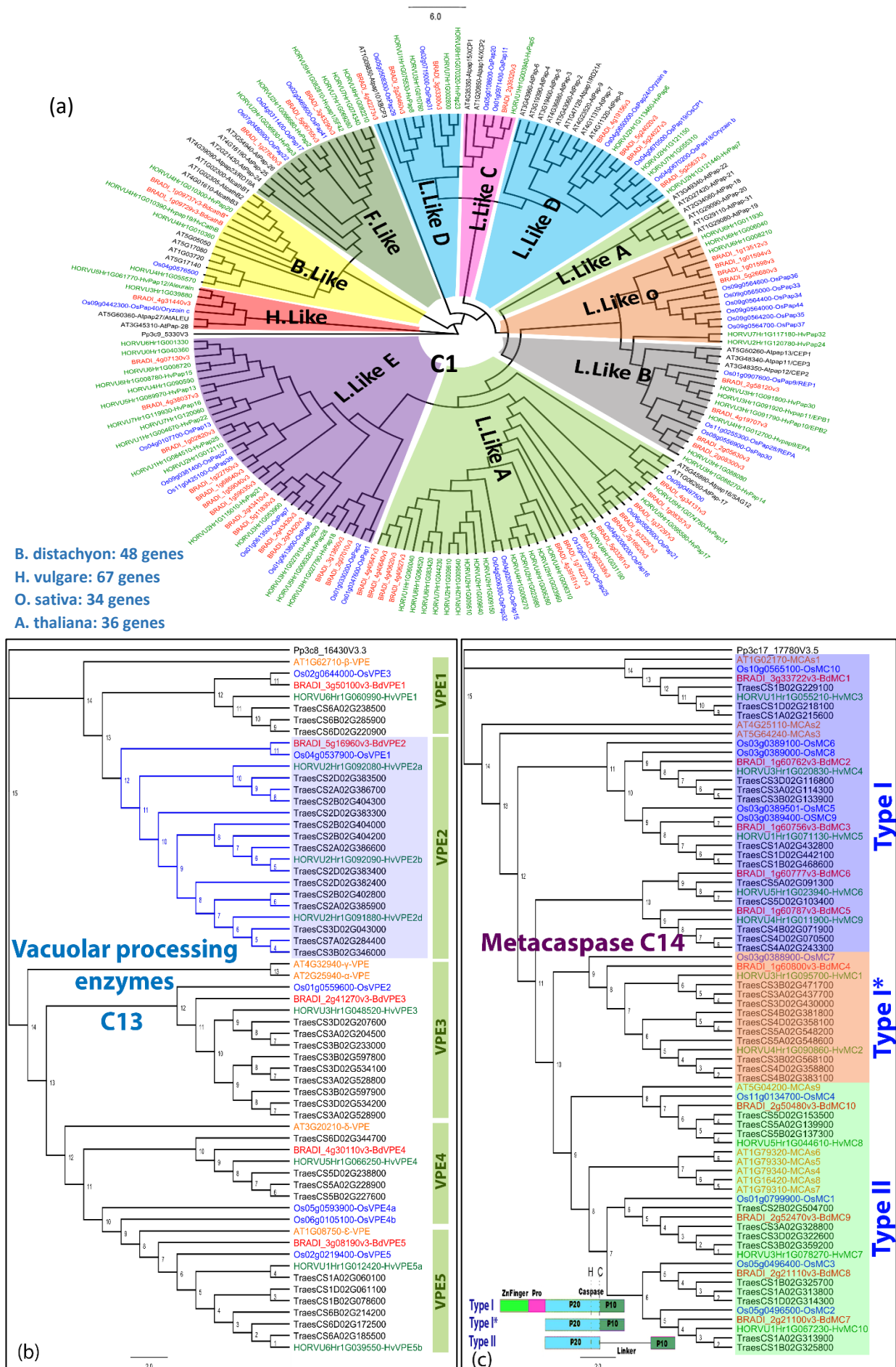


Figure 6-1: Phylogenetic Trees of papain C1 family, C13 Vacuolar processing enzyme (VPE) and Metacaspases C14 . (a) Phylogenetic tree of *Oryza sativa*, *Brachypodium distachyon*, *Hordeum vulgare* and *Arabidopsis* VPEs proteins. (b) Phylogenetic tree of C13 family vascular Processing Enzymes genes (VPE) in the listed four species. (c) C14 family metacaspases proteins orthologs in *Oryza sativa*, *Brachypodium distachyon*, *Hordeum vulgare* *Arabidopsis* Maximum Likelihood phylogenetic tree. The trees were generated using MAFFT alignment then Maximum Likelihood with 1000 replicates, Pro: proline domain, Physcomitrella orthologues used as an outgroup in C13 and C14, the metacaspase structure types were adapted from (Fagundes *et al.*, 2015).

The phylogenetic tree illustrates the C13 family Vacuolar processing enzymes (VPE) in three species, in order to identify the orthologue of each well-identified gene from barley and rice. The tree cluster two clades, which splits into two subclades. The first subclade includes the *HvVPE2* and its orthologues in *Brachypodium* and rice. *HvVPE2* (known as nucellain) express exclusively in the nucellus during grain development and might be involved in nucellar PCD (Radchuk *et al.*, 2010), In addition to *OsVPE1/ GLUP3* which was functionally characterised and found to play a crucial role in PCD process and required for an efficient glutelin processing, this is a major storage protein in rice (Wang *et al.*, 2009). The second subclade comprises *HvVPE1* which transcribes in starchy endosperm and increases during seed maturation. This sub-cluster also contains *OsVPE3* which promoted vacuole rupture and play a key role in stomatal development in rice (Lu *et al.*, 2016). These results suggest that their orthologues in *Brachypodium* may have a similar function, *BdVPE2* (BRADI_5g16960v3) is an orthologue of *HvVPE2* and *OsVPE1/ GLUP3*, whereas, *BdVPE1* (BRADI_3g50100v3) is an orthologue of *HvVPE1* (HORVU6Hr1G060990) and *OsVPE3*; however, additional genetic and biochemical investigations are required to validate this hypothesis.

The second clade in the tree groups *HvVPE5* (HORVU1Hr1G012420) with its orthologues in rice *OsVPE5* (Os02g0219400) and *Brachypodium* *BdVPE5* (BRADI_3g08190v3), whereas, *HvVPE4* (HORVU5Hr1G066250) showed a high similarity to *OsVPE4* (Os05g0593900) and *BdVPE4* (BRADI_4g30110v3). *HvVPE4* is exclusively expressed in the degeneration pericarp parenchyma between 4 and 12 days after anthesis and associated with apoptotic DNA degradation (Hatsugai *et al.*, 2015).

In rice, previous studies (Deng *et al.*, 2011) on vacuolar processing enzyme *OsVPEs* could identify only four paralogues. In addition to that, we could identify two new homologues

(Os02g0219400 and Os06g0105100) in rice, and we call it OsVPE4b and OsVPEc respectively, as it showed a close homology to the OsVPE4a.

Characterisation of the barley vacuolar processing enzyme (VPE) gene family identified seven transcripts (Radchuk *et al.*, 2010), however, when we blast these sequences against the new genome assembly (IBSC_v2), three of these transcripts (*HvVPE2a*, *b*, *c*) has been identified as one gene HORVU2Hr1G092080. Our study identified three new paralogues of *HvVPEs* in barley and named based on their similarity of the previously identified genes *HvVPE2e* (HORVU2Hr1G092090) or as additional number *HvVPE5a* (HORVU1Hr1G012420), *HvVPE5b* (HORVU6Hr1G039550).

Our study was successful in identifying five paralogues in *Brachypodium distachyon* and named based on their closeness to their known orthologues in barley.

The metacaspase C14 family, for *Hordeum vulgare*, identified protein sequences from (Bostancioglu *et al.*, 2018) study, were blaster against the IBSC_v2 genome assembly and was recognised with new IDs. Whereas for rice and *Arabidopsis*, was selected from (Fagundes *et al.*, 2015) the rice gene ID were converted into the new assembly IRGSP-1.0gene ID.

The C14 metacaspases family were divided into three categories based on their enzyme structure. The type I has a metacaspase domain at the C-terminus region, presenting a zinc finger motif in the N-terminus region and a proline domain. Similarly, type I* has both domains P20 and P10 however its N terminal is missing Zn Finger and proline domain. Metacaspase type II does not present the proline domain and the zinc finger but has a linker between caspase-like catalytic domains of 20 kDa (p20) and 10 kDa (p10) (Figure 6-1).

1.1. Expression profile analysis

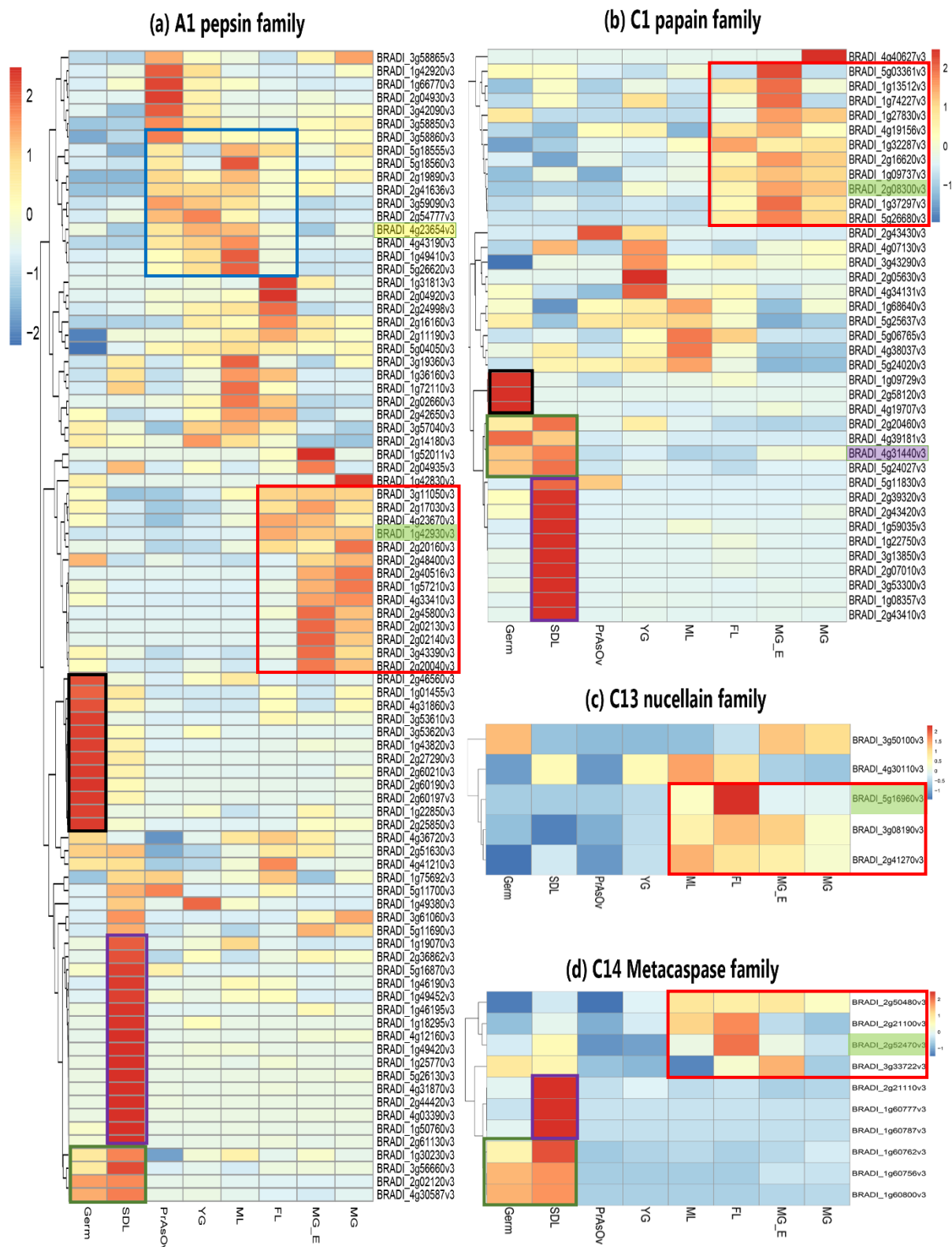


Figure 6-2: Hierarchically Clustered Expression Profile of Protease genes from pepsin-like (aspartic protease) family A1, papain family C1, vacuolar processing enzymes family C13 and metacaspase family C14 in vegetative and developing grain tissues of *Brachypodium*, Germ: germination, SDL: seedling, PrAsOv: pre-anthesis ovary, YG: young grain, ML: mid-length, FL: full-length, MG-E: mature grain without embryo, MG: mature grain. Rows are centred; unit variance scaling is applied to rows. Rows are clustered using correlation distance and average linkage. Highlighted green are the candidate genes tested by RT-qPCR.

In aspartic protease A1, 11 genes were highly expressed in grain development stages (Figure 6-2), grouped by a blue rectangle, this group contains *Bd* nucellin (BRADI_4g23654v3) the orthologues of the well-known nucellin in barley and *OsAsp1* in rice yellow shaded, 14 genes marked the late development and maturity stage (Figure 6-2) are assembled by a red rectangle, the expression of one candidate (BRADI_1g42930v3) from this group was validated using RT-qPCR. The expression showed that there was a cluster of 12 genes that are highly expressed in the germination tissues (Figure 6-2 black rectangle), and another clade was marked by high expression in the seedling stage (Figure 6-2 purple rectangle), four genes were exclusively expressed in the germination and seedling stages (green rectangle).

The C1 papain family heat-map illustrate the expression profile of all members of this family and group them in four categories, based on their expression pattern. The first cluster of eleven genes (Figure 6-2, red rectangle) that showed a high expression in Full and mature grain, this expression was validated by RT-qPCR relative expression for one candidate gene (BRADI_2g08300). The second clade (purple rectangle), was presented by ten genes that expressed exclusively in the seedling stage (Figure 2-6). In addition of other three genes that exclusively express in the germination (black rectangle). These results suggest that these genes are specific to only one tissue and mark the vegetative stages. Four genes appear to be transcribed highly in both samples (germination and seedling) (green rectangle), including *Bd* aleurain (BRADI_4g31440v3) (highlighted in purple), the orthologue of the well-characterised gene aleurain in barley, this gene start expression gradually from mature grain to reach its peak in the seedling. The rest of the members of the C1 papain family were expressed with different levels in all the tested samples.

The C13 nucellain family contain five members, and three of them showed high expression in developing grain including the *Bd* nucellain (BRADI_5g16960v3) which peaks in the full-length stage (highlighted green).

The C14 metacaspase family had three genes that were highly expressed in seedling tissues (purple rectangle), in addition to another three that were transcribed in both germinating and seedling (green rectangle). The rest of the family showed a high expression starting from the mid-length stages until the mature grain and was validated by RT-qPCR for one candidate gene (BRADI_2g52470v3).

6.4.2 Expression pattern of proteases and RNA-seq validation using RT-qPCR

In order to validate the RNA-seq experiment expression, RT-qPCR was performed. One candidate gene from each family was chosen based on their high expression in developing grain illustrated on the heat-map generated in (Figure 6-3) highlighted in green.

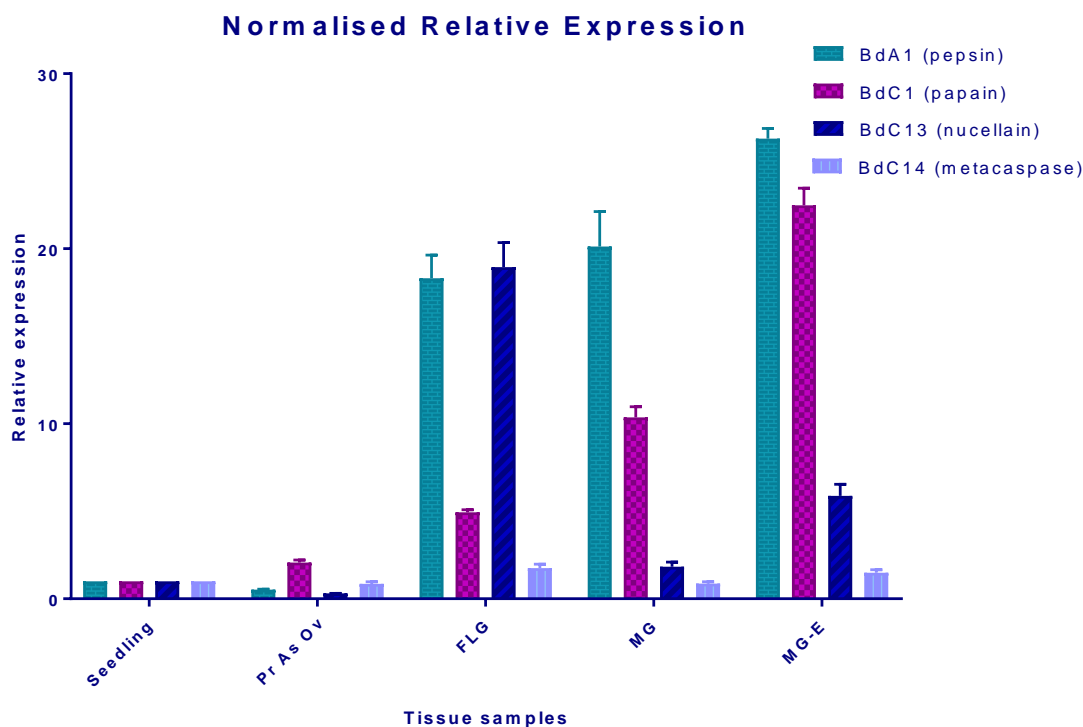


Figure 6-3: Normalised relative expression analysis of selected peptidase candidates using RT-qPCR, in Seedling, pre-anthesis (Pre-anthesis), Full-length grain (FL), mature grain without embryo (MG-E), and mature grain (MG). *BdA1*-pepsin-like (aspartic protease): BRADI_1g42930v3, *BdC1*-papain: BRADI_2g08300v3, *BdC13*-nucellain: BRADI_5g16960v3, *BdC14*-metacaspase (*BdMC9*): BRADI_2g52470v3, seedling tissue was used as a calibrator. Error bars indicate \pm SD ($n = 3$). Two way ANOVA was run to determine the effect of tissue difference and growth stage on gene expression, there was a significant variation in the time and tissues samples on genes expression $F(4, 12) = 3.06$, $P < 0.05$, and there was a significant variation in between genes with $F(3, 12) = 2.786$, $P < 0.08625$.

The RT-qPCR analysis showed that all selected genes from the four peptidases family were lowly expressed in vegetative tissues (seedling) and also reproductive tissues (Pr-anthesis), then start to increase with different levels and rates after anthesis in the grain filling and maturation.

The aspartic protease *BdA1* expression was found to be abundant in late stages of grain development with a peak in mature grain tissues. Similarly, the *BdC1*-papain follows the

same pattern with a slight difference in the full-length stage where the *BdA1* is higher than *BdC1*. In addition, The *BdC13* nucellain mRNA was observed after anthesis and reached its peak in the full-length tissues then decreases gradually during maturity. However, the *BdC14* (*BdMC9*) was lowly expressed in all stages.

6.4.3 Spatial distribution and progress of DNA fragmentation in the developing grain.

The programmed cell death is an essential cellular process that allows the organism to eliminate unwanted cells and develop normally. This happens through a cascade of reaction and leads to DNA fragmentation of the undesirable cell nucleus which can be labelled using the TUNEL assay to localise the apoptosis.

At pre-anthesis, The DNA cleavage was observed in all mesocarp, distributor uniformly around the ovule, this suggests that the carpel is preparing space for the ovule to cellularise and expand well before fertilisation (Figure 6-4(a)). After anthesis, the PCD was restricted only to the lateral side of the mesocarp and start with strong signal along the integuments as well as in the debris of degraded nucellar cells which appears before the endosperm, during young grain and mid-length stages, however there was no PCD signal in the palea at this stage (Figure 6-4 (c-e)). As the caryopsis develop, the DNA fragmentation continues in the mesocarp allowing the endosperm to fill the space at a full-length stage in addition to a weak signal of PCD along with the nucellar epidermis; the aleurone remains viable (Figure 6-4 (g)). At maturity stage, the DNA fragmentation appears in the starchy endosperm and the remaining cells of the ventral mesocarp cells adjacent to the vascular bundle, in addition, a strong signal of PCD event was observed in the palea especially around its vascular bundle in the spongy parenchyma cells and along the silicified cells, similar to the grain signal (Figure 6-4 (i)). There was no DNA degradation noticed in the nucellus cells which remains viable at maturity stage, it persists and occupies a considerable and valuable space from the endosperm.

These results suggest that the programmed cell death happens throughout the grain development with a different rate in different tissues in order to allow the generation of new cells and layers and clear the unwanted cells.

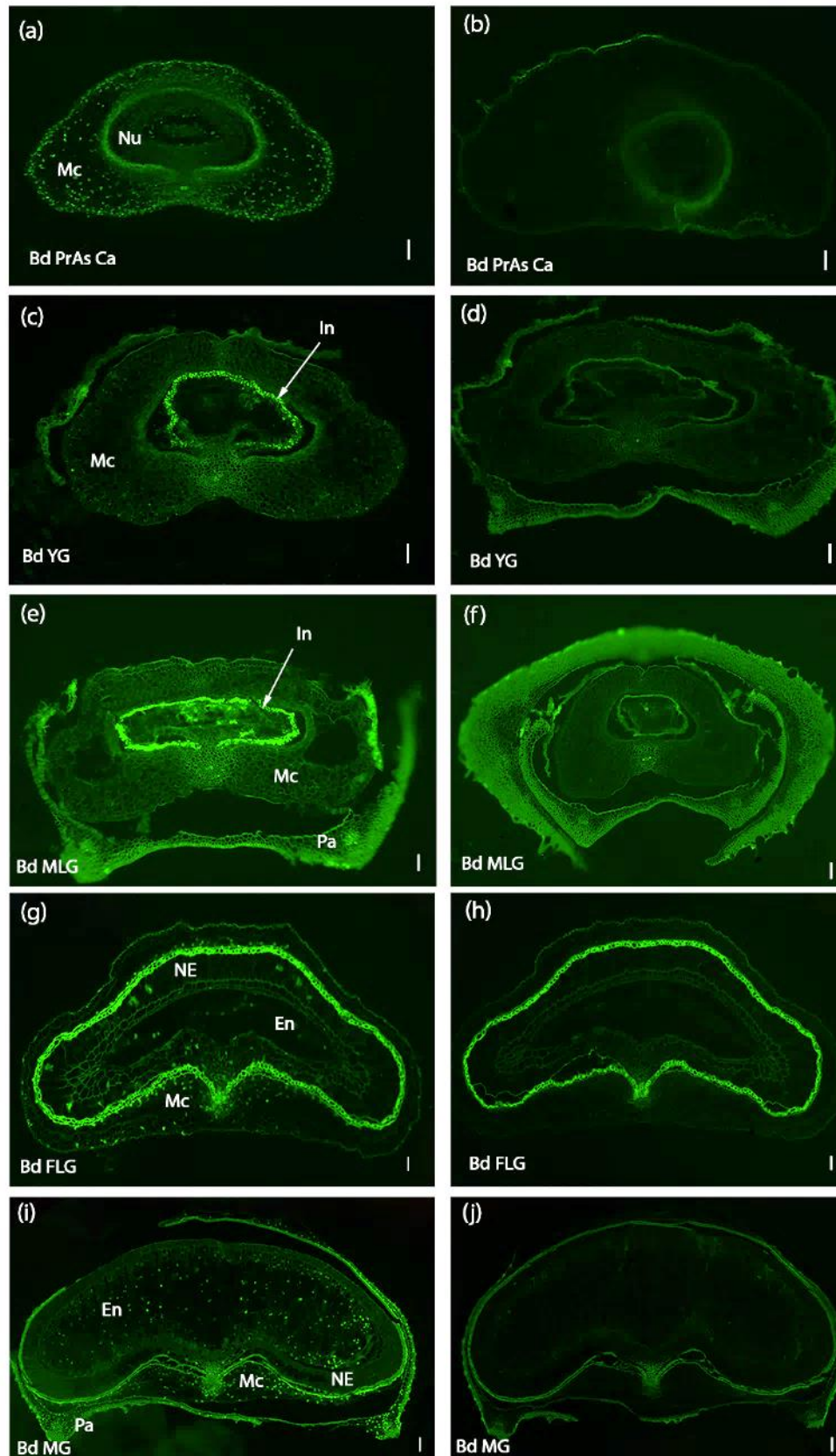


Figure 6-4: DNA fragmentation detected in *Brachypodium* grain section in five developmental stages, (a) *Bd* PrAs Ca: pre-anthesis carpel, (c) *Bd* YG: young grain 2 days post-anthesis (2 DPA), (e) mid-length grain at (6 DPA), (g) *Bd* FL: full length grain (10 DPA), (i) *Bd* MG: mature grain (20 DPA), (b-d-f-h-j) negative control, NE: nucellar epidermis, Nu: nucellus parenchyma cells, Lo: lodicule, Mc: mesocarp, Pa: palea, Ec: endocarp, In: integument, En: endosperm, NL: nucellar lysate, Scale bar 50 μ m.

6.4.4 Expression pattern of *BdMADS29* in the life cycle of the *Brachypodium* plant in different tissues samples using RT-PCR

The study carried out by (Yang *et al.*, 2012), showed that the knock-down transgenic seeds of *OsMADS29* had a significant decrease of the level of the only *OsVPE1*, which suggest that *OsVPE1* is a direct target of *OsMADS29*. Therefore, a comprehensive study was carried out to identify the expression pattern of their orthologue in *Brachypodium* *BdMADS29*.

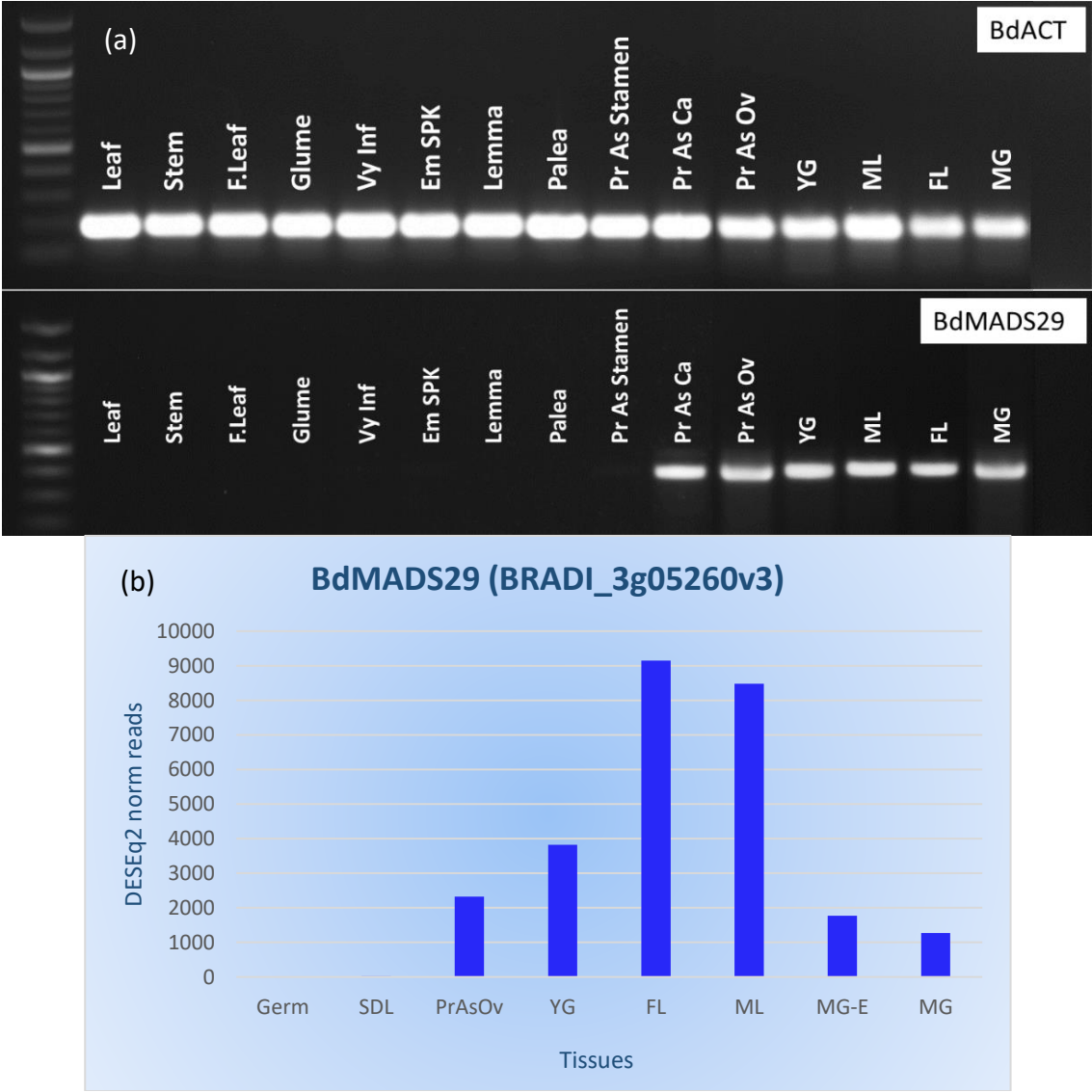


Figure 6-5: (a) RT-PCR illustrating the expression profile of *BdMADS29* in different tissues samples. (b) Read counts of *BdMADS29* transcript. CPM: Counts per million, Germ: germination, SDL: seedling, PrAsOv: Pre-anthesis ovary, YG: young grain, FL: full-length grain, ML: Mid-length grain, MG-E: mature grain without the embryo, MG: mature grain.

The results from the RT-PCR showed that *BdMADS29* expression (Figure 6-5 (a)) is restricted to the reproductive organs, precisely the carpel and the developing grain which represent the fourth whorl in flower. Starting from the Pre-anthesis carpel, the gene continues to express in the young grain, mid-length, full-length and mature grain with

variable levels. There was no expression detected in the tested vegetative tissues from early-stage till the flowering time, similarly the early stage of inflorescence, for instance, the very young inflorescence, the emerging spike, lemma, palea and pre-anthesis stamen. In order to quantify these expressions, an analysis of the normalised gene count from RNA-seq experiment was performed on similar stages and tissues. The bar chart (Figure 5-6 (b)) shows that there was no presence of the *BdMADS29* transcript in the two vegetative tissues, germination, and seedling. The detection of the expression started in the Pre-anthesis ovary with just above 2000 reads to continue to nearly double in the young grain with 4000 reads. The peak expression was observed in the grain during the mid and full-length stages with around 9000 reads. The expression plum in the following stages with a decrease of approximately 8000 reads to reach its lowest at maturity with just above 1000 reads.

6.4.5 The localisation of mRNA of *BdMADS29* in grain development from Pre-anthesis to mature grain

Due to the importance of *OsMADS29* during the PCD in rice, we selected its orthologue in *Brachypodium* and analysed its expression then localise this expression within the tissues using the mRNA *in situ* hybridisation (ISH).

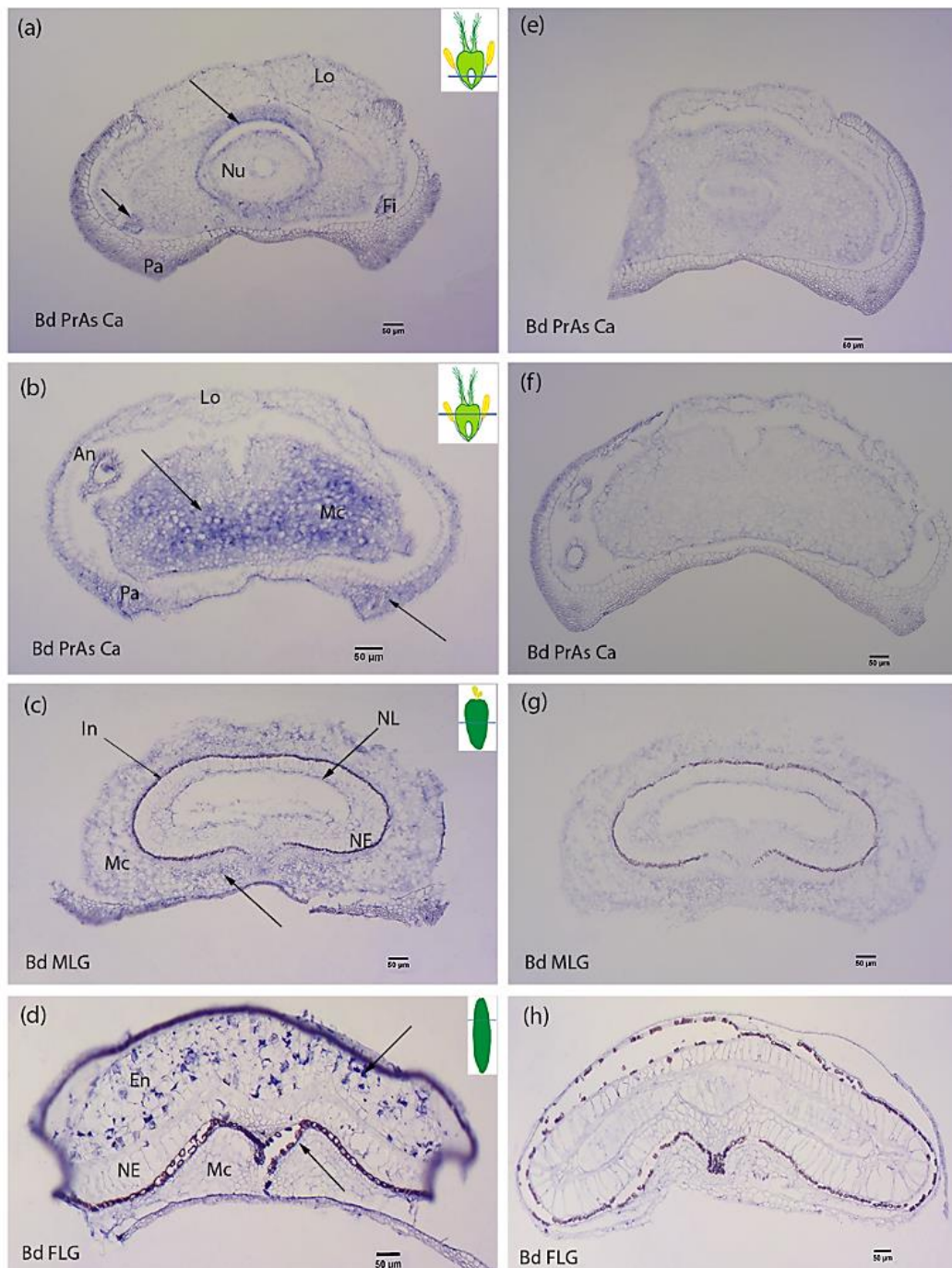


Figure 6-6: *in situ* hybridisation of *BdMADS29* mRNA in different grain developmental stages (a-d) antisense (e-h) sense. Arrows indicate the expression .An: anther, Fi: filament, NE: nucellar epidermis, Nu: nucellus parenchyma cells, Lo: lodicule, Mc: mesocarp, Pa: palea, Ec: endocarp, In: integument, En: endosperm, NL: nucellar lysate, Scale bar 50 µm.

The expression of *BdMADS29* mRNA was detected at Pre-anthesis carpel stage (Figure 6-6 (a)), in the endocarp adjacent to the developing nucellus, as well as, a very strong signal in the two filaments at the basal section of the carpel. There was no expression of *BdMADS29* in both lodicules nor the palea. At this stage, the carpel is enlarging and preparing for the

fertilisation of the ovule. This enlargement happened mostly at the apical side of the carpel; therefore, a section in this area was analysed and revealed that there was a significant expression of *BdMADS29* mRNA in the apical mesocarp adjacent to the two stigmas as well as, the surrounding of the vascular bundle of the palea. After anthesis especially at the mid-length stage, the mesocarp starts degenerating and reducing in size to leave space for the developing embryo and their nutritional supplements (endosperm), this is marketed with a detectable signal of *BdMADS29* in the whole mesocarp (Figure 6-6 (b)). As the grain increase in size, there was a degradation of the nucellar lysate to allow the endosperm to develop at the centre, and this was detected by the presence of the *BdMADS29* transcript. No signal was detected in the developing nucellar epidermis at the mid-length stage. The following full-length stage section showed a strong expression of the *BdMADS29* mRNA in the starchy endosperm as it undergoes maturation and dehydration; however, there was no expression in the nucellar epidermis nor in the aleurone at this stage (Figure 6-6 (d)). These results suggest that *BdMADS29* play a crucial role during the grain development by regulation the lifespan of the cells in different grain tissues. Despite focussing on the expression of *BdMADS29* in the development grain, our section revealed significant expression of in the spongy parenchymatous cells surrounding the vascular bundle.

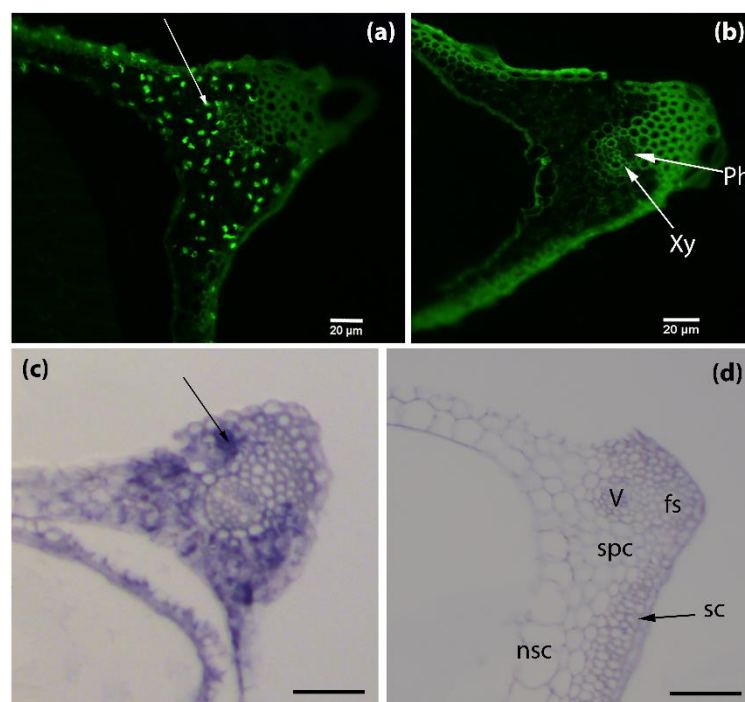


Figure 6-7: palea vascular bundle, (a-b) TUNEL assay and its negative control, (c-d) *in situ* hybridisation of *BdMADS29* transcript and its sense hybridisation (negative control), white arrow indicates the nucleus DNA fragmentation, black arrow indicates the signal of *BdMADS29* mRNA, **sc**: Upper epidermis (silicified cells), **nsc**: Lower epidermis (non-silicified cells), **fs**: fibrous sclerenchyma, **spc**: spongy parenchymatous cells, **Xy**: xylem, **Ph**: phloem, **V**: vascular bundle. Scale bar: a-b= 20 µm, c-d= 50µm.

At anthesis stage, the figure 6-7 shows a spatial distribution matching of the expression of *BdMADS29* mRNA with the fragmentation of the DNA in the palea spongy parenchymatous cells. These results suggest that *BdMADS29* may have a direct/ or indirect effect on the nucleus DNA to be fragmented and eventually death of the palea cells by the maturity stage.

6.5 Discussion

Researches in the past decades have highlighted the accuracy in time and tissues specific that undergo PCD in the lifespan of the plant cycle, which may be considered as an ordered process of selective cells removal, and how it is crucial to the normal development in the life cycle. These studies have focused mainly on domesticated species such as rice and barley showing That the PCD is an essential process of cereal seed development and germination (Young *et al.*, 1997; Yin & Xue, 2012; Radchuk *et al.*, 2018).

Our project is studying the PCD in a wild species, *Brachypodium*, and investigation the expression pattern and the spatial distribution in developing grain of *BdMADS29*, an orthologue of *OsMADS29*, shown to activate *OsVPE1* in rice. In addition, to the localisation of the DNA fragmentation reflecting the PCD during grain development, we revealed that while there is similarity in tissues specificity, between domesticated cereals and wild grass *Brachypodium*, there were differences in timing, pattern and progression of PCD in the caryopsis.

In the present study, the PCD was detected in the pre-anthesis carpel in *Brachypodium* which have not been previously reported before in the other species. Our results suggest that PCD is activated in *Brachypodium* nucellar and mesocarp cells before fertilisation. While in barley, it appears that initiation of PCD in nucellar cells was detected at anthesis, and in mesocarp occurs later between 2 and 6 days after anthesis (Radchuk *et al.*, 2010).

Brachypodium nucellar degradation happens in a radial pattern similar to barley, rice and wheat. However, the rate of nucellar cells breakdown is slow compared to barley. Rapid degeneration of the nucellus is thought to provide growth space and resources for the emerging endosperm in cereals (Domínguez & Cejudo, 2014). In maize, early endosperm development around the period of nucellar PCD influences growth and yield (Leroux *et al.*, 2014). We suggest that it is possible that slow nucellar disintegration in *Brachypodium*

delays its early endosperm development because of space reduction and low nutrient supply from disintegrated cells. In addition, *Brachypodium* nucellar epidermis enlarges greatly before undergoing PCD six days after anthesis, just to reduce in size; however, the nucellar epidermis does not collapse or disintegrate throughout grain development. As it has been shown that the nucellar epidermis serves as assimilate transport channel towards the endosperm in rice (ELLIS & CHAFFEY, 1987).

On the other hand, mesocarp degradation occurs faster in *Brachypodium* compared to barley. This corresponds with the earlier observation that mesocarp PCD is initiated earlier in *Brachypodium* compared to barley. Because barley mesocarp PCD is coordinated with endosperm development and only detected after endosperm cellularisation (Radchuk *et al.*, 2010). Similar harmonisation was confirmed in (Okada *et al.*, 2017) study for wheat as well beside showing that in unfertilised carpel the mesocarp cells do not degenerate, instead it swell sideways and push the lemma and palea apart in order to increase the chances of cross-pollination and fertilisation.

Our results suggest this harmonisation of endosperm development with mesocarp PCD is absent in *Brachypodium*. In addition, *Brachypodium* grain mesocarp cells degenerate rapidly at the lateral sides of the carpel. This is different from the centrifugal mesocarp degeneration pattern reported in barley, rice and wheat (Domínguez & Cejudo, 2014; Dominguez *et al.*, 2001; Radchuk *et al.*, 2010; Yin & Xue, 2012). We propose that this lateral degradation may result in the flat shape of the grain by allowing the endosperm to expand laterally and fill the space left by the dead mesocarp cells.

The endosperm PCD was observed after 15 days after anthesis in *Brachypodium*; however, it did not follow any systematic pattern. The same Random progression of endosperm PCD has also been detected in wheat, while the maize endosperm PCD proceeds in an organised manner. The difference in the pattern of endosperm PCD has been attributed to grain size. It has been suggested that while PCD can proceed randomly in the comparatively smaller endosperm of wheat, the large endosperm of maize require PCD to be executed in an organised fashion (Young & Gallie, 2000; Young & Gallie, 1999; Young *et al.*, 1997).

Proteases are known to contribute to the execution of PCD. Although the details of the contribution of individual genes are yet to be elucidated, the expression profile of several protease genes strongly coincides with PCD in different grain tissues (Buono *et al.*, 2019).

The programmed cell death regulation has been associated to a wide range of proteases in the different stages of the life cycle of the plant; however, functional evidence for the roles and contributions of plant proteases in PCD remains restricted. In the present study, we identified all *Brachypodium* genes that belonged to protease families and have been reported to play a role in the PCD. The families including A1, C1, C13 and C14. Phylogenetic analyses allowed the identification of *Brachypodium* orthologs of the selected families and name them based on their closeness to characterised and well-known genes in barley. Interestingly, the vacuolar processing enzyme C13 family tree was very diverse in barley, for instance, VPE2 was presented in rice (OsVPE1) and *Brachypodium* (BdVPE2) with only one version of gene, however, in barley it was presented with three paralogues *HvVPE2a*, *b* and *d*, the same was applied to the wheat. This could explain the rapid degeneration of the nucellus in barley and wheat, and on the other hand, the persistence of the nucellar epidermis in *Brachypodium*. It may also explain the size difference between the domesticated and wild species.

An RNA-seq investigation of the expression profile of genes in the four families revealed that aspartic protease family expression could be grouped in some protease that are tissues specific. The orthologue of the barley nucellin, was expressed mainly during grain development controlling the nucellin lifespan in the grain. Interestingly, the orthologue of the barley aleurain belonging to the papain family was expressed in the late stage of maturity to peak during germination and seedling which suggests that the aleurain is degenerating tissues to provide sufficient nutrient to the developing embryo. The nucellain *BdVPE2* from the vacuolar processing enzyme family C13, do not start straight after anthesis, mRNA of the gene appears during the mid-length then peak highly in the full-length stages where the endosperm is expanding and filling the degenerated nucellus cells. Similarly, the meta-caspase family presented three genes (*BdMC5*, *6* and *8*) that are seedling specific, suggesting that they play a vital role in the early stage of the plant development. In addition to two genes that are grain specific (*BdMC7* and *BdMC10*), however, *BdMC1* and *9* were abundant in the vegetative and reproductive tissues with high expression in full length and mature grain tissues. This family (VPEs) remarkably has genes that are the most highly expressed in the four families surveyed.

It was reported that *OsMADS29* is a direct regulator of *OsVPE1* (nucellain) (Os04g0537900), one of the vacuolar processing enzyme family C13 (Yang *et al.*, 2012), this gene is playing a crucial role in the PCD processes and essential for efficient glutelin hydrolysis. It was also shown that *OsMADS29* is a direct regulator of *Ospap31* (Cysteine protease) (Os02g0715000), from papain family C1 (Yin & Xue, 2012), which is close orthologue to *Hvpap23* (HORVU1Hr1G003940) in barley and belong to the cathepsin L-Like D group. Their close orthologues in *Arabidopsis* are the sisters, xylem cysteine protease1 and 2 (*XCP1*: AT4G35350, *XCP2*: AT1G20850) which, participate in the central vacuole hydrolysis in roots during xylogenesis. Due to the importance of *OsMADS29* during the PCD in rice, we selected its orthologue in *Brachypodium* and analysed the expression pattern using RT-PCR to specify the time and tissues where the gene is active, then localise this expression within the tissues using the mRNA *in situ* hybridisation (ISH). Our result revealed a relative overlapping pattern of the DNA degradation with the expression of *BdMADS29* mRNA suggesting its role in the PCD process.

6.6 Conclusion

Our study provides detailed PCD progress in a wild species, *Brachypodium*. The results and discussion present a comparative perspective to what is previously found in cereal grain PCD and highlight similarities and crucial, differences in the timing, pattern and progression of grain PCD between wild and domesticated cereals species. It appears that the rapid degeneration of the nucellar after anthesis in cereals stimulates fast endosperm cellularisation and expansion, resulting in large grains in domesticated cereals. We believe that the slow degenerating *Brachypodium* nucellus contributes to small grain size. Our theory is supported by the fact that *Brachypodium* presents just one ortholog of nucellar specific VPE *BdVPE2*, whereas barley and wheat have three (*HvVPE2a-d*). The lateral pattern of *Brachypodium* mesocarp degeneration appears to contribute to its dorsiventrally flattened grain shape.

Chapter 7

Summary and conclusion

Chapter 7. Summary and conclusion

Brachypodium distachyon is a great reference system to study cereals not only due to its high physiological and genetic similarity but also as an important species for tracing the origins of cereal grain diversity.

The ABCDE floral model that includes *MADS-box* genes was first established in *Arabidopsis thaliana* to illustrate that these genes control the morphogenesis of different floral organs. The recently increasing number of studies are in investigating the role of these genes in monocot species. In our project we demonstrate that some of *MADS-box* genes have essential roles beyond flowering in the wild grass *Brachypodium* as well as the cereals; wheat and barley and this might be generalised to all monocot's species.

Chapter 2 focused on diversification and conservation of *MADS-box* transcription factors across five species. This study has successfully identified all *MADS-box* MICK-type transcription factors (Figure 2-4) and characterised members of C and D class genes within this family in Wheat barley and *Brachypodium*. The results showed *MADS-box* genes were located across chromosomes in wheat rice and *Arabidopsis* (Figure 2-5) without any selection of a specific region. This study also proposes that some *MADS-box* sub-families witness gene duplication events after the divergence of *Brachypodieae* to have multiple copies in barley and wheat, such as *SOC*, *SEP* and *AGL17* sub-families (table 2-3). *SOC* sub-family presented by eight copies in wheat and barley compare to only six copies in *Brachypodium* and rice, and also *ANR1/AGL17* sub-family has expanded, with seven copies in wheat, but only three were represented in *Brachypodium*. Transcriptome analysis (figure 2-14) shows that some of these genes are highly expressed in during grain development in wheat and barley (figure 2-18) which propose that during the duplication events, neo-functionalisation might occur in these new copies giving rise to traits that are economically important in the crops (big endosperm size and more protein and fibre) that these crops have developed after domestication.

Gene duplication and potential functional divergence

The rice *OsMADS56* (Os10g0536100) orthologue was duplicated in wheat and its orthologues showed functional divergence as the three original homoeologues in A, B and

D (TraesCS1A02G199600, TraesCS1B02G214500 and TraesCS1D02G203300) have a significant expression in glumes and stamens (Figure S14), however, its paralogue (TraesCS1B02G214600, TraesCS1D02G203400 and TraesCS1A02G199900) (Figure 2-4) presented a totally different expression pattern with high activation in the seed coat and rachis (wheat-expression.com).

Similarly, barley has two orthologues of the rice SOC (Os06g0108500) gene that displayed some subfunctionalisation, while HORVU3Hr1G095080 expression spread across tissues such as germinating embryo, root-seedling, internode and inflorescence (Figure 2-18), its paralog (HORVU3Hr1G095090) expression was restricted to the inflorescence and the internodes in addition to being active in the seed at 15 days after anthesis, at this stage, the endosperm cells start accumulating proteins and starch granules (International Barley Genome Sequencing Consortium, 2012) (Figure 2-4).

These findings suggest that the new duplicated copies in these families gain a function in the caryopsis in both wheat and barley and potentially contribute to agronomic values to these crops. Therefore, it would be very interesting to explore the targets genes for these new duplicates.

On the other hand, expression quantification of the C and D class showed some extent of conservation between the wild grass *Brachypodium* and the domesticated cereals wheat and barley, except that barley lacks the ortholog of *MADS21*. All genes displayed similar trend of the transcript expression during flowering and grain development. *MADS3* was generally abundant in stamens while its paralog *MADS58* was expressed more in the carpel with significant activity in the young grain mid-length and full-length. This transcript was localised in *Brachypodium* grain section in the syncytial and cellularising endosperm during seed filling and in the aleurone layers which accumulates protein bodies and hydrolytic enzymes. These results suggest that the function of C-class genes has expanded in grasses to involve in different aspect of grain development and function from flower to flour.

Similarly, the D-class genes were expressed in comparable pattern across these species, where *MADS13* was restricted to the fourth whorl and continues to express highly during filling stages. However, *MADS21*, which is missing in barley, showed persistent expression in *Brachypodium* during grain development while in wheat it presented a high peak in only

pre-anthesis carpel (Figure 2-12, 3-11). This could suggest that domestication affected the expression and the function of *MADS21* in cereals.

Figure 7-1 summarises the expression pattern of all candidate genes and illustrates the function of *BdMADS3* in each stage in *Brachypodium*.

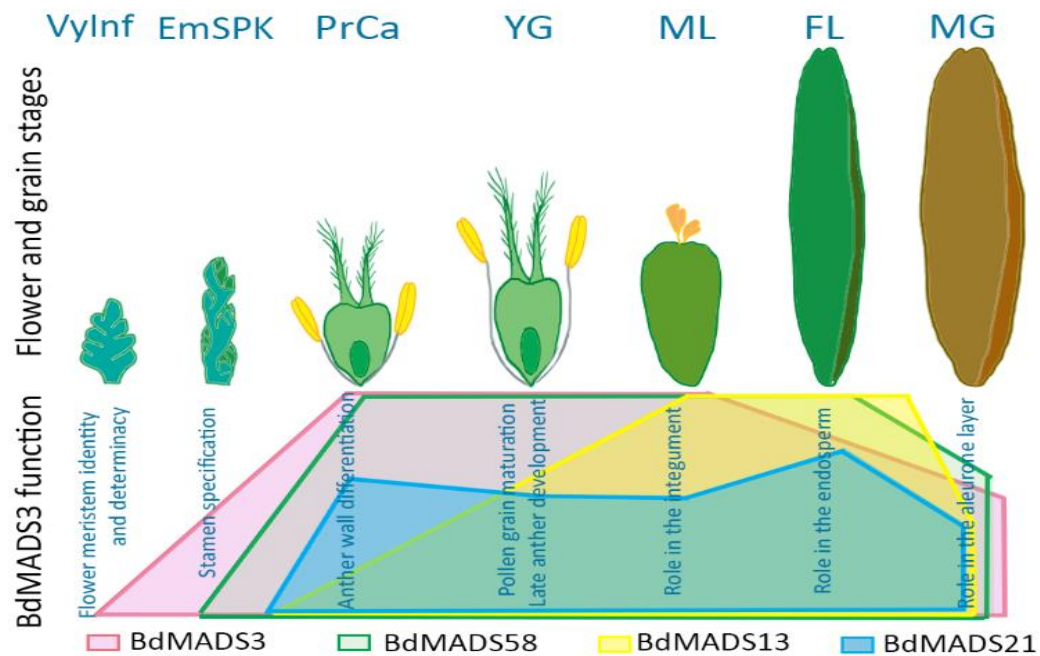


Figure 6-8: Illustration of the expression pattern of C- and D-lineage genes in *Brachypodium distachyon* and the function of *BdMADS3* during flower and grain development. Very young inflorescence (VyInf), Emerging spikelet (EmSPK), pre-anthesis carpel (PrCa), young grain (YG), mid-length grain (ML), full-length grain (FL), mature grain (MG).

In general, C and D-class gene expression patterns synchronised to essential developmental processes such as nucellar degeneration, endosperm proliferation, and differentiation into aleurone layer. The comparative analysis of the expression pattern of C-class and D-class genes between wheat, barley and *Brachypodium* showed some degree of similarity that was confirmed with RT-PCR, qRT-PCR and RNA-seq expression data. These findings agree with the conserved number of *AGAMOUS* genes found across these species (table 2-3) and confirm that *MADS58* has more function in ovule determinacy (Yamaguchi et al., 2006; Li et al., 2011), unlike its paralog *MADS3* that is more specific to stamen (Figure 2-13 and 2-17). In addition, we revealed that there is a significant expression of these genes in wheat, barley and *Brachypodium* during grain development and the temporal and spatial expression localise the transcript of C and D-class genes in the endocarp parenchyma tissue

at pre-anthesis that transform into integument and also in the developing endosperm and aleurone layer.

Overall, the transcript localisation assays revealed that C-class signal were presented in the mesocarp and integument at pre-anthesis then appears in the nucellar lysate in the subsequent stages; however, D-class genes were restricted to the integument at early stages then overlap with C-class at full-length in the syncytial endosperm then in the aleurone layer at maturity (Figure 7-2).

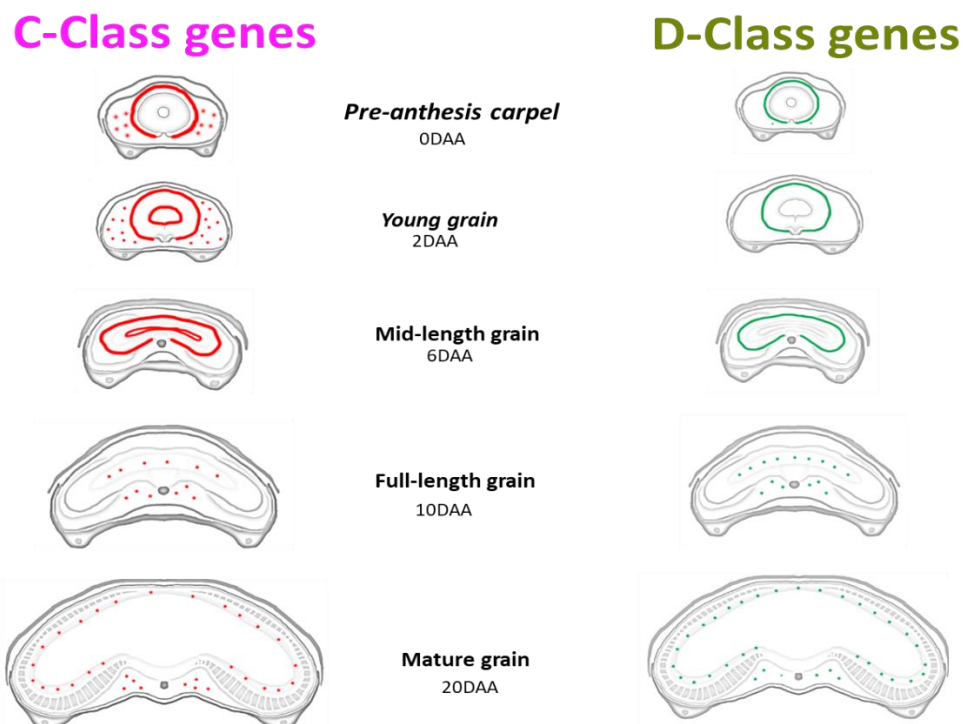


Figure 7-6-9: General illustration of the spatial localisation of the mRNA expression in grain sections in *Brachypodium*. DAA: days after anthesis.

These results propose that C-class genes are required during mesocarp and nucellar degradation, while D-class genes are essential for endocarp and integument formation and both play a role in the endosperm cellularisation and enzyme storage in the aleurone layer at maturity stage.

Chapter 4 was focused on the functional investigation of members of the C and D class. Functional analysis of *BdMADS3* and *BdMADS13* in *Brachypodium* revealed that the C-lineage gene *BdMADS3* is essential for proper spikelet formation and crucial for the late anther development and pollen maturation (Figure 4-3, 4-18). The loss-of-function of

BdMADS3 caused complete sterility and major interruption to the plant development; including, plant stature, flowering time, spikelet morphology, anther wall structure and pollen maturation. These results highlight that *BdMADS3* intervene in one or many biological processes that control these characteristics in the plant. It is worth mentioning that, this was the only line that showed a phenotype out of ten lines that carried the construct without displaying any physical alteration or dysfunction.

The *BdMADS13* transgenic line seems to be completely sterile with some malformation of the anther and decrease of the number of the floret (Figure 4-6). The reduction of the expression of the D-class gene *BdMADS13* affected the level of the reactive oxygen species scavenging genes which caused a delay in the tapetum layer PCD. The *BdMADS13* up-regulated genes were localised in the chloroplast where the photosynthesis occurs and these results are correlated with the *in situ* that showed expression of *BdMADS13* in the chlorenchyma cell of the carpel (Figure 3-8 (a-b)). These results were explained by the abundance of the ROS-scavenging genes that interrupt the redox homeostasis and cause the activation of programmed cell death proteases in the plant (Figure 5-12).

This programmed cell death was investigated in details in *Brachypodium* wild-type during grain development (Figure 6-4). The results showed that *Brachypodium* nucellar degradation happens in a radial pattern similar to barley, rice and wheat. However, the rate of nucellar cells breakdown is slow compared to barley. It appears that the rapid degeneration of the nucellus after anthesis in cereals stimulates fast endosperm cellularisation and expansion, resulting in large grains in domesticated cereals. This supposition is supported with the expansion in gene number in C13 family in wheat and barley for the nucellain gene (one copy in *Brachypodium* and four to five copies in wheat and barley). In addition, our findings propose that the lateral degradation of the mesocarp may result in the characteristic flat shape of the grain by allowing the endosperm to expand laterally and fill the space. Moreover, the temporal and spatial expression pattern of *BdMADS29* mRNA revealed a relative overlapping pattern with the DNA degradation suggesting its role in the PCD process.

Increasingly available wheat resources such as rapid lifecycle varieties, mutant collections (Wheat eFP, EMS mutation and Wheat TILLING) in addition to CRISPR-Cas9 System for genome editing that is gradually being available for wheat through institutes, for instance,

the National Institute of Agricultural Botany (NIAB) in the UK. Combined together with RNA-seq data sets frequently generated, would allow us to perform various comparative analyses between wild and domesticated grass to reveal the genetic control behind the common and unique features in grain development, and extend the understanding about the function of C and D-class in grain development.

This study has successfully established the expression profile at tissue and cell level, of C and D-class genes and highlighting the importance of some members using functional analysis in the cereal reference *Brachypodium*, it also suggested some candidate's genes that could be potential targets for these transcription factors.

The work presented in this thesis provide novel insights into key features of grain biology. However, further studies are required to confirm how these genes are affection the ROS homeostasis. Also, due to the availability of the wheat genome and the increasing number of the transgenic lines, it would be very useful to analyse the function of the genes that were duplicated in *SOC*, *SEP* and *AGL17* families.

References

Reference

Abdel-Aal, E.M., Young, J.C., Rabalski, I., Hucl, P., Fregeau-Reid, J., 2007. Identification and quantification of seed carotenoids in selected wheat species. *Journal of Agricultural and Food Chemistry*. **55**, 787-794.

Alvarez, J. & Smyth, D.R., 1999. CRABS CLAW and SPATULA, two *Arabidopsis* genes that control carpel development in parallel with *AGAMOUS*. *Development*. **126**, 2377-2386.

Alvarez-Buylla, E.R., Pelaz, S., Liljegren, S.J., Gold, S.E., Burgeff, C., Ditta, G.S., De Pouplana, L.R., Martínez-Castilla, L., Yanofsky, M.F., 2000. An ancestral *MADS-box* gene duplication occurred before the divergence of plants and animals. *Proceedings of the National Academy of Sciences*. **97**, 5328-5333.

Almaghrabi B, Ali MA, Zahoor A, Shah KH, Bohlmann H (2019) *Arabidopsis* thionin-like genes are involved in resistance against the beet-cyst nematode (*Heterodera schachtii*) 140:55-67

An, L. & You, R., 2004. Studies on nuclear degeneration during programmed cell death of synergid and antipodal cells in *Triticum aestivum*. *Sexual Plant Reproduction*. **17**, 195-201.

Andrews, S. & FastQC, A., 2015. A quality control tool for high throughput sequence data. 2010. URL: [Http://Www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc](http://www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc).

Andersson, A.A., Andersson, R., Piironen, V., Lampi, A., Nyström, L., Boros, D., Fraś, A., Gebruers, K., Courtin, C.M., Delcour, J.A., 2013. Contents of dietary fibre components and their relation to associated bioactive components in whole grain wheat samples from the HEALTHGRAIN diversity screen. *Food Chemistry*. **136**, 1243-1248.

Angenent, G.C., Franken, J., Busscher, M., van Dijken, A., van Went, J.L., Dons, H.J., van Tunen, A.J., 1995. A novel class of *MADS box* genes is involved in ovule development in *Petunia*. *The Plant Cell*. **7**, 1569-1582.

Arruda, P., Sodek, L., da Silva, W.J., 1982. Lysine-ketoglutarate reductase activity in developing maize endosperm. *Plant Physiology*. **69**, 988-989.

Asakura, T., Watanabe, H., Abe, K., Arai, S., 1995. Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases. *European Journal of Biochemistry*. **232**, 77-83.

Aune, D., Keum, N., Giovannucci, E., Fadnes, L.T., Boffetta, P., Greenwood, D.C., Tonstad, S., Vatten, L.J., Riboli, E., Norat, T., 2016. Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: systematic review and dose-response meta-analysis of prospective studies. *Bmj*. **353**, i2716.

References

- Barberá, P., Romero-Zarco, C., Aedo, C., 2017. Taxonomic revision of *Trisetum* section *Acrospelion* (Poaceae: Pooideae: Aveninae) from Eurasia. *Systematic Botany*. **42**, 754-781.
- Barton, M.K., 2010. Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. *Developmental Biology*. **341**, 95-113.
- Becker, A., Kaufmann, K., Freialdenhoven, A., Vincent, C., Li, M., Saedler, H., Theissen, G., 2002. A novel *MADS-box* gene subfamily with a sister-group relationship to class B floral homeotic genes. *Molecular Genetics and Genomics*. **266**, 942-950.
- Becker, A. & Theissen, G., 2003. The major clades of *MADS-box* genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution*. **29**, 464-489.
- Bemer, M., Gordon, J., Weterings, K., Angenent, G.C., 2009. Divergence of Recently Duplicated M γ -Type *MADS-box* Genes in *Petunia*. *Molecular Biology and Evolution*. **27**, 481-495.
- Bemer, M., Heijmans, K., Airoidi, C., Davies, B., Angenent, G.C., 2010. An atlas of type I *MADS* box gene expression during female gametophyte and seed development in *Arabidopsis*. *Plant Physiology*. **154**, 287-300.
- Bemer, M., Wolters-Arts, M., Grossniklaus, U., Angenent, G.C., 2008. The *MADS* domain protein DIANA acts together with *AGAMOUS-LIKE80* to specify the central cell in *Arabidopsis* ovules. *The Plant Cell*. **20**, 2088-2101.
- Bennett, M.D., Smith, J.B., Barclay, I., 1975. Early seed development in the Triticeae. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*. **272**, 199-227.
- Bercy, J., Dubois, E., Messenguy, F., 1987. Regulation of arginine metabolism in *Saccharomyces cerevisiae*: expression of the three *ARGR* regulatory genes and cellular localization of their products. *Gene*. **55**, 277-285.
- Bhandari, N.N., 1984. The microsporangium. *Embryology of angiosperms*. Springer. 53-121.
- Bhargava, A., Clabaugh, I., To, J.P., Maxwell, B.B., Chiang, Y., Schaller, G.E., Loraine, A., Kieber, J.J., 2013. Identification of cytokinin-responsive genes using microarray meta-analysis and RNA-Seq in *Arabidopsis*. *Plant Physiology*. **162**, 272-294.
- Bi, X., Khush, G.S., Bennett, J., 2005. The rice nucellin gene ortholog *OsAsp1* encodes an active aspartic protease without a plant-specific insert and is strongly expressed in early embryo. *Plant and Cell Physiology*. **46**, 87-98.
- Bo Wei, Rong-Zhi Zhang, Juan-Juan Guo, Dan-Mei Liu, Ai-Li Li, Ren-Chun Fan, Long Mao, Xiang-Qi Zhang, 2014. Genome-Wide Analysis of the *MADS-box* Gene Family in *Brachypodium distachyon*. *PLoS One*. **9**, e84781.

References

- Borrill, P., Harrington, S.A., Simmonds, J., Uauy, C., 2019. Identification of transcription factors regulating senescence in wheat through gene regulatory network modelling. *Plant Physiology*. pp. 00380.2019.
- Bostancioglu, S.M., Tombuloglu, G., Tombuloglu, H., 2018. Genome-wide identification of barley MCs (metacaspases) and their possible roles in boron-induced programmed cell death. *Molecular Biology Reports*. **45**, 211-225.
- Bouchard, C., Staller, P., Eilers, M., 1998. Control of cell proliferation by Myc. *Trends in Cell Biology*. **8**, 202-206.
- Bouchez, O., Huard, C., Lorrain, S., Roby, D., Balague, C., 2007. Ethylene Is One of the Key Elements for Cell Death and Defense Response Control in the *Arabidopsis* Lesion Mimic Mutant vad1. *Plant Physiology*. **145**, 465-477.
- Bowman, J.L., Drews, G.N., Meyerowitz, E.M., 1991. Expression of the *Arabidopsis* floral homeotic gene *AGAMOUS* is restricted to specific cell types late in flower development. *The Plant Cell*. **3**, 749-758.
- Bowman, J.L. & Smyth, D.R., 1999. CRABS CLAW, a gene that regulates carpel and nectary development in *Arabidopsis*, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development*. **126**, 2387-2396.
- Bowman, J.L., Smyth, D.R., Meyerowitz, E.M., 1989. Genes directing flower development in *Arabidopsis*. *The Plant Cell*. **1**, 37-52.
- Brown, R.C. & Lemmon, B.E., 2001. The cytoskeleton and spatial control of cytokinesis in the plant life cycle. *Protoplasma*. **215**, 35-49.
- Buono, R.A., Hudecek, R., Nowack, M.K., 2019. Plant proteases during developmental programmed cell death. *Journal of Experimental Botany*. **70**, 2097-2112.
- Burke, S.V., Wysocki, W.P., Zuloaga, F.O., Craine, J.M., Pires, J.C., Edger, P.P., Mayfield-Jones, D., Clark, L.G., Kelchner, S.A., Duvall, M.R., 2016. Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). *BMC Plant Biology*. **16**, 140.
- Cai, Q., Yuan, Z., Chen, M., Yin, C., Luo, Z., Zhao, X., Liang, W., Hu, J., Zhang, D., 2014. Jasmonic acid regulates spikelet development in rice. *Nature Communications*. **5**, 3476.
- Callens, C., Tucker, M.R., Zhang, D., Wilson, Z.A., 2018. Dissecting the role of *MADS-box* genes in monocot floral development and diversity. *Dissecting the Role of MADS-box Genes in Monocot Floral Development and Diversity*.
- Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., Garcia-Ponce, B., Sánchez, María de la Paz, Álvarez-Buylla, E.R., Garay-Arroyo, A., 2019. *MADS-box* genes are key components of genetic regulatory networks involved in abiotic stress and plastic developmental responses in plants. *Frontiers in Plant Science*. **10**, 853.

References

- Catalán, P., Kellogg, E.A., Olmstead, R.G., 1997. Phylogeny of *Poaceae* Subfamily Pooideae Based on ChloroplastndhF Gene Sequences. *Molecular Phylogenetics and Evolution*. **8**, 150-166.
- Catalan, P., López-Álvarez, D., Díaz-Pérez, A., Sancho, R. and López-Herránz, M.L., 2015. Phylogeny and evolution of the genus *Brachypodium*. *Genetics and genomics of Brachypodium*. Springer. 9-38.
- Catalán, P., Müller, J., Hasterok, R., Jenkins, G., Mur, L.A., Langdon, T., Betekhtin, A., Siwinska, D., Pimentel, M., López-Alvarez, D., 2012. Evolution and taxonomic split of the model grass *Brachypodium distachyon*. *Annals of Botany*. **109**, 385-405.
- Catalán, P. & Olmstead, R.G., 2000. Phylogenetic reconstruction of the genus *Brachypodium* P. Beauv. (*Poaceae*) from combined sequences of chloroplastndhF gene and nuclear ITS. *Plant Systematics and Evolution*. **220**, 1-19.
- Catalán, P., Shi, Y., Armstrong, L., Draper, J., Stace, C.A., 1995. Molecular phylogeny of the grass genus *Brachypodium* P. Beauv. based on RFLP and RAPD analysis. *Botanical Journal of the Linnean Society*. **117**, 263-280.
- Cejudo, F.J., Ghose, T.K., Stabel, P., Baulcombe, D.C., 1992. Analysis of the gibberellin-responsive promoter of a cathepsin B-like gene from wheat. *Plant Molecular Biology*. **20**, 849-856.
- Cercós, M., Gómez-Cadenas, A., Ho, T.D., 1999. Hormonal regulation of a cysteine proteinase gene, EPB-1, in barley aleurone layers: cis-and trans-acting elements involved in the co-ordinated gene expression regulated by gibberellins and abscisic acid. *The Plant Journal*. **19**, 107-118.
- Charmet, G., 2011. Wheat domestication: lessons for the future. *Comptes Rendus Biologies*. **334**, 212-220.
- Charoensawan, V., Wilson, D., Teichmann, S.A., 2010. Genomic repertoires of DNA-binding transcription factors across the tree of life. *Nucleic Acids Research*. **38**, 7364-7377.
- Chae, E., Tan, Q.K., Hill, T.A., Irish, V.F., 2008. An *Arabidopsis* F-box protein acts as a transcriptional co-factor to regulate floral development. *Development*. **135**, 1235-1245.
- Chen, F. & Foolad, M.R., 1997. Molecular organization of a gene in barley which encodes a protein similar to aspartic protease and its specific expression in nucellar cells during degeneration. *Plant Molecular Biology*. **35**, 821-831.
- Chen, J., Ding, J., Ouyang, Y., Du, H., Yang, J., Cheng, K., Zhao, J., Qiu, S., Zhang, X., Yao, J., 2008. A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proceedings of the National Academy of Sciences*. **105**, 11436-11441.

References

- Chen, J., Janssen, B., Williams, A., Sinha, N., 1997. A gene fusion at a homeobox locus: alterations in leaf shape and implications for morphological evolution. *The Plant Cell*. **9**, 1289-1304.
- Chen, R., Shen, L., Wang, D., Wang, F., Zeng, H., Chen, Z., Peng, Y., Lin, Y., Tang, X., Deng, M., 2015. A gene expression profiling of early rice stamen development that reveals inhibition of photosynthetic genes by *OsMADS58*. *Molecular Plant*. **8**, 1069-1089.
- Chen, Z., Liu, X., Wang, D., Chen, R., Zhang, X., Xu, Z., Bai, S., 2018. Transcription factor *OsTGA10* is a target of the MADS protein *OsMADS8* and is required for tapetum development. *Plant Physiology*. **176**, 819-835.
- Cho, S., Yoo, S., Zhang, H., Pandeya, D., Koh, H., Hwang, J., Kim, G., Paek, N., 2013. The rice narrow leaf2 and narrow leaf3 loci encode WUSCHEL-related homeobox 3 A (*OsWOX3A*) and function in leaf, spikelet, tiller and lateral root development. *New Phytologist*. **198**, 1071-1084.
- Chochois, V., Vogel, J.P., Watt, M., 2012. Application of *Brachypodium* to the genetic improvement of wheat roots. *Journal of Experimental Botany*. **63**, 3467-3474.
- Ciaffi, M., Paolacci, A.R., Tanzarella, O.A., Porceddu, E., 2011. Molecular aspects of flower development in grasses. *Sexual Plant Reproduction*. **24**, 247-282.
- Clavijo, B.J., Venturini, L., Schudoma, C., Accinelli, G.G., Kaithakottil, G., Wright, J., Borrill, P., Kettleborough, G., Heavens, D., Chapman, H., 2017. An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Research*. **27**, 885-896.
- Coen, E.S. & Meyerowitz, E.M., 1991. The war of the whorls: genetic interactions controlling flower development. *Nature*. **353**, 31.
- Coll, N.S., Vercammen, D., Smidler, A., Clover, C., Van Breusegem, F., Dangl, J.L., Epple, P., 2010. *Arabidopsis* type I metacaspases control cell death. *Science*. **330**, 1393-1397.
- Collins, E.J., 1918. The structure of the integumentary system of the barley grain in relation to localized water absorption and semi-permeability. *Annals of Botany*. **32**, 381-414.
- Collins, H.M., Burton, R.A., Topping, D.L., Liao, M., Bacic, A., Fincher, G.B., 2010. Variability in fine structures of noncellulosic cell wall polysaccharides from cereal grains: potential importance in human health and nutrition. *Cereal Chemistry*. **87**, 272-282.
- Colombo, L., Franken, J., Koetje, E., van Went, J., Dons, H.J., Angenent, G.C., van Tunen, A.J., 1995. The *Petunia* MADS box gene *FBP11* determines ovule identity. *The Plant Cell*. **7**, 1859-1868.

References

- Colombo, M., Masiero, S., Vanzulli, S., Lardelli, P., Kater, M.M., Colombo, L., 2008. AGL23, a type I MADS-box gene that controls female gametophyte and embryo development in *Arabidopsis*. *The Plant Journal*. **54**, 1037-1048.
- Coordinators, N.R., 2017. Database resources of the national center for biotechnology information. *Nucleic Acids Research*. **45**, D12.
- Cosio, C., Ranocha, P., Francoz, E., Burlat, V., Zheng, Y., Perry, S.E., Ripoll, J., Yanofsky, M., Dunand, C., 2017. The class III peroxidase PRX 17 is a direct target of the MADS-box transcription factor *AGAMOUS-LIKE15* (AGL 15) and participates in lignified tissue formation. *New Phytologist*. **213**, 250-263.
- Craine, J.M., Ocheltree, T.W., Nippert, J.B., Towne, E.G., Skibbe, A.M., Kembel, S.W., Fargione, J.E., 2013. Global diversity of drought tolerance and grassland climate-change resilience. *Nature Climate Change*. **3**, 63-67.
- Cui, R., Han, J., Zhao, S., Su, K., Wu, F., Du, X., Xu, Q., Chong, K., Theißen, G., Meng, Z., 2010. Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). *The Plant Journal*. **61**, 767-781.
- Daneva, A., Gao, Z., Van Durme, M., Nowack, M.K., 2016. Functions and regulation of programmed cell death in plant development. *Annual Review of Cell and Developmental Biology*. **32**, 441-468.
- Davidson RM, Gowda M, Moghe G, Lin H, Vaillancourt B, Shiu S, Jiang N, Robin Buell C (2012) Comparative transcriptomics of three *Poaceae* species reveals patterns of gene expression evolution 71:492-502. doi: 10.1111/j.1365-3113X.2012.05005.x
- Davis, J.I. & Soreng, R.J., 2007. A preliminary phylogenetic analysis of the grass subfamily Pooideae (Poaceae), with attention to structural features of the plastid and nuclear genomes, including an intron loss in GBSSI. *Aliso*.
- De Bodt, S., Raes, J., Florquin, K., Rombauts, S., Rouze, P., Theißen, G., Van de Peer, Y., 2003. Genomewide structural annotation and evolutionary analysis of the type I *MADS-box* genes in plants. *Journal of Molecular Evolution*. **56**, 573-586.
- de Camargos, L.F., Fraga, O.T., Oliveira, C.C., da Silva, Jose Cleydson Ferreira, Fontes, E.P.B., Reis, P.A.B., 2019. Development and cell death domain-containing asparagine-rich protein (DCD/NRP): an essential protein in plant development and stress responses. *Theoretical and Experimental Plant Physiology*. **31**, 59-70.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. **29**, 15-21.
- Doig, R.I., Colborne, A.J., Morris, G., Laidman, D.L., 1975. The induction of glyoxysomal enzyme activities in the aleurone cells of germinating wheat. *Journal of Experimental Botany*. **26**, 387-398.

References

- Domínguez, F. & Cejudo, F.J., 2014. Programmed cell death (PCD): an essential process of cereal seed development and germination. *Frontiers in Plant Science*. **5**, 366.
- Domínguez, F. & Cejudo, F.J., 1998. Germination-related genes encoding proteolytic enzymes are expressed in the nucellus of developing wheat grains. *The Plant Journal*. **15**, 569-574.
- Dominguez, F., Moreno, J., Cejudo, F.J., 2001. The nucellus degenerates by a process of programmed cell death during the early stages of wheat grain development. *Planta*. **213**, 352-360.
- Dong, Q., Wang, F., Kong, J., Xu, Q., Li, T., Chen, L., Chen, H., Jiang, H., Li, C., Cheng, B., 2019. Functional analysis of ZmMADS1 reveals its role in regulating starch biosynthesis in maize endosperm. *Scientific Reports*. **9**, 3253.
- Draper, J., Mur, L.A.J., Jenkins, G., Ghosh-Biswas, G.C., Bablak, P., Hasterok, R., Routledge, A.P.M., 2001. *Brachypodium distachyon*. A New Model System for Functional Genomics in Grasses. *Plant Physiology*. **127**, 1539-1555.
- Drea, S., Leader, D.J., Arnold, B.C., Shaw, P., Dolan, L., Doonan, J.H., 2005. Systematic spatial analysis of gene expression during wheat caryopsis development. *The Plant Cell*. **17**, 2172-2185.
- Dreni, L., Jacchia, S., Fornara, F., Fornari, M., Ouwerkerk, P.B.F., An, G., Colombo, L., Kater, M.M., 2007. The D-lineage MADS-box gene *OsMADS13* controls ovule identity in rice. *The Plant Journal*. **52**, 690-699.
- Dreni, L. & Kater, M.M., 2014. MADS reloaded: evolution of the *AGAMOUS* subfamily genes. *New Phytologist*. **201**, 717-732.
- Dreni, L., Osnato, M., Kater, M.M., 2013. The ins and outs of the rice *AGAMOUS* subfamily. *Molecular Plant*. **6**, 650-664.
- Drews, G.N., Bowman, J.L., Meyerowitz, E.M., 1991. Negative regulation of the *Arabidopsis* homeotic gene *AGAMOUS* by the *APETALA2* product. *Cell*. **65**, 991-1002.
- Dubois, E., Bercy, J., Descamps, F., Messenguy, F., 1987. Characterization of two new genes essential for vegetative growth in *Saccharomyces cerevisiae*: nucleotide sequence determination and chromosome mapping. *Gene*. **55**, 265-275.
- Dubcovsky, J. & Dvorak, J., 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*. **316**, 1862-1866.
- Dyer, B.W., Ferrer, F.A., Klinedinst, D.K., Rodriguez, R., 2000. A noncommercial dual luciferase enzyme assay system for reporter gene analysis. *Analytical Biochemistry*. **282**, 158-161.

References

- Ehlers, K., Bhide, A.S., Tekleyohans, D.G., Wittkop, B., Snowdon, R.J., Becker, A., 2016. The MADS box genes ABS, SHP1, and SHP2 are essential for the coordination of cell divisions in ovule and seed coat development and for endosperm formation in *Arabidopsis thaliana*. *PloS One*. **11**, e0165075.
- El Baidouri, M., Murat, F., Veyssiere, M., Molinier, M., Flores, R., Burlot, L., Alaux, M., Quesneville, H., Pont, C., Salse, J., 2017. Reconciling the evolutionary origin of bread wheat (*Triticum aestivum*). *New Phytologist*. **213**, 1477-1486.
- ELLIS, J.R. & CHAFFEY, N.J., 1987. Structural differentiation of the nucellar epidermis in the caryopsis of rice (*Oryza sativa*). *Annals of Botany*. **60**, 671-675.
- Engell, K., 1994. Embryology of barley. IV. Ultrastructure of the antipodal cells of *Hordeum vulgare* L. cv. Bomi before and after fertilization of the egg cell. *Sexual Plant Reproduction*. **7**, 333-346.
- Evers, A.D. & Reed, M., 1988. Some novel observations by scanning electron microscopy on the seed coat and nucellus of the mature wheat grain. *Cereal Chemistry*.
- Fagundes, D., Bohn, B., Cabreira, C., Leipelt, F., Dias, N., Bodanese-Zanettini, M.H., Cagliari, A., 2015. Caspases in plants: metacaspase gene family in plant stress responses. *Functional & Integrative Genomics*. **15**, 639-649.
- Feldman, M., Lupton, F., Miller, T.E., 1995. Evolution of crop Plants. Edited by: Smartt J, Simmonds NW. Harlow Essex: Longman Scientific & Technical. 184-192.
- Fjellheim, S., Boden, S., Trevaskis, B., 2014. The role of seasonal flowering responses in adaptation of grasses to temperate climates. *Frontiers in Plant Science*. **5**, 431.
- Fornara, F., Pařenicová, L., Falasca, G., Pelucchi, N., Masiero, S., Ciannamea, S., Lopez-Dee, Z., Altamura, M.M., Colombo, L., Kater, M.M., 2004. Functional characterization of OsMADS18, a member of the AP1/SQUA subfamily of MADS box genes. *Plant Physiology*. **135**, 2207-2219.
- Forestan, C., Meda, S., Varotto, S., 2010. ZmPIN1-mediated auxin transport is related to cellular differentiation during maize embryogenesis and endosperm development. *Plant Physiology*. **152**, 1373-1390.
- Francin-Allami, M., Alvarado, C., Daniel, S., Geairon, A., Saulnier, L., Guillon, F., 2019. Spatial and temporal distribution of cell wall polysaccharides during grain development of *Brachypodium distachyon*. *Plant Science*. **280**, 367-382.
- Freeman, P.L. & Palmer, G.H., 1984. The structure of the pericarp and testa of barley. *Journal of the Institute of Brewing*. **90**, 88-94.
- Fuller, D.Q., 2007. Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. *Annals of Botany*. **100**, 903-924.

References

- Galimba, K.D. & Di Stilio, V.S., 2015. Sub-functionalization to ovule development following duplication of a floral organ identity gene. *Developmental Biology*. **405**, 158-172.
- Gao, X., Liang, W., Yin, C., Ji, S., Wang, H., Su, X., Guo, C., Kong, H., Xue, H., Zhang, D., 2010. The SEPALLATA-like gene *OsMADS34* is required for rice inflorescence and spikelet development. *Plant Physiology*. **153**, 728-740.
- Gechev, T.S., Van Breusegem, F., Stone, J.M., Denev, I., Laloi, C., 2006. Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *Bioessays*. **28**, 1091-1101.
- Geigenberger, P., Thormählen, I., Daloso, D.M., Fernie, A.R., 2017. The unprecedented versatility of the plant thioredoxin system. *Trends in Plant Science*. **22**, 249-262.
- Geurts, P., 2018. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. *Scientific Reports*. **8**, 3384.
- Glémin, S. & Bataillon, T., 2009. A comparative view of the evolution of grasses under domestication. *New Phytologist*. **183**, 273-290.
- González-Calle, V., Iglesias-Fernández, R., Carbonero, P., Barrero-Sicilia, C., 2014. The *BdGAMYB* protein from *Brachypodium distachyon* interacts with *BdDOF24* and regulates transcription of the *BdCathB* gene upon seed germination. *Planta*. **240**, 539-552.
- Gonzalez, D.H., 2015. *Plant transcription factors: evolutionary, structural and functional aspects*. [e-book]. Academic Press.
- Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*. **36**, 3420-3435.
- Gramzow, L. & Theissen, G., 2010. A hitchhiker's guide to the MADS world of plants. *Genome Biology*. **11**, 214.
- Gramzow, L., Ritz, M.S., Theißen, G., 2010. On the origin of MADS-domain transcription factors. *Trends in Genetics*. **26**, 149-153.
- Gregis, V., Sessa, A., Dorca-Fornell, C., Kater, M.M., 2009. The *Arabidopsis* floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *The Plant Journal*. **60**, 626-637.
- Grennan, A.K., 2011. Metallothioneins, a diverse protein family. *Plant Physiology*. **155**, 1750-1751.
- Guillon, F., Larre, C., Petipas, F., Berger, A., Moussawi, J., Rogniaux, H., Santoni, A., Saulnier, L., Jamme, F., Miquel, M., 2011. A comprehensive overview of grain development in *Brachypodium distachyon* variety Bd21. *Journal of Experimental Botany*. **63**, 739-755.

References

- Guittou, A. & Berger, F., 2004. Control of reproduction by Polycomb Group complexes in animals and plants. *International Journal of Developmental Biology*. **49**, 707-716.
- Hands, P. & Drea, S., 2012. A comparative view of grain development in *Brachypodium distachyon*. *Journal of Cereal Science*. **56**, 2-8.
- Hands, P., Kourmpetli, S., Sharples, D., Harris, R.G., Drea, S., 2012. Analysis of grain characters in temperate grasses reveals distinctive patterns of endosperm organization associated with grain shape. *Journal of Experimental Botany*. **63**, 6253-6266.
- Hands, P., Rabiger, D.S., Koltunow, A., 2016. Mechanisms of endosperm initiation. *Plant Reproduction*. **29**, 215-225.
- Hatsugai, N., Yamada, K., Goto-Yamada, S., Hara-Nishimura, I., 2015. Vacuolar processing enzyme in plant programmed cell death. *Frontiers in Plant Science*. **6**, 234.
- He, Y., Wang, C., Higgins, J.D., Yu, J., Zong, J., Lu, P., Zhang, D., Liang, W., 2016. MEIOTIC F-BOX is essential for male meiotic DNA double-strand break repair in rice. *The Plant Cell*. **28**, 1879-1893.
- Heinze, K., 2017. From Grain to Granule: The Biomechanics of Wheat Grain Fractionation with a Focus on the Role of Starch Granules. Thesis. 10.13140/RG.2.2.29266.25284.
- Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Münster, T., Theißen, G., 2002. Two ancient classes of MIKC-type *MADS-box* genes are present in the moss *Physcomitrella patens*. *Molecular Biology and Evolution*. **19**, 801-814.
- Heslop-Harrison, J. & Mackenzie, A., 1967. Autoradiography of soluble [2-¹⁴C] thymidine derivatives during meiosis and microsporogenesis in *Lilium* anthers. *Journal of Cell Science*. **2**, 387-400.
- Hillman, G.C. & Davies, M.S., 1990. 6. Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society*. **39**, 39-78.
- Hirabayashi, C. & Murai, K., 2009. Class C *MADS-box* gene *AGAMOUS* was duplicated in the wheat genome. *Wheat Inf.Serv*. **107**, 13-16.
- Holtman, W.L., Heistek, J.C., Mattern, K.A., Bakhuizen, R., Douma, A.C., 1994. β -Oxidation of fatty acids is linked to the glyoxylate cycle in the aleurone but not in the embryo of germinating barley. *Plant Science*. **99**, 43-53.
- Hong, S., Seo, P.J., Yang, M., Xiang, F., Park, C., 2008. Exploring valid reference genes for gene expression studies in *Brachypodium distachyon* by real-time PCR. *BMC Plant Biology*. **8**, 112.
- Honma, T. & Goto, K., 2001. Complexes of *MADS-box* proteins are sufficient to convert leaves into floral organs. *Nature*. **409**, 525.

References

- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., Zimmermann, P., 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics*. **2008**, .
- Hsu, H., Hsieh, W., Chen, M., Chang, Y., Yang, C., 2010. C/D class MADS box genes from two monocots, orchid (*Oncidium Gower Ramsey*) and lily (*Lilium longiflorum*), exhibit different effects on floral transition and formation in *Arabidopsis thaliana*. *Plant and Cell Physiology*. **51**, 1029-1045.
- Hu, L., Liang, W., Yin, C., Cui, X., Zong, J., Wang, X., Hu, J., Zhang, D., 2011. Rice MADS3 regulates ROS homeostasis during late anther development. *The Plant Cell*. **23**, 515-533.
- Hu, Y., Liang, W., Yin, C., Yang, X., Ping, B., Li, A., Jia, R., Chen, M., Luo, Z., Cai, Q., 2015. Interactions of *OsMADS1* with floral homeotic genes in rice flower development. *Molecular Plant*. **8**, 1366-1384.
- Huang, K., Louis, J.M., Donaldson, L., Lim, F., Sharrocks, A.D., Clore, G.M., 2000. Solution structure of the MEF2A–DNA complex: structural basis for the modulation of DNA bending and specificity by MADS-box transcription factors. *The EMBO Journal*. **19**, 2615-2628.
- Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., Gornicki, P., 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences*. **99**, 8133-8138.
- Hubbard, C.E., 1954. Grasses. A guide to their structure, identification, uses, and distribution in the British Isles. *Grasses. A Guide to their Structure, Identification, Uses, and Distribution in the British Isles*.
- Iglesias-Fernández, R., Wozny, D., Iriondo-de Hond, M., Oñate-Sánchez, L., Carbonero, P., Barrero-Sicilia, C., 2014. The *AtCathB3* gene, encoding a cathepsin B-like protease, is expressed during germination of *Arabidopsis thaliana* and transcriptionally repressed by the basic leucine zipper protein GBF1. *Journal of Experimental Botany*. **65**, 2009-2021.
- Ingouff, M., Haseloff, J., Berger, F., 2005. Polycomb group genes control developmental timing of endosperm. *The Plant Journal*. **42**, 663-674.
- International Barley Genome Sequencing Consortium, 2012. A physical, genetic and functional sequence assembly of the barley genome. *Nature*. **491**, 711.
- International *Brachypodium* Initiative, 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*. **463**, 763.
- Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I., Okada, K., 2001. The DEFECTIVE IN ANTHIER DEHISCENCE1 gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. *The Plant Cell*. **13**, 2191-2209.

References

- Ito, T., Ng, K., Lim, T., Yu, H., Meyerowitz, E.M., 2007. The homeotic protein *AGAMOUS* controls late stamen development by regulating a jasmonate biosynthetic gene in *Arabidopsis*. *The Plant Cell*. **19**, 3516-3529.
- Ito, T., Wellmer, F., Yu, H., Das, P., Ito, N., Alves-Ferreira, M., Riechmann, J.L., Meyerowitz, E.M., 2004. The homeotic protein *AGAMOUS* controls microsporogenesis by regulation of *SPOROCTELESS*. *Nature*. **430**, 356.
- Irrthum A, Wehenkel L, Geurts P (2010) Inferring regulatory networks from expression data using tree-based methods 5:e12776
- Jalmi SK, Sinha AK (2015) ROS mediated MAPK signaling in abiotic and biotic stress-striking similarities and differences 6:769
- Jager, M., Hassanin, A., Manuel, M., Guyader, H.L., Deutsch, J., 2003. *MADS-box* genes in *Ginkgo biloba* and the evolution of the *AGAMOUS* family. *Molecular Biology and Evolution*. **20**, 842-854.
- Jeon, K.W., 1996. *International review of cytology*. [e-book]. Academic Press.
- Jian Ma, Yujie Yang, Wei Luo, Congcong Yang, Puyang Ding, Yaxi Liu, Linyi Qiao, Zhijian Chang, Hongwei Geng, Penghao Wang, Qiantao Jiang, Jirui Wang, Guoyue Chen, Yuming Wei, Youliang Zheng, Xiujin Lan, 2017. Genome-wide identification and analysis of the *MADS-box* gene family in bread wheat (*Triticum aestivum* L.). *PLoS One*. **12**, e0181443.
- Jones, R.L., 1972. Fractionation of the enzymes of the barley aleurone layer: evidence for a soluble mode of enzyme release. *Planta*. **103**, 95-109.
- Jones, R.L., 1969. The fine structure of barley aleurone cells. *Planta*. **85**, 359-375.
- Julián, I., Gandullo, J., Santos-Silva, L.K., Diaz, I., Martinez, M., 2013. Phylogenetically distant barley legumains have a role in both seed and vegetative tissues. *Journal of Experimental Botany*. **64**, 2929-2941.
- Kang, H., Jeon, J., Lee, S., An, G., 1998. Identification of class B and class C floral organ identity genes from rice plants. *Plant Molecular Biology*. **38**, 1021-1029.
- Kang, I., Steffen, J.G., Portereiko, M.F., Lloyd, A., Drews, G.N., 2008. The AGL62 *MADS* domain protein regulates cellularization during endosperm development in *Arabidopsis*. *The Plant Cell*. **20**, 635-647.
- Kanno, A., Hienuki, H., Ito, T., Nakamura, T., Fukuda, T., Yun, P., Song, I., Kamimura, T., Ochiai, T., Yokoyama, J., 2006. The structure and expression of *SEPALLATA*-like genes in *Asparagus* species (*Asparagaceae*). *Sexual Plant Reproduction*. **19**, 133-144.
- Kater, M.M., Dreni, L., Colombo, L., 2006. Functional conservation of *MADS-box* factors controlling floral organ identity in rice and *Arabidopsis*. *Journal of Experimental Botany*. **57**, 3433-3444.

References

- Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*. **30**, 3059-3066.
- Katoh, K. & Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*. **30**, 772-780.
- Kaufmann, K., Melzer, R., Theißen, G., 2005. MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene*. **347**, 183-198.
- Keijzer, C.J., 1987. THE PROCESSES OF ANTHOR DEHISCENCE AND POLLEN DISPERSAL: II. THE FORMATION AND THE TRANSFER MECHANISM OF POLLENKITT, CELL-WALL DEVELOPMENT OF THE LOCULUS TISSUES AND A FUNCTION OF ORBICULES IN POLLEN DISPERSAL. *New Phytologist*. **105**, 499-507.
- Kellogg, E.A., 2001. Evolutionary history of the grasses. *Plant Physiology*. **125**, 1198-1205.
- Kellogg, E.A., 2015. Flowering plants. Monocots: Poaceae. [e-book]. Springer.
- Kellogg, E.A., Camara, Paulo E A S, Rudall, P.J., Ladd, P., Malcomber, S.T., Whipple, C.J., Doust, A.N., 2013. Early inflorescence development in the grasses (Poaceae). *Frontiers in Plant Science*. **4**, 250.
- Kelly, S.A., Summerbell, C.D., Brynes, A., Whittaker, V., Frost, G., 2007. Wholegrain cereals for coronary heart disease. *Cochrane Database of Systematic Reviews*.
- Khan, M.A. & Stace, C.A., 1999. Breeding relationships in the genus *Brachypodium* (Poaceae: Pooideae). *Nordic Journal of Botany*. **19**, 257-269.
- Khanday, I., Yadav, S.R., Vijayraghavan, U., 2013. Rice LHS1/OsMADS1 controls floret meristem specification by coordinated regulation of transcription factors and hormone signaling pathways. *Plant Physiology*. **161**, 1970-1983.
- Kladnik, A., Chamusco, K., Dermastia, M., Chourey, P., 2004. Evidence of programmed cell death in post-phloem transport cells of the maternal pedicel tissue in developing caryopsis of maize. *Plant Physiology*. **136**, 3572-3581.
- Kobayashi, K., Yasuno, N., Sato, Y., Yoda, M., Yamazaki, R., Kimizu, M., Yoshida, H., Nagamura, Y., Kyojuka, J., 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS box genes and PAP2, a SEPALLATA MADS box gene. *The Plant Cell*. **24**, 1848-1859.
- Kourmpetli, S. & Drea, S., 2014. The fruit, the whole fruit, and everything about the fruit. *Journal of Experimental Botany*. **65**, 4491-4503.
- Kramer, E.M., Jaramillo, M.A., Di Stilio, V.S., 2004. Patterns of Gene Duplication and Functional Evolution During the Diversification of the *AGAMOUS* Subfamily of MADS Box Genes in Angiosperms. *Genetics*. **166**, 1011-1023.

References

- Kumar, S., Soukup, M., Elbaum, R., 2017. Silicification in grasses: variation between different cell types. *Frontiers in Plant Science*. **8**, 438.
- Kyozuka, J., Kobayashi, T., Morita, M., Shimamoto, K., 2000. Spatially and temporally regulated expression of rice MADS box genes with similarity to *Arabidopsis* class A, B and C genes. *Plant and Cell Physiology*. **41**, 710-718.
- Lai, C., Huang, L., Chen, L.O., Chan, M., Shaw, J., 2017. Genome-wide analysis of GDSL-type esterases/lipases in *Arabidopsis*. *Plant Molecular Biology*. **95**, 181-197.
- Lachman, J., Hejtmánková, K., Kotíková, Z., 2013. Tocols and carotenoids of einkorn, emmer and spring wheat varieties: Selection for breeding and production. *Journal of Cereal Science*. **57**, 207-214.
- Lee, S., Jung, K., An, G., Chung, Y., 2004. Isolation and characterization of a rice cysteine protease gene, *OsCP1*, using T-DNA gene-trap system. *Plant Molecular Biology*. **54**, 755-765.
- Lee, S., Jeon, J., An, K., Moon, Y., Lee, S., Chung, Y., An, G., 2003. Alteration of floral organ identity in rice through ectopic expression of *OsMADS16*. *Planta*. **217**, 904-911.
- Lenhard, M., Bohnert, A., Jürgens, G., Laux, T., 2001. Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between WUSCHEL and AGAMOUS. *Cell*. **105**, 805-814.
- Leroux, B.M., Goodyke, A.J., Schumacher, K.I., Abbott, C.P., Clore, A.M., Yadegari, R., Larkins, B.A., Dannenhoffer, J.M., 2014. Maize early endosperm growth and development: from fertilization through cell type differentiation. *American Journal of Botany*. **101**, 1259-1274.
- Levine, M. & Tjian, R., 2003. Transcription regulation and animal diversity. *Nature*. **424**, 147-151.
- Li, H., Liang, W., Hu, Y., Zhu, L., Yin, C., Xu, J., Dreni, L., Kater, M.M., Zhang, D., 2011. Rice MADS6 interacts with the floral homeotic genes SUPERWOMAN1, MADS3, MADS58, MADS13, and DROOPING LEAF in specifying floral organ identities and meristem fate. *The Plant Cell*. **23**, 2536-2552.
- Li, H., Liang, W., Yin, C., Zhu, L., Zhang, D., 2011. Genetic interaction of *OsMADS3*, DROOPING LEAF, and *OsMADS13* in specifying rice floral organ identities and meristem determinacy. *Plant Physiology*. **156**, 263-274.
- Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. **25**, 1754-1760.
- Li, L., Li, Y., Song, S., Deng, H., Li, N., Fu, X., Chen, G., Yuan, L., 2015. An anther development F-box (ADF) protein regulated by tapetum degeneration retardation (TDR) controls rice anther development. *Planta*. **241**, 157-166.

References

- Liao, Y., Smyth, G.K., Shi, W., 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*. **47**, e47.
- Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., Yanofsky, M.F., 2000. SHATTERPROOF *MADS*-box genes control seed dispersal in *Arabidopsis*. *Nature*. **404**, 766.
- Linnestad, C., Doan, D.N., Brown, R.C., Lemmon, B.E., Meyer, D.J., Jung, R., Olsen, O., 1998. Nucellain, a barley homolog of the dicot vacuolar-processing protease, is localized in nucellar cell walls. *Plant Physiology*. **118**, 1169-1180.
- Liu, C., Chen, H., Er, H.L., Soo, H.M., Kumar, P.P., Han, J., Liou, Y.C., Yu, H., 2008. Direct interaction of AGL24 and SOC1 integrates flowering signals in *Arabidopsis*. *Development*. **135**, 1481-1491.
- Liu, C., Wang, J., Mei, X., Deng, X., Yu, T., Liu, X., Wang, G., Liu, Z., Cai, Y., 2015. Characterization of the imprinting and expression patterns of ZAG2 in maize endosperm and embryo. *The Crop Journal*. **3**, 74-79.
- Liu, X., 2011. *AGAMOUS* terminates floral stem cell maintenance in *Arabidopsis* by directly repressing WUSCHEL through recruitment of Polycomb Group proteins. *The Plant Cell*. **23**, 3654-3670.
- Livak, K.J. & Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods*. **25**, 402-408.
- Lohmann, J.U., Hong, R.L., Hobe, M., Busch, M.A., Parcy, F., Simon, R., Weigel, D., 2001. A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. *Cell*. **105**, 793-803.
- Lopez-Dee, Z.P., Wittich, P., Enrico Pe, M., Rigola, D., Del Buono, I., Gorla, M.S., Kater, M.M., Colombo, L., 1999. *OsMADS13*, a novel rice *MADS*-box gene expressed during ovule development. *Developmental Genetics*. **25**, 237-244.
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. **15**, 550.
- Lu, W., Deng, M., Guo, F., Wang, M., Zeng, Z., Han, N., Yang, Y., Zhu, M., Bian, H., 2016. Suppression of *OsVPE3* enhances salt tolerance by attenuating vacuole rupture during programmed cell death and affects stomata development in rice. *Rice*. **9**, 65.
- Ludovico Dreni, Alessandro Pilatone, Dapeng Yun, Stefano Erreni, Alice Pajoro, Elisabetta Caporali, Dabing Zhang, Martin M. Kater, 2011. Functional Analysis of All *AGAMOUS* Subfamily Members in Rice Reveals Their Roles in Reproductive Organ Identity Determination and Meristem Determinacy. *The Plant Cell*. **23**, 2850-2863.

References

- Martinez, M. & Diaz, I., 2008. The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. *BMC Evolutionary Biology*. **8**, 198.
- Meguro, A., Takumi, S., Ogiwara, Y., Murai, K., 2003. WAG, a wheat *AGAMOUS* homolog, is associated with development of pistil-like stamens in alloplasmic wheats. *Sexual Plant Reproduction*. **15**, 221-230.
- Mehrian-Shai, R., Yalon, M., Simon, A.J., Eyal, E., Pismenyuk, T., Moshe, I., Constantini, S., Toren, A., 2015. High metallothionein predicts poor survival in glioblastoma multiforme. *BMC Medical Genomics*. **8**, 68.
- Messenguy, F. & Dubois, E., 2003. Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene*. **316**, 1-21.
- Miller, G. & Mittler, R., 2006. Could heat shock transcription factors function as hydrogen peroxide sensors in plants? *Annals of Botany*. **98**, 279-288.
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H., El-Gebali, S., Fraser, M.I., 2018. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research*. **47**, D351-D360.
- Mittler R (2017) ROS are good. *Trends Plant Sci* 22:11-19
- Mittler R, Vanderauwera S, Gollery M, Van Breusegem F (2004) Reactive oxygen gene network of plants. *Trends Plant Sci* 9:490-498
- Mizukami, Y. & Ma, H., 1995. Separation of AG function in floral meristem determinacy from that in reproductive organ identity by expressing antisense AG RNA. *Plant Molecular Biology*. **28**, 767-784.
- Mizukami, Y., Huang, H., Tudor, M., Hu, Y., Ma, H., 1996. Functional domains of the floral regulator *AGAMOUS*: characterization of the DNA binding domain and analysis of dominant negative mutations. *The Plant Cell*. **8**, 831-845.
- Mizzotti, C., Ezquer, I., Paolo, D., Rueda-Romero, P., Guerra, R.F., Battaglia, R., Rogachev, I., Aharoni, A., Kater, M.M., Caporali, E., 2014. SEEDSTICK is a master regulator of development and metabolism in the *Arabidopsis* seed coat. *PLoS Genetics*. **10**, e1004856.
- Morrison, I.N., O'brien, T.P., Kuo, J., 1978. Initial cellularization and differentiation of the aleurone cells in the ventral region of the developing wheat grain. *Planta*. **140**, 19-30.
- Mur, L.A., Allainguillaume, J., Catalán, P., Hasterok, R., Jenkins, G., Lesniewska, K., Thomas, I., Vogel, J., 2011. Exploiting the *Brachypodium* Tool Box in cereal and grass research. *New Phytologist*. **191**, 334-347.

References

- Murai, K., 2013. Homeotic genes and the ABCDE model for floral organ formation in wheat. *Plants*. **2**, 379-395.
- Nagasawa, N., Miyoshi, M., Sano, Y., Satoh, H., Hirano, H., Sakai, H., Nagato, Y., 2003. SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice. *Development*. **130**, 705-718.
- Nawaschin, S., 1898. Resultate einer Revision der Befruchtungsvorgänge bei *Lilium martagon* und *Fritillaria tenella*. *Известия Российской Академии Наук. Серия Математическая*. **9**, 377-382.
- Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M., Lepiniec, L., 2002. The TRANSPARENT TESTA16 locus encodes the *ARABIDOPSIS* BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *The Plant Cell*. **14**, 2463-2479.
- Nevo, E., Fu, Y., Pavlicek, T., Khalifa, S., Tavasi, M., Beiles, A., 2012. Evolution of wild cereals during 28 years of global warming in Israel. *Proceedings of the National Academy of Sciences*. **109**, 3412-3415.
- Nguyen, N.H. & Lee, H., 2016. MYB-related transcription factors function as regulators of the circadian clock and anthocyanin biosynthesis in *Arabidopsis*. *Plant Signaling & Behavior*. **11**, e1139278.
- Norstog, K., 1974. Nucellus during early embryogeny in barley: fine structure. *Botanical Gazette*. **135**, 97-103.
- Ogbe, R.J., Ochalefu, D.O., Mafulul, S.G., Olaniru, O.B., 2015. A review on dietary phytosterols: Their occurrence, metabolism and health benefits. *Asian J.Plant Sci.Res.* **5**, 10-21.
- Okada, T., Jayasinghe, J.R.M., Nansamba, M., Baes, M., Warner, P., Kouidri, A., Correia, D., Nguyen, V., Whitford, R., Baumann, U., 2017. Unfertilized ovary pushes wheat flower open for cross-pollination. *Journal of Experimental Botany*. **69**, 399-412.
- Olsen, O., 2004. Nuclear endosperm development in cereals and *Arabidopsis thaliana*. *The Plant Cell*. **16**, S214-S227.
- Olson, E.N., Perry, M., Schulz, R.A., 1995. Regulation of muscle differentiation by the MEF2 family of MADS box transcription factors. *Developmental Biology*. **172**, 2-14.
- Opanowicz Magdalena, Philip Hands, Donna Betts, Mary L. Parker, Geraldine A. Toole, E. N. Clare Mills, John H. Doonan, Sinéad Drea, 2011. Endosperm development in *Brachypodium distachyon*. *Journal of Experimental Botany*. **62**, 735-748.
- Opanowicz, M., Vain, P., Draper, J., Parker, D., Doonan, J.H., 2008. *Brachypodium distachyon*: making hay with a wild grass. *Trends in Plant Science*. **13**, 172-177.

References

- Oparka, K.J. & Gates, P., 1981. Transport of assimilates in the developing caryopsis of rice (*Oryza sativa* L.). *Planta*. **151**, 561-573.
- Ozdemir, B.S., Hernandez, P., Filiz, E., Budak, H., 2008. *Brachypodium* genomics. *International Journal of Plant Genomics*. **2008**, .
- Ozyigit, I.I., Filiz, E., Vatansever, R., Kurtoglu, K.Y., Koc, I., Öztürk, M.X., Anjum, N.A., 2016. Identification and comparative analysis of H₂O₂-scavenging enzymes (ascorbate peroxidase and glutathione peroxidase) in selected plants employing bioinformatics approaches. *Frontiers in Plant Science*. **7**, 301.
- Paolacci, A.R., Tanzarella, O.A., Porceddu, E., Varotto, S., Ciaffi, M., 2007. Molecular and phylogenetic analysis of *MADS-box* genes of MIKC-type and chromosome location of SEP-like genes in wheat (*Triticum aestivum* L.). *Molecular Genetics and Genomics*. **278**, 689-708.
- Papini, A., Mosti, S., Brighigna, L., 1999. Programmed-cell-death events during tapetum development of angiosperms. *Protoplasma*. **207**, 213-221.
- Pařenicová, L., De Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., 2003. Molecular and phylogenetic analyses of the complete *MADS-box* transcription factor family in *Arabidopsis*: new openings to the MADS world. *The Plant Cell*. **15**, 1538-1551.
- Paquet, N., Bernadet, M., Morin, H., Traas, J., Dron, M., Charon, C., 2005. Expression patterns of TEL genes in Poaceae suggest a conserved association with cell differentiation. *Journal of Experimental Botany*. **56**, 1605-1614.
- Park, S.K., Howden, R., Twell, D., 1998. The *Arabidopsis thaliana* gametophytic mutation gemini pollen1 disrupts microspore polarity, division asymmetry and pollen cell fate. *Development*. **125**, 3789-3799.
- Pařenicová, L., De Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., 2003. Molecular and phylogenetic analyses of the complete *MADS-box* transcription factor family in *Arabidopsis*: new openings to the MADS world. *The Plant Cell*. **15**, 1538-1551.
- Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*. **457**, 551-556.
- Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E., Yanofsky, M.F., 2000. B and C floral organ identity functions require SEPALLATA *MADS-box* genes. *Nature*. **405**, 200.
- Perez-Prat, E. & van Lookeren Campagne, Michiel M, 2002. Hybrid seed production and the challenge of propagating male-sterile plants. *Trends in Plant Science*. **7**, 199-203.
- Percival, J., 1921. The wheat plant; a monograph, by John Percival...

References

- Peterson, R., Slovin, J.P., Chen, C., 2010. A simplified method for differential staining of aborted and non-aborted pollen grains. *International Journal of Plant Biology*. **1**, e13.
- Piergiovanni, A.R., Rizzi, R., Pannacciulli, E., Gatta, C.D., 1997. Mineral composition in hulled wheat grains: a comparison between emmer (*Triticum dicoccon* Schrank) and spelt (*T. spelta* L.) accessions. *International Journal of Food Sciences and Nutrition*. **48**, 381-386.
- Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E., Yanofsky, M.F., 2003. Assessing the redundancy of *MADS-box* genes during carpel and ovule development. *Nature*. **424**, 85.
- Pireyre, M. & Burow, M., 2015. Regulation of MYB and bHLH transcription factors: a glance at the protein level. *Molecular Plant*. **8**, 378-388.
- Polowick, P.L., 1992. Ultrastructural changes in the cell wall, nucleus and cytoplasm of pollen mother cells during meiotic prophase I in *Lycopersicon esculentum* (Mill). *Springer*.
- Pont, C., Leroy, T., Seidel, M., Tondelli, A., Duchemin, W., Armisen, D., Lang, D., Bustos-Korts, D., Goué, N., Balfourier, F., 2019. Tracing the ancestry of modern bread wheats. *Nature Genetics*. **51**, 905.
- Prasad, K., Zhang, X., Tobón, E., Ambrose, B.A., 2010. The *Arabidopsis* B-sister *MADS-box* protein, GORDITA, represses fruit growth and contributes to integument development. *The Plant Journal*. **62**, 203-214.
- Qian-Hao, Z., Ramm, K., Shivakkumar, R., Dennis, E.S., Upadhyaya, N.M., 2004. The ANTHIER INDEHISCENCE1 Gene Encoding a Single MYB Domain Protein Is Involved in Anther Development in Rice1. *Plant Physiology*. **135**, 1514.
- Qu, G., Quan, S., Mondol, P., Xu, J., Zhang, D., Shi, J., 2014. Comparative metabolomic analysis of wild type and *mads3* mutant rice anthers. *Journal of Integrative Plant Biology*. **56**, 849-863.
- Radchuk, V., Tran, V., Radchuk, R., Diaz-Mendoza, M., Weier, D., Fuchs, J., Riewe, D., Hensel, G., Kumlehn, J., Munz, E., 2018. Vacuolar processing enzyme 4 contributes to maternal control of grain size in barley by executing programmed cell death in the pericarp. *New Phytologist*. **218**, 1127-1142.
- Radchuk, V., Weier, D., Radchuk, R., Weschke, W., Weber, H., 2010. Development of maternal seed tissue in barley is mediated by regulated cell expansion and cell disintegration and coordinated with endosperm growth. *Journal of Experimental Botany*. **62**, 1217-1227.
- Raghavan, V., 2006. *Double fertilization: embryo and endosperm development in flowering plants*. [e-book]. Springer Science & Business Media.

References

- Ranjan, R., Khurana, R., Malik, N., Badoni, S., Parida, S.K., Kapoor, S., Tyagi, A.K., 2017. bHLH142 regulates various metabolic pathway-related genes to affect pollen development and anther dehiscence in rice. *Scientific Reports*. **7**, 43397.
- Ramírez-González, R.H., Borrill, P., Lang, D., Harrington, S.A., Brinton, J., Venturini, L., Davey, M., Jacobs, J., Van Ex, F., Pasha, A., 2018. The transcriptional landscape of polyploid wheat. *Science*. 361, eaar6089.
- Rawlings, N.D., Barrett, A.J., Bateman, A., 2011. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*. **40**, D343-D350.
- RCore, T., 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL [Http://Www.R-Project.Org](http://www.R-project.org).
- Riechmann, J.L. & Meyerowitz, E.M., 1997. MADS domain proteins in plant development. *Biological Chemistry*. **378**, 1079-1102.
- Riechmann, J.L., Wang, M., Meyerowitz, E.M., 1996. DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. *Nucleic Acids Research*. **24**, 3134-3141.
- Rijkema, A.S., Gerats, T., Vandenbussche, M., 2007. Evolutionary complexity of MADS complexes. *Current Opinion in Plant Biology*. **10**, 32-38.
- Rijkema, A.S., *et al*, 2010, Variations on a theme: changes in the floral ABCs in angiosperms, *Seminars in cell & developmental biology*, 2010, Elsevier pp100-107.
- Rocha, A.J., Soares, E.L., Costa, J.H., Costa, W.L., Soares, A.A., Nogueira, F.C., Domont, G.B., Campos, F.A., 2013. Differential expression of cysteine peptidase genes in the inner integument and endosperm of developing seeds of *Jatropha curcas* L.(Euphorbiaceae). *Plant Science*. **213**, 30-37.
- Rogers, J.C., Dean, D., Heck, G.R., 1985. Aleurain: a barley thiol protease closely related to mammalian cathepsin H. *Proceedings of the National Academy of Sciences*. **82**, 6512-6516.
- Roosevelt, T., 2013. *Autobiography of Theodore Roosevelt*. [e-book]. Simon and Schuster.
- Rounsley, S.D., Ditta, G.S., Yanofsky, M.F., 1995. Diverse roles for MADS box genes in *Arabidopsis* development. *The Plant Cell*. **7**, 1259-1269.
- Sage, T.L., Bagha, S., Lundsgaard-Nielsen, V., Branch, H.A., Sultmanis, S., Sage, R.F., 2015. The effect of high temperature stress on male and female reproduction in plants. *Field Crops Research*. **182**, 30-42.
- Salvi, S., Porfiri, O., Ceccarelli, S., 2013. Nazareno Strampelli, the 'Prophet' of the green revolution. *The Journal of Agricultural Science*. **151**, 1-5.

References

- Schilling, S., Gramzow, L., Lobbes, D., Kirbis, A., Weilandt, L., Hoffmeier, A., Junker, A., Weigelt-Fischer, K., Klukas, C., Wu, F., 2015. Non-canonical structure, function and phylogeny of the B sister MADS-box gene *O s MADS 30* of rice (*O ryza sativa*). *The Plant Journal*. **84**, 1059-1072.
- Schippmann, U., 1991. *Revision der europäischen Arten der Gattung Brachypodium Palisot de Beauvois (Poaceae)*. [e-book]. Geneve.
- Schmid, M., Simpson, D., Gietl, C., 1999. Programmed cell death in castor bean endosperm is associated with the accumulation and release of a cysteine endopeptidase from ricinosomes. *Proceedings of the National Academy of Sciences*. **96**, 14159-14164.
- Schmidt, R.J., Veit, B., Mandel, M.A., Mena, M., Hake, S., Yanofsky, M.F., 1993. Identification and molecular characterization of ZAG1, the maize homolog of the *Arabidopsis* floral homeotic gene *AGAMOUS*. *The Plant Cell*. **5**, 729-737.
- Schmitz, J., Franzen, R., Ngyuen, T.H., Garcia-Maroto, F., Pozzi, C., Salamini, F., Rohde, W., 2000. Cloning, mapping and expression analysis of barley *MADS-box* genes. *Plant Molecular Biology*. **42**, 899-913.
- Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A., 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science*. **326**, 1112-1115.
- Schneider, J., Döring, E., Hilu, K.W., Röser, M., 2009. Phylogenetic structure of the grass subfamily Pooideae based on comparison of plastid *matK* gene–3' *trnK* exon and nuclear ITS sequences. *Taxon*. **58**, 405-424.
- Scholthof, K.G., Irigoyen, S., Catalan, P., Mandadi, K.K., 2018. *Brachypodium*: a monocot grass model genus for plant biology. *The Plant Cell*. **30**, 1673-1694.
- Schreiber, D.N., Bantin, J., Dresselhaus, T., 2004. The MADS box transcription factor ZmMADS2 is required for anther and pollen maturation in maize and accumulates in apoptotic bodies during anther dehiscence. *Plant Physiology*. **134**, 1069-1079.
- Schubert, M., Marcussen, T., Meseguer, A.S., Fjellheim, S., 2019. The grass subfamily Pooideae: Cretaceous–Palaeocene origin and climate-driven Cenozoic diversification. *Global Ecology and Biogeography*. **28**, 1168-1182.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., Cavalli, G., 2007. Genome regulation by polycomb and trithorax proteins. *Cell*. **128**, 735-745.
- Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., Sommer, H., 1990. Genetic control of flower development by homeotic genes in *Antirrhinum majus*. *Science*. **250**, 931.
- Sharma, A., Singh, M.B., Bhalla, P.L., 2015. Anther ontogeny in *Brachypodium distachyon*. *Protoplasma*. **252**, 439-450.

References

- Sherf, B.A., Navarro, S.L., Hannah, R.R., Wood, K.V., 1996. Dual-luciferase reporter assay: an advanced co-reporter technology integrating firefly and Renilla luciferase assays. *Promega Notes*. **57**, .
- Shewry, P.R., 2009. Wheat. *Journal of Experimental Botany*. **60**, 1537-1553.
- Shewry, P.R. & Hey, S., 2015a. Do “ancient” wheat species differ from modern bread wheat in their contents of bioactive components? *Journal of Cereal Science*. **65**, 236-243.
- Shewry, P.R. & Hey, S.J., 2015b. The contribution of wheat to human diet and health. *Food and Energy Security*. **4**, 178-202.
- Shitsukawa, N., Ikari, C., Shimada, S., Kitagawa, S., Sakamoto, K., Saito, H., Ryuto, H., Fukunishi, N., Abe, T., Takumi, S., 2007. The einkorn wheat (*Triticum monococcum*) mutant, maintained vegetative phase, is caused by a deletion in the VRN1 gene. *Genes & Genetic Systems*. **82**, 167-170.
- Smaczniak, C., Immink, R.G., Angenent, G.C., Kaufmann, K., 2012. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development*. **139**, 3081-3098.
- Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., Arnaiz, O., Awedh, M.H., Baldock, R., Barbiera, G., 2015. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Research*. **43**, W589-W598.
- Sorensen, A., Kröber, S., Unte, U.S., Huijser, P., Dekker, K., Saedler, H., 2003. The *Arabidopsis* ABORTED MICROSPORES (AMS) gene encodes a MYC class transcription factor. *The Plant Journal*. **33**, 413-423.
- Sparkes, I.A., Runions, J., Kearns, A., Hawes, C., 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols*. **1**, 2019.
- Sreenivasulu, N., Radchuk, V., Strickert, M., Miersch, O., Weschke, W., Wobus, U., 2006. Gene expression patterns reveal tissue-specific signaling networks controlling programmed cell death and ABA-regulated maturation in developing barley seeds. *The Plant Journal*. **47**, 310-327.
- Stegmaier, P., Kel, A.E., Wingender, E., 2004. Systematic DNA-binding domain classification of transcription factors. *Genome Informatics*. **15**, 276-286.
- Stieglitz, H., 1977. Role of β -1, 3-glucanase in postmeiotic microspore release. *Developmental Biology*. **57**, 87-97.
- Sullivan, P., Arendt, E., Gallagher, E., 2013. The increasing use of barley and barley by-products in the production of healthier baked goods. *Trends in Food Science & Technology*. **29**, 124-134.

References

- Sun, B. & Ito, T., 2010. No title. *Floral Stem Cells: From Dynamic Balance Towards Termination*.
- Tanno, K. & Willcox, G., 2006. How fast was wild wheat domesticated? *Science*. **311**, 1886.
- Terrell, E.E. & Wergin, W.P., 1981. Epidermal features and silica deposition in lemmas and awns of *Zizania* (Gramineae). *American Journal of Botany*. **68**, 697-707.
- Theißen, G., 2001. Development of floral organ identity: stories from the MADS house. *Current Opinion in Plant Biology*. **4**, 75-85.
- Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Münster, T., Winter, K., Saedler, H., 2000. A short history of *MADS-box* genes in plants. *Plant Molecular Biology*. **42**, 115-149.
- Theißen, G., Kim, J.T., Saedler, H., 1996. Classification and phylogeny of the *MADS-box* multigene family suggest defined roles of *MADS-box* gene subfamilies in the morphological evolution of eukaryotes. *Journal of Molecular Evolution*. **43**, 484-516.
- Theißen, G., Melzer, R., Rümpler, F., 2016. MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development*. **143**, 3259-3271.
- Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., Narechania, A., 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research*. **13**, 2129-2141.
- Trafford, K., Haleux, P., Henderson, M., Parker, M., Shirley, N.J., Tucker, M.R., Fincher, G.B., Burton, R.A., 2013. Grain development in *Brachypodium* and other grasses: possible interactions between cell expansion, starch deposition, and cell-wall synthesis. *Journal of Experimental Botany*. **64**, 5033-5047.
- Tran, V., Weier, D., Radchuk, R., Thiel, J., Radchuk, V., 2014. Caspase-like activities accompany programmed cell death events in developing barley grains. *PloS One*. **9**, e109426.
- Tsanko S Gechev, *et al*, 2006. Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. United States: John Wiley & Sons, Inc.
- Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., Bozhkov, P.V., 2011. Metacaspases. *Cell Death and Differentiation*. **18**, 1279.
- Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A.N., Mittler, R., Shintani, D., 2009. Thiamin confers enhanced tolerance to oxidative stress in *Arabidopsis*. *Plant Physiology*. **151**, 421-432.
- Ullrich, S.E., 2010. *Barley: Production, improvement, and uses*. [e-book]. John Wiley & Sons.

References

- Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., Luscombe, N.M., 2009. A census of human transcription factors: function, expression and evolution. *Nature Reviews Genetics*. **10**, 252-263.
- Vercammen, D., Van De Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenbergh, I., Van Beeumen, J., Inzé, D., Van Breusegem, F., 2004. Type II metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* cleave substrates after arginine and lysine. *Journal of Biological Chemistry*. **279**, 45329-45336.
- Vicentini, A., Barber, J.C., Aliscioni, S.S., Giussani, L.M., Kellogg, E.A., 2008. The age of the grasses and clusters of origins of C4 photosynthesis. *Global Change Biology*. **14**, 2963-2977.
- Vogel, J. & Hill, T., 2008. High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. *Plant Cell Reports*. **27**, 471-478.
- Vogel, J.P., Gu, Y.Q., Twigg, P., Lazo, G.R., Laudencia-Chingcuanco, D., Hayden, D.M., Donze, T.J., Vivian, L.A., Stamova, B., Coleman-Derr, D., 2006. EST sequencing and phylogenetic analysis of the model grass *Brachypodium distachyon*. *Theoretical and Applied Genetics*. **113**, 186-195.
- Vogel, J. & Bragg, J., 2009. *Brachypodium distachyon*, a new model for the Triticeae. *Genetics and Genomics of the Triticeae*. Springer. 427-449.
- Vogel, K.P., Johnson, V.A., Mattern, P.J., 1976. Protein and Lysine Content of Grain, Endosperm, and Bran of Wheats from the USDA World Wheat Collection 1. *Crop Science*. **16**, 655-660.
- Vogel, J.P., Garvin, D.F., Mockler, T.C., Schmutz, J., Rokhsar, D., Bevan, M.W., Barry, K., Lucas, S., Harmon-Smith, M., Lail, K., 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*. **463**, 763-768.
- Wahl, V., Brand, L.H., Guo, Y., Schmid, M., 2010. The FANTASTIC FOUR proteins influence shoot meristem size in *Arabidopsis thaliana*. *BMC Plant Biology*. **10**, 285.
- Wang, D., Qin, B., Li, X., Tang, D., Zhang, Y., Cheng, Z., Xue, Y., 2016. Nucleolar DEAD-box RNA helicase TOGR1 regulates thermotolerant growth as a pre-rRNA chaperone in rice. *PLoS Genetics*. **12**, e1005844.
- Wang, H.L., Offler, C.E., Patrick, J.W., 1994. Nucellar projection transfer cells in the developing wheat grain. *Protoplasma*. **182**, 39-52.
- Wang, H.L., Offler, C.E., Patrick, J.W., Ugalde, T.D., 1994. The cellular pathway of photosynthate transfer in the developing wheat grain. I. Delineation of a potential transfer pathway using fluorescent dyes. *Plant, Cell & Environment*. **17**, 257-266.
- Wang, Q.H., Yang, Z.J., Wei, S.H., Jiang, Z.Y., Yang, Y.F., Hu, Z.S., Sun, Q.X., Peng, Z.S., 2015. Molecular cloning, characterization and expression analysis of WAG-1 in the pistillody line of common wheat. *Genetics and Molecular Research*. **14**, 12455-12465.

References

- Wang, W., Zhou, X., Xiong, H., Mao, W., Zhao, P., Sun, M., 2018. Papain-like and legumain-like proteases in rice: genome-wide identification, comprehensive gene feature characterization and expression analysis. *BMC Plant Biology*. **18**, 87.
- Wang, X., Tang, C., Zhang, H., Xu, J., Liu, B., Lv, J., Han, D., Huang, L., Kang, Z., 2011. TaDAD2, a negative regulator of programmed cell death, is important for the interaction between wheat and the stripe rust fungus. *Molecular Plant-Microbe Interactions*. **24**, 79-90.
- Wang, Y., Zhu, S., Liu, S., Jiang, L., Chen, L., Ren, Y., Han, X., Liu, F., Ji, S., Liu, X., 2009. The vacuolar processing enzyme OsVPE1 is required for efficient glutelin processing in rice. *The Plant Journal*. **58**, 606-617.
- Watanabe, H., Abe, K., Emori, Y., Hosoyama, H., Arai, S., 1991. Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (oryzains). *Journal of Biological Chemistry*. **266**, 16897-16902.
- Watanabe, N. & Lam, E., 2011. *Arabidopsis* metacaspase 2d is a positive mediator of cell death induced during biotic and abiotic stresses. *The Plant Journal*. **66**, 969-982.
- Waterkeyn, L., 1962. *Les paroies microsporocytaires de nature callosique chez Helleborus et Tradescantia*. [e-book]. Libraire C. Uystpruyst.
- Wei, B., Liu, D., Guo, J., Leseberg, C.H., Zhang, X., Mao, L., 2013. Functional divergence of two duplicated D-lineage *MADS-box* genes *BdMADS2* and *BdMADS4* from *Brachypodium distachyon*. *Journal of Plant Physiology*. **170**, 424-431.
- Wei, B., Zhang, R., Guo, J., Liu, D., Li, A., Fan, R., Mao, L., Zhang, X., 2014. Genome-wide analysis of the *MADS-box* gene family in *Brachypodium distachyon*. *PloS One*. **9**, e84781.
- Wei, S., Peng, Z., Zhou, Y., Yang, Z., Wu, K., Ouyang, Z., 2011. Nucleotide diversity and molecular evolution of the WAG-2 gene in common wheat (*Triticum aestivum* L) and its relatives. *Genetics and Molecular Biology*. **34**, 606-615.
- Wilson, Z.A., Morroll, S.M., Dawson, J., Swarup, R., Tighe, P.J., 2001. The *Arabidopsis* MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. *The Plant Journal*. **28**, 27-39.
- Wu, D., Liang, W., Zhu, W., Chen, M., Ferrandiz, C., Burton, R.A., Dreni, L., Zhang, D., 2018. Loss of LOFSEP transcription factor function converts spikelet to leaf-like structures in rice. *Plant Physiology*. **176**, 1646-1664.
- Wu, F., Shi, X., Lin, X., Liu, Y., Chong, K., Theißen, G., Meng, Z., 2017. The ABC s of flower development: mutational analysis of AP 1/FUL-like genes in rice provides evidence for a homeotic (A)-function in grasses. *The Plant Journal*. **89**, 310-324.
- Wu, H. & Cheung, A.Y., 2000. Programmed cell death in plant reproduction. *Programmed Cell Death in Higher Plants*. Springer. 23-37.

References

- Wu, X., Liu, J., Li, D., Liu, C., 2016. Rice caryopsis development I: Dynamic changes in different cell layers. *Journal of Integrative Plant Biology*. **58**, 772-785.
- Xi, Z., Liu, L., Rest, J.S., Davis, C.C., 2014. Coalescent versus concatenation methods and the placement of Amborella as sister to water lilies. *Systematic Biology*. **63**, 919-932.
- Xiao, H., Tang, J., Li, Y., Wang, W., Li, X., Jin, L., Xie, R., Luo, H., Zhao, X., Meng, Z., 2009. STAMENLESS 1, encoding a single C2H2 zinc finger protein, regulates floral organ identity in rice. *The Plant Journal*. **59**, 789-801.
- Xiong, F., Yu, X.R., Zhou, L., Wang, F., Xiong, A.S., 2013. Structural and physiological characterization during wheat pericarp development. *Plant Cell Reports*. **32**, 1309-1320.
- Xing, S., Salinas, M., Huijser, P., 2011. New players unveiled in early anther development. *Plant Signaling & Behavior*. **6**, 934-938.
- Xu, G. & Kong, H., 2007. Duplication and divergence of floral MADS-box genes in grasses: Evidence for the generation and modification of novel regulators. *Journal of Integrative Plant Biology*. **49**, 927-939.
- Xu, W., Fiume, E., Coen, O., Pechoux, C., Lepiniec, L., Magnani, E., 2016. Endosperm and nucellus develop antagonistically in *Arabidopsis* seeds. *The Plant Cell*. **28**, 1343-1360.
- Yamada, K., Saraike, T., Shitsukawa, N., Hirabayashi, C., Takumi, S., Murai, K., 2009. Class D and B sister *MADS-box* genes are associated with ectopic ovule formation in the pistil-like stamens of alloplasmic wheat (*Triticum aestivum* L.). *Plant Molecular Biology*. **71**, 1-14.
- Yamaguchi, T., Lee, D.Y., Miyao, A., Hirochika, H., An, G., Hirano, H., 2006. Functional Diversification of the Two C-Class MADS Box Genes *OsMADS3* and *OsMADS58* in *Oryza sativa*. *The Plant Cell*. **18**, 15-28.
- Yamaguchi, T., Nagasawa, N., Kawasaki, S., Matsuoka, M., Nagato, Y., Hirano, H., 2004. The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in *Oryza sativa*. *The Plant Cell*. **16**, 500-509.
- Yamaki, S., Nagato, Y., Kurata, N., Nonomura, K., 2011. Ovule is a lateral organ finally differentiated from the terminating floral meristem in rice. *Developmental Biology*. **351**, 208-216.
- Yang, W., Ye, D., Xu, J., Sundaresan, V., 1999. The SPOROCTELESS gene of *Arabidopsis* is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes & Development*. **13**, 2108-2117.
- Yang, X., Wu, F., Lin, X., Du, X., Chong, K., Gramzow, L., Schilling, S., Becker, A., Theißen, G., Meng, Z., 2012. Live and let die-the Bsister *MADS-box* gene *OsMADS29* controls the degeneration of cells in maternal tissues during seed development of rice (*Oryza sativa*). *PloS One*. **7**, e51435.

References

- Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A., Meyerowitz, E.M., 1990. The protein encoded by the *Arabidopsis* homeotic gene *AGAMOUS* resembles transcription factors. *Nature*. **346**, 35.
- Yao, S., Ohmori, S., Kimizu, M., Yoshida, H., 2008. Unequal genetic redundancy of rice PISTILLATA orthologs, *OsMADS2* and *OsMADS4*, in lodicule and stamen development. *Plant and Cell Physiology*. **49**, 853-857.
- Yasui, Y., Tanaka, W., Sakamoto, T., Kurata, T., Hirano, H., 2017. Genetic enhancer analysis reveals that *FLORAL ORGAN NUMBER2* and *OsMADS3* co-operatively regulate maintenance and determinacy of the flower meristem in rice. *Plant and Cell Physiology*. **58**, 893-903.
- Yi, J., Moon, S., Lee, Y., Zhu, L., Liang, W., Zhang, D., Jung, K., An, G., 2016. Defective Tapetum Cell Death 1 (DTC1) regulates ROS levels by binding to metallothionein during tapetum degeneration. *Plant Physiology*. **170**, 1611-1623.
- Yin, L. & Xue, H., 2012. The *MADS29* transcription factor regulates the degradation of the nucellus and the nucellar projection during rice seed development. *The Plant Cell*. **24**, 1049-1065.
- Yoo, S.K., Lee, J.S., Ahn, J.H., 2006. Overexpression of *AGAMOUS*-LIKE 28 (*AGL28*) promotes flowering by upregulating expression of floral promoters within the autonomous pathway. *Biochemical and Biophysical Research Communications*. **348**, 929-936.
- Yoshida, H. & Nagato, Y., 2011. Flower development in rice. *Journal of Experimental Botany*. **62**, 4719-4730.
- Young, T.E. & Gallie, D.R., 2000. Programmed cell death during endosperm development. *Programmed cell death in higher plants*. Springer. 39-57.
- Young, T.E. & Gallie, D.R., 1999. Analysis of programmed cell death in wheat endosperm reveals differences in endosperm development between cereals. *Plant Molecular Biology*. **39**, 915-926.
- Young, T.E., Gallie, D.R., DeMason, D.A., 1997. Ethylene-mediated programmed cell death during maize endosperm development of wild-type and *shrunk2* genotypes. *Plant Physiology*. **115**, 737-751.
- Yuan, J.S., Reed, A., Chen, F., Stewart, C.N., 2006. Statistical analysis of real-time PCR data. *BMC Bioinformatics*. **7**, 85.
- Zha, H., Liu, T., Zhou, J., Sun, H., 2013. MS-desi, a desiccation-related protein in the floral nectar of the evergreen velvet bean (*Mucuna sempervirens* Hemsl): molecular identification and characterization. *Planta*. **238**, 77-89.
- Zhang, D. & Wilson, Z.A., 2009. Stamen specification and anther development in rice. *Chinese Science Bulletin*. **54**, 2342-2353.

References

- Zhang, D. & Yang, L., 2014. Specification of tapetum and microsporocyte cells within the anther. *Current Opinion in Plant Biology*. **17**, 49-55.
- Zhang, D. & Yuan, Z., 2014. Molecular control of grass inflorescence development. *Annual Review of Plant Biology*. **65**, 553-578.
- Zhang, H., Xu, H., Feng, M., Zhu, Y., 2018. Suppression of *OsMADS7* in rice endosperm stabilizes amylose content under high temperature stress. *Plant Biotechnology Journal*. **16**, 18-26.
- Zhang, R., Tucker, M.R., Burton, R.A., Shirley, N.J., Little, A., Morris, J., Milne, L., Houston, K., Hedley, P.E., Waugh, R., 2016. The dynamics of transcript abundance during cellularization of developing barley endosperm. *Plant Physiology*. **170**, 1549-1565.
- Zhang S, Klessig DF (2001) MAPK cascades in plant defense signaling. *Trends Plant Sci* 6:520-527
- Zhao, F.J., Su, Y.H., Dunham, S.J., Rakszegi, M., Bedo, Z., McGrath, S.P., Shewry, P.R., 2009. Variation in mineral micronutrient concentrations in grain of wheat lines of diverse origin. *Journal of Cereal Science*. **49**, 290-295.
- Zhao, T., Ni, Z., Dai, Y., Yao, Y., Nie, X., Sun, Q., 2006. Characterization and expression of 42 *MADS-box* genes in wheat (*Triticum aestivum* L.). *Molecular Genetics and Genomics*. **276**, 334.
- ZOOK, E.G., GREENE, F.E., Morris, E.R., 1970. Nutrient composition of selected wheats and wheat products. 6. Distribution of manganese, copper, nickel, zinc, magnesium, lead, tin, cadmium, chromium, and selenium as determined by atomic absorption spectroscopy and colorimetry. *Cereal Chemistry*. **47**, 720-731.
- Zhou, L., Song, G., Li, H., Hu, Y., He, B., 2008. A *MADS-box* transcription factor related to fertility conversion in male sterile wheat lines. *Acta Agronomica Sinica*. **34**, 598.

Appendix

Supplementary Tables:

Table S1: Primer sets used in this study for RT- PCR experiments in *Brachypodium*.

Gene name	Gene ID	Polarity	Amplicon Size (bp)	Primer sequences (5'-3')	Start-end
<i>BdACT7</i>	<i>Bradi4g41850</i>	Forward	188 bp	CCTGAAGTCCTTTCCAGCC	929-948
		Reverse		AGGGCAGTGATCTCCTTGCT	1097-1116
<i>BdMADS3</i>	<i>Bradi2g06330</i>	Forward	259 bp	TCCCATCAGCTGCAGCAACT	958-978
		Reverse		CTAGCGGTAGCACTTAGTAC	1197-1216
<i>BdMADS58</i>	<i>Bradi2g32910</i>	Forward	284 bp	AGCTGCAGAAATGACAACCTTG	693-712
		Reverse		TTCAGTAGCTTCAGTCTCAC	957-976
<i>BdMADS13</i>	<i>Bradi4g40350</i>	Forward	150 bp	AGTGCTTCTCCCCGCCAATC	936-956
		Reverse		TTAGAAATGGTGAGTTTGGTCGCC	1061-1084
<i>BdMADS21</i>	<i>Bradi2g25090</i>	Forward	177bp	GGAGATCGAGTACATGCAG	640-658
		Reverse		TCGTGCAGCTGCATGTTTAC	797-816
<i>Hpt-p2</i>	Hygromycin phosphotransferase	Forward	394 bp	CTCCAGTCAATGACCGCTGT	
		Reverse		TGACCTATTGCATCTCCCGC	
<i>Bd H4</i>	BRADI_2g23004v3	Forward	212 bp	AAGGTCCTCCGCGACAACAT	135-154
		Reverse		AGGGCGTAGACGACGTCCATG	326-346
<i>BdMT (promoter)</i>	BRADI_2g28254v3	Forward	565 bp	TTTAATCGATGATGCACCCG	1424-1453
		Reverse		TATTTATAAGGGAACAAGGC	-1959-19889
<i>BdPex (promoter)</i>	BRADI_2g12216v3	Forward	1986 bp	CCTACATGGAAACGATCTATCT	-15-84
		Reverse		GAGACGAGGAATAGTACTTG	1974-2000
<i>BdMADS7 (promoter)</i>	BRADI_3g41260v3	Forward	1789 bp	GTCCGCCGTTCGGACAAT	-212-229
		Reverse		CGTCCCGTCCTTGCCCGG	-1983-2000
<i>BdMADS7-CDS</i>	BRADI_3g41260v3	Forward	793 bp	GGAGAGGCAGGAGGAGGCG	302-320
		Reverse		TCATGGTAACCACGGCGGCAT	1074-1094

Table S2: Primer sets used in *In situ* hybridization (ISH) for *Brachypodium*.

Name	Primers sequence
BdMADS3 T7 R	5' GAATTGTAATACGACTCACTATAGGG CTAGCGGTAGCACTTAGTAC 3'
BdMADS58 T7 R	5' GAATTGTAATACGACTCACTATAGGG TTAGTAGCTTCAGTCTCAC 3'
BdMADS13 T7 R	5' GAATTGTAATACGACTCACTATAGGG AACGATGCCTGCGCTGAGC 3'
BdMADS21 T7 R	5' GAATTGTAATACGACTCACTATAGGG TCGTGCAGCTGCATGTTTAC 3'
Bd8d2RT7 (BdH4)	5' AATAGGACTCACTATAGGAATCCGAAGC 3'
BdMADS29	GAATTGTAATACGACTCACTATAGGG CCTCCACCGTGACCTTCTTA

Table S3: qPCR primer sets used in this study in *Brachypodium*.

Gene name	Gene ID	Polarity	Amplicon Size (bp)	Primer sequences (5'-3')	Start-end
<i>qBdACT7</i>	<i>Bradi4g41850</i>	Forward	101 bp	GTGACCTAACTGACTGCTTGAT	705-726
		Reverse		GAGCTTCTCCTTGATATCCCTTAC	782-805
<i>BdGAPDH</i>	BRADI_3G14120v3	Forward	236 bp	TTGCTCTCCAGAGCGATGAC	139-158
		Reverse		CTCCACGACATAATCGGCAC	374-355
<i>BdUBC18</i>	BRADI_4G00660v3	Reverse	193 bp	GGGCACCAGTCAACAACTG	1171-1153
		Forward		GGAGGCACCTCAGGTCATTT	206-225
<i>BdMADSq3</i>	<i>Bradi2g06330</i>	Forward	106 bp	GGATTCTGTCAGCACCATGA	645-664
		Reverse		GCATACAGCAGCTCATTCTTTC	728-749
<i>BdMADSq58</i>	<i>Bradi2g32910</i>	Forward	102 bp	GCTAAGTTGAGGCACCAGATAA	500-521
		Reverse		CTCCAAGTCTTAAGGTCTCTG	580-601
<i>BdMADSq13</i>	<i>Bradi4g40350</i>	Forward	102 bp	AGCAGGTGACAATAGCAAGG	870-889
		Reverse		CCGCTGCATCGAAGAGATT	953-971
<i>BdMADSq21</i>	<i>Bradi2g25090</i>	Forward	104 bp	CTTGAGAACAGGCTGGAGAAG	575-595
		Reverse		TGCAGATCCGCTTCCATTT	660-678
<i>BdA1 pepsin</i>	BRADI_1g42930v3	Forward	117 bp	GGCAGTCAACAACTTCCT	(1494-1413)
		Reverse		CGGTTCAACATCATTACCC	(1591-1610)
<i>BdC1 papain</i>	BRADI_2g08300v3	Forward	139 bp	CATGATGAGCTCTGATGCCAT	(1226-1247)
		Reverse		GTGGTGGTGTACTGACATACTG	(1343-1364)
<i>BdC13 nucellain</i>	BRADI_5g16960v3	Forward	96 bp	GCAGGACAATGCCCTCTGATTT	(2305-2326)
		Reverse		GGACTCTAGGAGGATCACAACCTTAC	(2375-2400)
<i>BdC14 metacaspase</i>	BRADI_2g52470v3	Forward	96 bp	TTCAGAGTGCTGGTGAGGTTTATG	(1222-1245)
		Reverse		CTGCTGATGTTTGGCTGGTTTG	(1296-1317)
<i>BdMADS29</i>	BRADI_3g05260v3	Forward	377 bp	AACACTCTCCTGTGCCGCAT	(1083-1102)
		Reverse		CCTCCACCGTGACCTTCTTA	(1440-1459)
<i>Bd papain</i>	BRADI_1g08357v3	Forward	78 bp	CCAGGCATGGCTTAAACTAGA	1317-1337
		Reverse		TGTACCGGCCCTACTTCTAT	1375-1394
<i>Bd metallothionein</i>	BRADI_4g09257v3	Forward	108 bp	AAATAAGCGCCGAGTCAGAG	436-455
		Reverse		CACACACAGACACAGATACA	522-543
<i>Bd peroxidase</i>	BRADI_2g11320v3	Forward	113 bp	CACGCTGCCGTACTGATAAA	1147-1166
		Reverse		GAACTCCAAGCTAGTCCACAC	1239-1259
<i>Bd MYB</i>	BRADI_2g13920v3	Forward	115 bp	TGCAGATGGCAGTTGAGATAG	927-947
		Reverse		GGTTAGGAACATTAGGCTGTAGAT	1018-1041
<i>Bd thioredoxin</i>	BRADI_3g00210v3	Forward	115 bp	GCTACTCAATGCTGGCTACAA	1779-1799
		Reverse		TCGAGCTGACCACTCTAACA	1874-1893
<i>Bd F-box</i>	BRADI_3g48970v3	Forward	117 bp	TCAAAGTCTGGTACAGGGAAC	5336-5357
		Reverse		GAGGCATCTAGCGTCCATTT	5433-5452

Table S4: Numbers of genes found in each species for each family.

			<i>Brachypodium distachyon</i>	<i>Hordeum vulgare</i>	<i>Arabidopsis thaliana</i>	<i>Oryza sativa</i>	<i>Triticum aestivum</i>
families	Uniprot ID	Panther ID					
A1 pepsin	IPR001461	PTHR13683	93	128	69	102	450
C1 papain	IPR000668	PTHR12411	48	65	39	34	185
C13 nucellain	IPR001096	PTHR12000	5	8	5	6	36
C14 metacaspase	PTHR31773	PTHR31810	10	10	9	10	22

Table S5: qPCR primers used in this work for wheat and barley.

Gene name	Gene ID	Polarity	Amplicon Size (bp)	Primer sequences (5' – 3')	Start-end
<i>HvGAPDH</i>	HORVU6Hr1G054520	Forward	121 bp	GTTGGCAAGGTGCTCCCAGA	(780–799)
		Reverse		CAAGCCAGCCACCTATGAGC	(881–900)
<i>HvMADS3</i>	HORVU3Hr1G026650	Forward	100 bp	GTCCACGCACGCTAGATAATAC	(1050–1071)
		Reverse		TGATCCACAAATTCGGTTACT	(1128–1149)
<i>HvMADS58</i>	HORVU1Hr1G029220	Forward	114 bp	AGAAATACCAAGGGCCACAG	(17–36)
		Reverse		GGAAGGCAAATGAGGGAGAA	(111–130)
<i>HvMADS13</i>	HORVU1Hr1G023620	Forward	85 bp	CTCTGGCTTGGGTCTCTATAAAC	(52–73)
		Reverse		AAGAGCAGACGATGGACAAG	(117–136)
<i>TaGAPDH_A</i>	TraesCS6A02G213700	Forward	187 bp	GGTGCCAAGAAGGTCATCAT	(438–457)
		Reverse		TGGTCATCAAACCTCAACA	(1296–1317)
<i>TaMADS3_A</i>	TraesCS3A02G314300	Forward	122 bp	AGAATGAGAGGGGTCAGCAGC	(605–624)
		Reverse		CTGCTGCATGATGTTTCGCCTGC	(831–852)
<i>TaMADS58_D</i>	TraesCS1D02G127700.2	Forward	86 bp	ACAAAGGCCTTGGAAGATTAG	(647–668)
		Reverse		TTATTGTGTACCCTTCTCTGC	(711–732)
<i>TaMADS13_A</i>	TraesCS5A02G117500.1	Forward	102 bp	AGAATGAGCTCCTATCTTCTGAG	(562–584)
		Reverse		GCTCCTCCTCCGCAATC	(647–663)
<i>TaMADS21_A</i>	TraesCS1A02G262700.2	Forward	122 bp	AGCTGCTGTTTCGCGGAG	(731–751)
		Reverse		TTACAGTACACTGAGAAGCATGCAGGA	(831–852)

Appendix

Table S6: K means clustering correlation coefficient results for selected genes from STRING (www.string-db.org/) calculated based on Davidson et al 2012 study expressions. The yellow highlights are MADS-box genes and blue are corresponding to the MAPK genes.

	BRADI2 G0633 0	BRADI2 G1562 0	BRADI2 G1633 7	BRADI2 G2509 0	BRADI2 G3291 0	BRADI2 G3647 0	BRADI2 G4435 0	BRADI2 G4501 0	BRADI2 G4587 0	BRADI3 G0378 0	BRADI3 G1656 0	BRADI3 G3200 0	BRADI4 G2491 2	BRADI4 G4035 0
BRADI2 G0633 0	1													
BRADI2 G1562 0	0.903	1												
BRADI2 G1633 7	-0.489	-0.389	1											
BRADI2 G2509 0	0.107	-0.312	-0.258	1										
BRADI2 G3291 0	0.365	-0.067	-0.330	0.938	1									
BRADI2 G3647 0	0.448	0.477	0.245	-0.253	-0.011	1								
BRADI2 G4435 0	0.901	0.999	-0.380	-0.320	-0.072	0.501	1							
BRADI2 G4501 0	0.905	1.000	-0.392	-0.308	-0.063	0.477	0.999	1						
BRADI2 G4587 0	0.223	0.363	0.680	-0.423	-0.288	0.737	0.375	0.360	1					
BRADI3 G0378 0	0.643	0.509	-0.613	0.201	0.434	0.248	0.509	0.511	-0.076	1				
BRADI3 G1656 0	-0.440	-0.279	0.843	-0.423	-0.430	0.505	-0.257	-0.283	0.715	-0.311	1			
BRADI3 G3200 0	-0.499	-0.511	0.075	-0.011	-0.036	0.247	-0.492	-0.510	-0.155	-0.037	0.402	1		
BRADI4 G2491 2	0.156	0.006	-0.200	0.157	0.306	0.432	0.024	0.009	-0.158	0.059	-0.085	0.534	1	
BRADI4 G4035 0	0.062	-0.325	-0.286	0.937	0.869	-0.339	-0.332	-0.322	-0.468	0.282	-0.409	-0.069	-0.012	1

Table S7: Gene names and IDs used in figure 4 and Figure S5, grouped in C13 and C14 families for each specie.

Species	family	Gene Name	Gene ID	UniProt	Ref
<i>Arabidopsis thaliana</i>	C13	At-gamma-VPE	AT4G32940	A0A178UU68	(Kinoshita, Nishimura and Hara-Nishimura, 1995)
		At-beta-VPE	AT1G62710	A0A178W0Z7	
		At-alpha-VPE	AT2G25940	A0A178VQP4	
		At-delta-VPE	AT3G20210	A0A178VB13	
		At-delta-VPE variant	AT1G08750	A0A178WNN3	
	C14	AtMCAs1	At1g02170	A0A178W8H4	(Fagundes <i>et al</i> , 2015)
		AtMCAs2	At4g25110	A0A178V2G1	
		AtMCAs3	At5g64240	F4KDK6	
		AtMCAs4	At1g79340	A0A178WK95	
		AtMCAs5	At1g79330	A0A178WIC7	
		AtMCAs6	At1g79320	A0A178W108	
		AtMCAs7	At1g79310	A0A178WN22	
		AtMCAs8	At1g16420	A0A178WDC5	
		AtMCAs9	At5g04200	A0A178U6S6	
<i>Oryza sativa</i>	C13	OsVPE1/ GLUP3	Os04g0537900	Q84LM2	(Deng <i>et al</i> , 2011)
		OsVPE2	Os01g0559600	Q7F1B4	
		OsVPE3	Os02g0644000	Q8GS39	
		OsVPE4a	Os05g0593900	Q6L4R2	our study
		OsVPE4b	Os06g0105100	Q9LWZ3	
		OsVPE5	Os02g0219400	Q6Z6K3	(Wang <i>et al</i> , 2018)
	C14	OsMC1	Os03g0388900	Q75LQ1	(Wang and Zhang, 2014) and (Fagundes <i>et al</i> , 2015)
		OsMC2	Os03g0389400	A0A0P0VY86	
		OsMC3	Os03g0389000	A0A0P0VY92	
		OsMC4	Os05g0496400	A0A0P0WP05	
		OsMC5	Os05g0496500	Q84VF0	
		OsMC6	Os01g0799900	Q8LJ88	
		OsMC7	Os11g0134700	Q2RAW9	
		OsMC8	Os03g0389100	Q75LQ8	
		OsMC9	Os03g0389501	A0A0P0VYB1	our study
		OsMC10	Os10g0565100	A0A0P0XYC7	
<i>Hordeum vulgare</i>	C13	HvLeg-1 /HvVPE1	HORVU6Hr1G060990	B4ESD9	(Radchuk <i>et al</i> , 2010) and (Julián <i>et al</i> , 2013)
		HvLeg-2 /HvVPE2b	HORVU2Hr1G092080	B4ESE0	

Appendix

		HvLeg-3 /HvVPE2d	HORVU2Hr1G091880	B4ESE1	
		HvLeg-4 /HvVPE3	HORVU3Hr1G048520	B4ESE2	
		HvLeg-5 /HvVPE4	HORVU5Hr1G066250	B4ESE3	
		HvLeg-6 /HvVPE2a, Nucellain	HORVU2Hr1G092080	E5AXU4	
		HvLeg-7 /HvVPE2c		E5AXU6	
		HvLeg-8		F2DJF5	
	C14	HvMC1	HORVU3Hr1G095700	A0A287MDW0	(Bostancioglu, Tombuloglu and Tombuloglu, 2018)
		HvMC2	HORVU4Hr1G090860	A0A287Q5C5	
		HvMC3	HORVU1Hr1G055210	A0A287FP58	
		HvMC4	HORVU3Hr1G020830	A0A287KD76	
		HvMC5	HORVU1Hr1G071130	A0A287G509	
		HvMC6	HORVU5Hr1G023940	A0A287QP80	
		HvMC7	HORVU3Hr1G078270	A0A287LT80	
		HvMC8	HORVU5Hr1G044610	A0A287R060	
		HvMC9	HORVU4Hr1G011900	A0A287N8J8	
		HvMC10	HORVU1Hr1G067230	A0A287G1X9	
<i>Brachypodium distachyon</i>	C13	BdVPE1	BRADI_3g50100v3	I1IC16	our study
		BdVPE2	BRADI_5g16960v3	I1J043	
		BdVPE3	BRADI_2g41270v3	I1HNN1	
		BdVPE4	BRADI_4g30110v3	I1IQ27	
		BdVPE5	BRADI_3g08190v3	I1HYS5	
	C14	BdMC1	Bradi_3g33722v3	A0A2K2D0W7	(Fagundes <i>et al</i> , 2015)
		BdMC2	Bradi_1g60762v3	I1H4W5	
		BdMC3	Bradi_1g60756v3	I1H4W4	
		BdMC4	Bradi_1g60800v3	I1H4W9	
		BdMC5	Bradi_1g60787v3	I1H4W8	
		BdMC6	Bradi_1g60777v3	I1H4W7	
		BdMC7	Bradi_2g21100v3	I1HI23	
		BdMC8	Bradi_2g21110v3	I1HI24	
		BdMC9	Bradi_2g52470v3	I1HSI1	
		BdMC10	Bradi_2g50480v3	I1HRT2	

Supplementary figures:

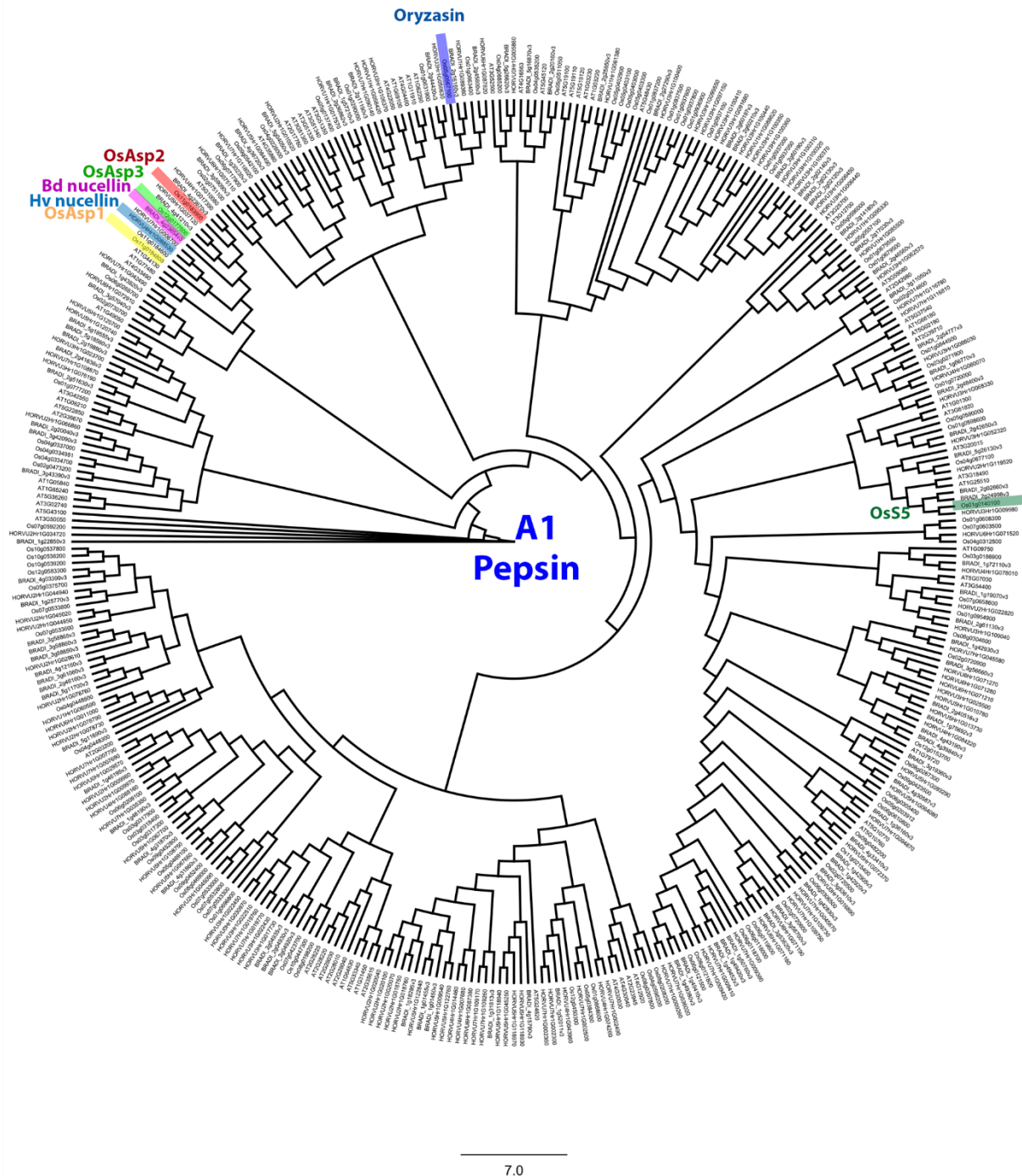


Figure S1: Phylogenetic Tree of pepsin-like A1 family. It includes aspartic protease from *Arabidopsis thaliana*, *Oryza sativa*, *Brachypodium distachyon* and *Hordeum vulgare*. Identified genes in Hv, Bd and Os were highlighted. JTT+G was used as protein model and the sequences were aligned with MAFFT v7.308 (Kato and Standley, 2013), then inferred using Maximum Likelihood method with a bootstrap of 1,000 replicates. The tree was generated using Geneious R10.

Figure S2: co-expression analysis of MADS3 in *Arabidopsis*, *wheat* rice and *barley*, using GENEVESTIGATOR.



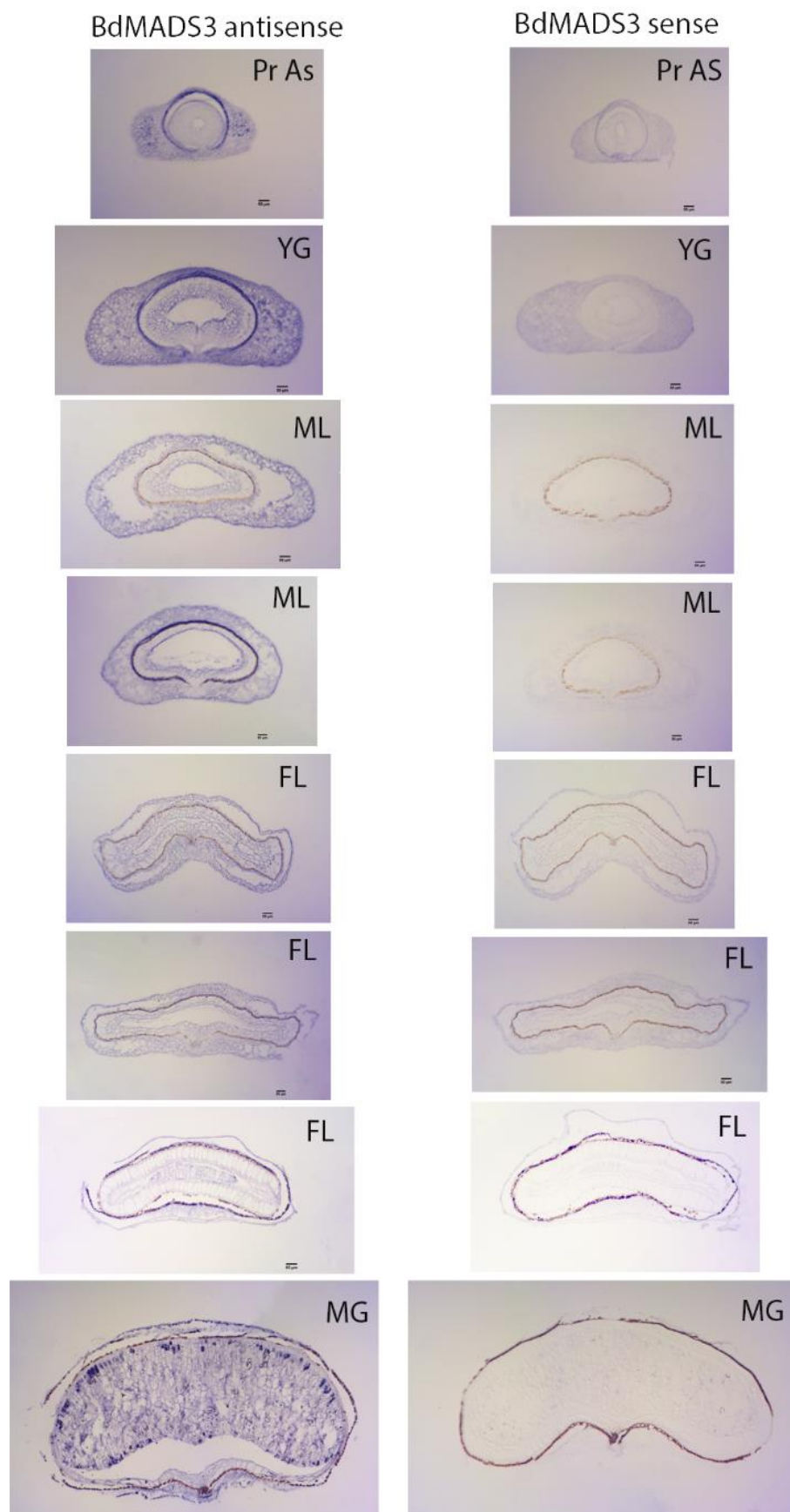


Figure S3: *In situ* hybridisation of sense and anti-sense of BdMADS3 in different grain development stages in Brachypodium.

Appendix

Table S8: List of studies used by wheat expression browser to generate the heat map for TaMADS box genes expression in different tissues.

number	Name of the study
1	Grain tissue-specific developmental timecourse
2	Developmental time-course of Chinese Spring
3	Chinese Spring seedling (leaves and roots) and spikes at anthesis
4	Chinese Spring leaves and roots from seven leaf stage
5	Six unreplicated tissues from Chinese Spring.
6	Chinese Spring flag leaves, 6 timepoints
7	Chinese Spring early meiosis, early prophase
8	Developmental time-course of Azhurnaya
9	Phosphate starvation in roots and shoots
10	Drought and heat stress time-course in seedlings
11	Spikes with water stress
12	Seedlings with PEG to simulate drought
13	Shoots after 2 weeks cold
14	Fusarium head blight infected spikelets
15	Stripe rust infected seedlings
16	Septoria tritici infected seedlings
17	Stripe rust and powdery mildew timecourse of infection in seedlings
18	PAMP inoculation of seedlings
19	Timecourse of spikelets inoculated with fusarium head blight
20	Coleoptile infection with Fusarium pseudograminearum (crown rot)
21	Leaves naturally infected with Magnaporthe oryzae
22	spikes inoculated with fusarium head blight and ABA/GA
23	Zymoseptoria tritici infected seedlings
24	CS spikes inoculated with fusarium head blight
25	Shoots from NILs segregating for crown rot resistance
26	Stripe rust infected seedlings
27	Flag leaf downregulation of GPC
28	Grain tissue-specific developmental timecourse
29	Shoots and leaves of nulli tetra group 1 and group 5
30	Grain tissue-specific developmental timecourse
31	Comparison of stamen, pistil and pistilloidy expression
32	Developmental time-course of synthetic hexaploid
33	Grain tissue-specific expression at 12 days post anthesis
34	Grain developmental timecourse with 4A dormancy QTL
35	Microspores in tissue culture and cold
36	Gene expression during a timecourse of flag leaf senescence

MADS-box domain	<div>100110120130140150</div> <div>MGRGRIEIKRIENTTQRQVTFCKRRNGLLKKAYELSVLCDAEVALIVFSSRGRLYEYS</div> <div>MGRGRIEIKRIENTTQRQVTFCKRRNGLLKKAYELSVLCDAEVALIVFSSRGRLYEYS</div>
I domain	<div>160170180</div> <div>NSVKATIERYKKANxDTSNSG-xVAEVNA-QY</div> <div>NSVKATIERYKKANxDTSNSG-xVAEVNA-QY</div>
K domain	<div>190200210220230240250</div> <div>QESAKLRHQIQSLQNSN-RHLVGDSVGTMSLXLKQLEGRLEKGIXKIRARKNELLEXAEIYMQKREME</div> <div>QESAKLRHQIQSLQNSN-RHLVGDSVGTMSLXLKQLEGRLEKGIXKIRARKNELLEXAEIYMQKREME</div>

Figure S4: MADS-box transcription factor main domain consensus.

MADS-box domain consensus: MGRGRIEIKRIENTTQRQVTFCKRRNGLLKKAYELSVLCDAEVALIVFSSRGRLYEYS

I domain consensus: NSVKATIERYKKANxDTSNSG-xVAEVNAQY

K domain consensus:
QESAKLRHQIQSLQNSN-RHLVGDSVGTMSLXLKQLEGRLEKGIXKIRARKNELLEXAEIYMQKREME

Appendix

Table S9: Go term annotation of the overlap (Vyinf and PrAs samples) list of the genes that are up-regulated by BdMADS13, IEA: Inferred from Electronic Annotation, IDA: inferred by direct assay, IBA: Inferred from Biological aspect of Ancestor, IMP: inferred by mutant phenotype

Sequence Name	Sequence Description	Mapping GO Term	Annotation GO Term	Enzyme Name	InterPro GO Term
BRADI_5g26380 v3					
BRADI_5g01574 v3	receptor like protein kinase		protein serine/threonine kinase activity		Protein kinase domain
BRADI_4g13987 v3	Disease resistance RPP13-like		cytoplasm vesicle-mediated transport exocytosis protein-containing complex ion binding		ADP binding; Domain of unknown function DUF1618; Leucine-rich repeat domain superfamily; NB-ARC; P-loop containing nucleoside triphosphate hydrolase
BRADI_4g13380 v3					
BRADI_3g48970 v3	probable inactive histone-lysine N-methyltransferase		cellular protein modification process histone-lysine N-methyltransferase activity zinc ion binding chromosome organization	Histone-lysine N-methyltransferase	F-box-like domain superfamily; Pre-SET domain; histone-lysine N-methyltransferase activity/nucleus/zinc ion binding/histone lysine methylation
BRADI_3g41691 v3	1-aminocyclopropane-1-carboxylate oxidase		oxidoreductase activity metal ion binding		oxidoreductase activity/oxidation-reduction process; Oxoglutarate/iron-dependent dioxygenase; Isopenicillin N synthase-like; Non-haem dioxygenase N-terminal domain
BRADI_3g08301 v3	ATP-dependent RNA helicase drs1		helicase activity translation factor activity RNA binding ion binding	Nucleoside-triphosphate phosphatase	Domain of unknown function DUF4283
BRADI_3g01143 v3	Helicase protein MOM1		DNA binding DNA-binding transcription factor activity nucleus biosynthetic process		Non-haem dioxygenase N-terminal domain; Chromo-like domain

Appendix

			cellular nitrogen compound metabolic process metal ion binding HDA1 complex		
BRADI_3g00210 v3	protein disulfide isomerase-like 1-4		mitochondrion vacuole, vacuolar membrane endoplasmic reticulum protein folding response to oxidative stress plastid protein disulfide isomerase activity homeostatic process	Protein disulfide- isomerase	Myb-like transcription factor; Thioredoxin domain ; cell redox homeostasis.
BRADI_2g34547 v3	DCD (Development and Cell Death) domain protein		Kelch motif		Development/cell death domain
BRADI_2g33690 v3	predicted protein		retrotransposon nucleocapsid-IEA		Development/cell death domain
BRADI_2g27197 v3	predicted protein		membrane-IEA integral component of membrane-IEA regulation of transcription, DNA- templated-IEA		Senescence regulator S40; F- box domain
BRADI_2g23797 v3	A/G-SPECIFIC ADENINE GLYCOSYLASE/ENDONUCLE ASE III	DNA repair-IEA DNA demethylation- IDA base-excision repair-IEA	molecular function DNA metabolic process response to stress		HhH-GPD domain
BRADI_2g23341 v3	transcription factor MYB3R-2	DNA binding- IEA membrane-IEA integral component of membrane-IEA nucleus-IEA	DNA binding nucleus		The fantastic four family; Myb-like transcription factor
BRADI_2g13920 v3	hypothetical protein BRADI_2g13920v3		integral cellular component of membrane		Myb-like transcription factor
BRADI_2g07820 v3	POLYCOMB PROTEIN EED	protein binding- IPI negative regulation of transcription by RNA polymerase II- IBA PcG protein complex-IDA histone methyltransferase activity-IBA cell differentiation- IEA histone H3- K27 methylation- IMP	reproduction nucleoplasm cellular protein modification process methyltransferase activity biosynthetic process developmental maturation cell differentiation protein-containing complex cellular nitrogen compound metabolic process anatomical structure development	Histone-lysine N- methyltransferase	WD domain, G- beta repeat

Appendix

		nucleosome binding-IBA maintenance of seed dormancy-IMP multicellular organism development-IEA endosperm development-IMP ESC/E(Z) complex-IBA post-embryonic plant organ development-IMP histone methyltransferase activity (H3-K27 specific)-IBA chromatin silencing complex-IMP	chromosome organization		
BRADI_2g05802 v3	F-box/LRR-repeat protein				Senescence regulator S40
BRADI_1g67440 v3	protein FAF-like, chloroplastic		Fantastic Four meristem regulator		The fantastic four family
BRADI_1g62007 v3	retrotransposon protein	histone acetyltransferase complex-IEA nucleotide binding-IEA histone acetylation-IEA phosphorylation-IEA transferase activity-IEA kinase activity-IEA transferase activity, transferring hexosyl groups-IEA zinc ion binding-IEA DNA integration-IEA retrotransposon nucleocapsid-IEA pollen development-IEA protein serine/threonine kinase activity-IEA	nucleus biological process ion binding Alpha-N-acetylglucosaminidase (NAGLU) tim-barrel domain, N-terminal domain, C-terminal domain.		Reverse transcriptase, RNA-dependent DNA polymerase; Ribonuclease H superfamily; Retrotransposon Ty1/copia-like; Integrase, catalytic core; Ribonuclease H-like superfamily; nucleic acid binding; retrotransposon nucleocapsid; DNA integration

Appendix

		ATP binding-IEA protein phosphorylation-IEA plasma membrane-IEA			
BRADI_1g59720 v3	predicted protein		Senescence regulator		Senescence regulator S40
BRADI_1g21455 v3	TPA: orf y	nucleic acid phosphodiester bond hydrolysis-IEA RNA-directed DNA polymerase activity-IEA endonuclease activity-IEA RNA-dependent DNA biosynthetic process-IEA transferase activity-IEA hydrolase activity-IEA nucleic acid binding-IEA	nuclease activity DNA metabolic process biosynthetic process Nucleotidyl-transferase activity zinc ion binding	RNA-directed DNA polymerase	

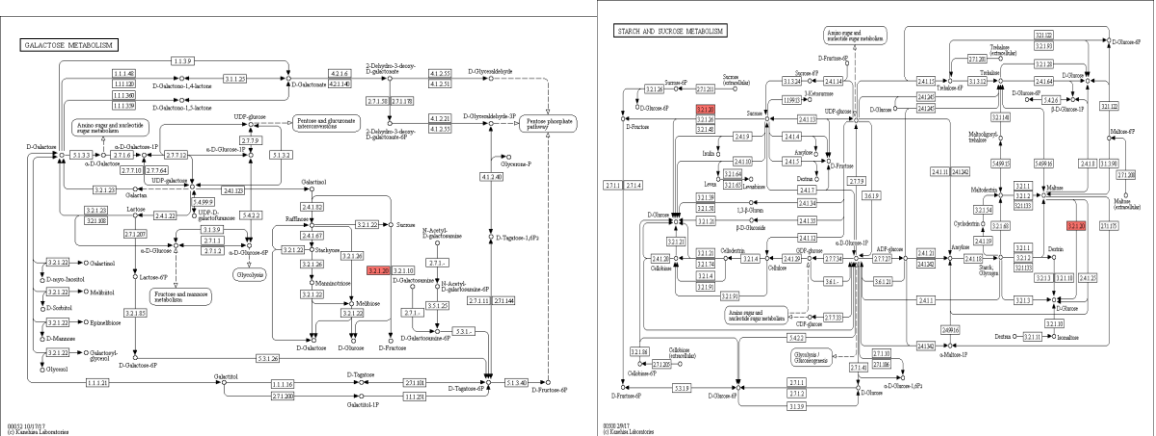


Figure S5: KEGG pathway analysis of the down-regulated genes list.

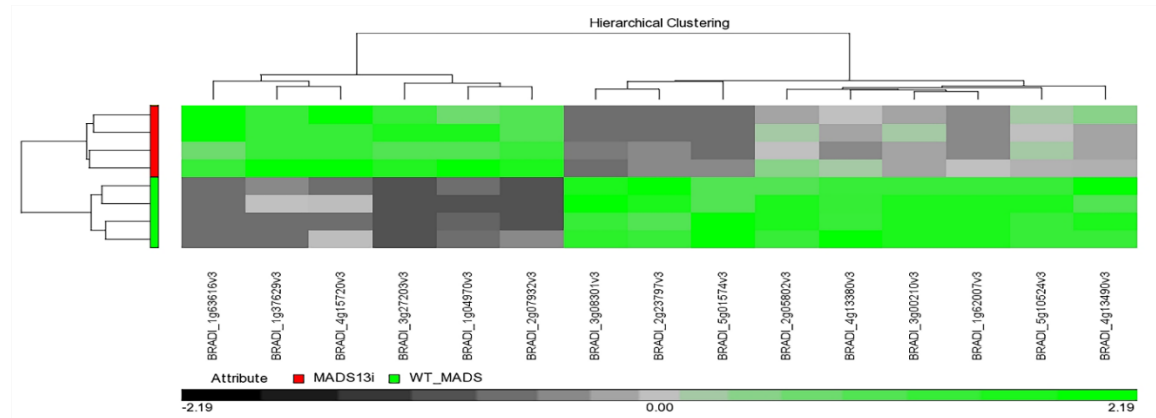


Figure S6: down regulated genes list from Partek analysis.

Appendix

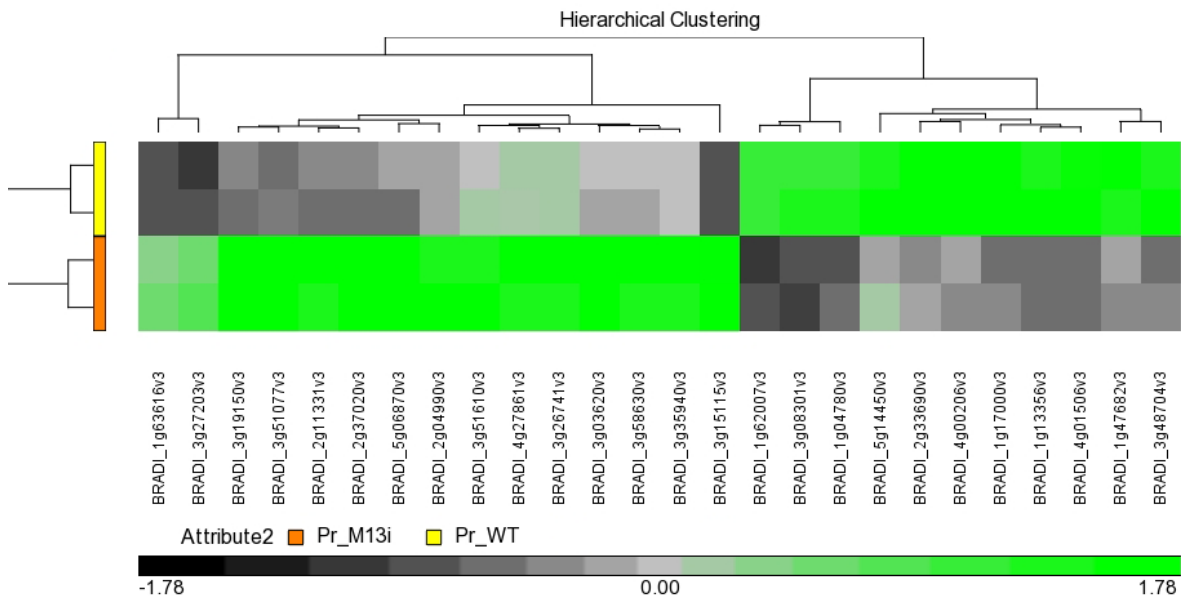


Figure S7: PrM13i high in Peroxidase and papain, and down in RNA HELICASE and Protein kinase domain



Figure S8: Constructs used in this stud chapter 4.

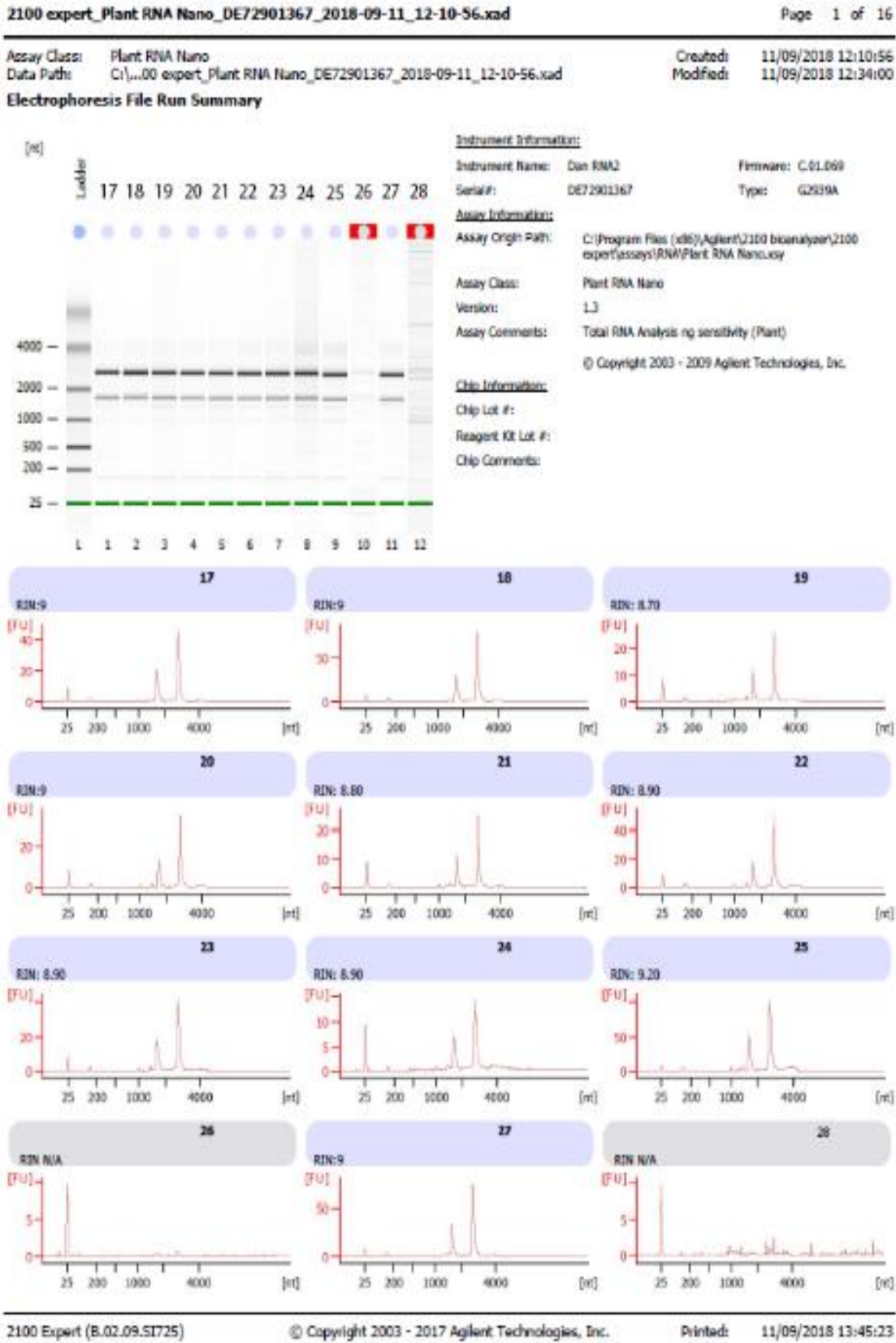


Figure S9: RNA Integrity Number (RIN), for samples used in the RNA-seq.

Appendix



Appendix

Table S10: Co-express genes for only **TaMADS3** (TraesCS3D01G140200, TraesCS3A01G314300 and TraesCS3B01G157500). Top 50 genes.

downstream.gene	TF	directed_edge_weight	Description
TraesCS6D01G267900	TraesCS3D01G140200	0.024969082	Early nodulin-like protein 1
TraesCS1B01G005400	TraesCS3A01G314300	0.023268056	Cyclotide
TraesCS2A01G414400	TraesCS3A01G314300	0.021735942	Pollen Ole e 1 allergen and extensin family protein, putative
TraesCS6D01G142300	TraesCS3A01G314300	0.021337561	Centrosomal protein of 76 kDa
TraesCS5A01G507000	TraesCS3D01G140200	0.020793543	Purple acid phosphatase
TraesCS7D01G496000	TraesCS3D01G140200	0.02040841	Carboxypeptidase
TraesCS6D01G267900	TraesCS3A01G314300	0.020286659	Early nodulin-like protein 1
TraesCS1B01G139700	TraesCS3A01G314300	0.019828942	Cytochrome P450 family protein, expressed
TraesCS3B01G132100	TraesCS3D01G140200	0.019738369	bZIP transcription factor, putative (DUF1664)
TraesCS2B01G258700	TraesCS3B01G157500	0.019678622	ROP-interactive CRIB motif protein
TraesCSU01G146900	TraesCS3D01G140200	0.019607932	Proline-rich protein
TraesCS5A01G173600	TraesCS3A01G314300	0.019604876	Mitochondrial import inner membrane translocase subunit Tim22
TraesCS7A01G383800	TraesCS3A01G314300	0.019486138	MADS box transcription factor
TraesCS7B01G451200	TraesCS3D01G140200	0.019481348	Bidirectional sugar transporter SWEET
TraesCS4A01G109400	TraesCS3B01G157500	0.019240097	Syntaxin-binding protein 5
TraesCS5A01G511000	TraesCS3D01G140200	0.019039141	Phospholipase A1
TraesCS6B01G455800	TraesCS3A01G314300	0.019014175	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-3
TraesCS7B01G179800	TraesCS3B01G157500	0.018993137	Calcium binding protein
TraesCS1D01G344800	TraesCS3B01G157500	0.018988215	Soluble inorganic pyrophosphatase
TraesCS1D01G344800	TraesCS3B01G157500	0.018988215	Inorganic pyrophosphatase family protein
TraesCS2D01G390800	TraesCS3A01G314300	0.01882928	HXXXD-type acyl-transferase family protein, putative
TraesCS7B01G376800	TraesCS3B01G157500	0.018723818	Translation initiation factor IF-2
TraesCS3D01G359600	TraesCS3D01G140200	0.01867273	Small ubiquitin-related modifier
TraesCS4D01G150500	TraesCS3A01G314300	0.018602842	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
TraesCS3A01G400600	TraesCS3B01G157500	0.018583541	Protein kinase family protein
TraesCS3A01G400600	TraesCS3B01G157500	0.018583541	Protein kinase family protein
TraesCS7D01G190400	TraesCS3A01G314300	0.018432397	CASP-like protein
TraesCS5A01G238300	TraesCS3D01G140200	0.018370568	GDSL esterase/lipase
TraesCS2B01G333700	TraesCS3D01G140200	0.018337361	Glutamate decarboxylase
TraesCS2B01G333700	TraesCS3D01G140200	0.018337361	Glutamate decarboxylase
TraesCS5B01G520100	TraesCS3B01G157500	0.018319913	Subtilisin-like protease
TraesCS7B01G468800	TraesCS3D01G140200	0.018239608	11S globulin seed storage protein 2
TraesCS5A01G042900	TraesCS3D01G140200	0.01819788	Cationic amino acid transporter, putative
TraesCS5D01G487300	TraesCS3B01G157500	0.018128408	Glycine-rich protein A3
TraesCS3D01G140200	TraesCS3A01G314300	0.018056719	AGAMOUS MADS-box transcription factor
TraesCS3D01G140200	TraesCS3A01G314300	0.018056719	AGAMOUS MADS-box transcription factor
TraesCS3D01G140200	TraesCS3A01G314300	0.018056719	AGAMOUS MADS-box transcription factor
TraesCS3D01G140200	TraesCS3A01G314300	0.018056719	AGAMOUS MADS-box transcription factor
TraesCS1D01G020200	TraesCS3B01G157500	0.017974189	Phosphoinositide phosphatase family protein
TraesCS2A01G059500	TraesCS3D01G140200	0.017896598	Protease inhibitor/seed storage/lipid transfer family protein
TraesCS2B01G381900	TraesCS3D01G140200	0.017852509	Lung seven transmembrane receptor, putative
TraesCS5A01G001100	TraesCS3A01G314300	0.017767631	Sugar transporter, putative
TraesCS2A01G240300	TraesCS3B01G157500	0.01765377	ROP-interactive CRIB motif protein
TraesCS6B01G455800	TraesCS3D01G140200	0.017597338	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-3
TraesCS3B01G593800	TraesCS3A01G314300	0.017540287	11S globulin seed storage protein 2
TraesCS7B01G135400	TraesCS3A01G314300	0.017512195	BURP domain protein RD22
TraesCS1D01G165200	TraesCS3B01G157500	0.017471994	LIM domain containing protein
TraesCS1B01G220900	TraesCS3B01G157500	0.01747141	F-box and Leucine Rich Repeat domains containing protein, putative isoform 1
TraesCS5B01G171000	TraesCS3D01G140200	0.017457033	Mitochondrial import inner membrane translocase subunit Tim22
TraesCS3B01G477900	TraesCS3D01G140200	0.017340872	Membrane protein LEM1

Appendix

Table S11: Co-express genes for only **TaMADS58** (TraesCS1D01G127700, TraesCS1A01G125800 and TraesCS1B02G144800). Top 50 genes.

downstream.gene	TF	directed_edge_weight	Description
TraesCS3D01G052700	TraesCS1D01G127700	0.034877595	Photosystem II reaction center protein L
TraesCS3A01G215700	TraesCS1D01G127700	0.031737857	Photosystem II CP47 reaction center protein
TraesCS3D01G052500	TraesCS1A01G125800	0.030221975	Cytochrome f
TraesCS4D01G187400	TraesCS1A01G125800	0.029591266	Retrovirus-related Pol polyprotein from transposon 297 family
TraesCS1D01G181900	TraesCS1D01G127700	0.029164096	Photosystem II reaction center protein L
TraesCS1B01G385700	TraesCS1D01G127700	0.027892598	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B
TraesCS4B01G159300	TraesCS1A01G125800	0.027875224	Ribosomal protein S19
TraesCS3D01G052700	TraesCS1A01G125800	0.026847235	Photosystem II reaction center protein L
TraesCS1A01G384400	TraesCS1A01G125800	0.026816546	NAD(P)H-quinone oxidoreductase subunit 3, chloroplastic vacuolar ATP synthase catalytic subunit-related / V-ATPase-related / vacuolar proton pump-like protein
TraesCS3B01G584600	TraesCS1D01G127700	0.026558477	
TraesCS5A01G436700	TraesCS1D01G127700	0.026293801	50S ribosomal protein L2
TraesCS4A01G074600	TraesCS1A01G125800	0.025959987	sulfate transporter 91
TraesCS3D01G202200	TraesCS1A01G125800	0.025783594	carboxyl-terminal peptidase, putative (DUF239)
TraesCS1D01G087000	TraesCS1A01G125800	0.025738525	FBD-associated F-box protein
TraesCS2A01G210900	TraesCS1A01G125800	0.025535387	50S ribosomal protein L23
TraesCS5D01G010900	TraesCS1A01G125800	0.025437508	Photosystem I P700 chlorophyll a apoprotein A1
TraesCS5D01G010900	TraesCS1D01G127700	0.025174209	Photosystem I P700 chlorophyll a apoprotein A1
TraesCS4D01G137600	TraesCS1D01G127700	0.025100357	Mitochondrion orf173
TraesCS4A01G151900	TraesCS1D01G127700	0.024156854	ATP synthase subunit beta
TraesCS4A01G151900	TraesCS1A01G125800	0.024127208	ATP synthase subunit beta
TraesCS2A01G257900	TraesCS1A01G125800	0.024022304	Ribosomal protein L11 methyltransferase
TraesCS4A01G161600	TraesCS1A01G125800	0.023837991	Cytochrome b559 subunit alpha
TraesCS2A01G242000	TraesCS1A01G125800	0.023607226	sulfate transporter 91
TraesCS3D01G052500	TraesCS1D01G127700	0.023263289	Cytochrome f
TraesCS2A01G257900	TraesCS1D01G127700	0.023181525	Ribosomal protein L11 methyltransferase
TraesCS4D01G137600	TraesCS1A01G125800	0.023156161	Mitochondrion orf173
TraesCS4B01G026900	TraesCS1A01G125800	0.023061705	alpha-mannosidase 1
TraesCS1A01G220400	TraesCS1A01G125800	0.023032543	Cell number regulator 6
TraesCS2A01G210900	TraesCS1D01G127700	0.022948121	50S ribosomal protein L23
TraesCS6D01G079800	TraesCS1D01G127700	0.022867763	Maturase K
TraesCS3D01G202200	TraesCS1D01G127700	0.022840891	carboxyl-terminal peptidase, putative (DUF239)
TraesCS3B01G186500	TraesCS1D01G127700	0.022540789	Photosystem II reaction center protein L
TraesCS5D01G008200	TraesCS1D01G127700	0.022505651	30S ribosomal protein S16
TraesCS3D01G053400	TraesCS1D01G127700	0.022467296	Photosystem II CP47 reaction center protein
TraesCS4A01G156100	TraesCS1D01G127700	0.022385845	30S ribosomal protein S7
TraesCS4A01G024500	TraesCS1D01G127700	0.0223539	NAD(P)H-quinone oxidoreductase subunit 2, chloroplastic AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kinase family protein
TraesCS3A01G458800	TraesCS1D01G127700	0.022298512	
TraesCS7D01G261600	TraesCS1D01G127700	0.022168151	MADS-box transcription factor
TraesCS7D01G261600	TraesCS1D01G127700	0.022168151	MADS-box transcription factor
TraesCS1A01G125800	TraesCS1D01G127700	0.022088439	AGAMOUS MADS-box transcription factor
TraesCS1A01G125800	TraesCS1D01G127700	0.022088439	AGAMOUS MADS-box transcription factor
TraesCS6A01G192300	TraesCS1D01G127700	0.021992314	50S ribosomal protein L14
TraesCS6A01G192300	TraesCS1D01G127700	0.021992314	50S ribosomal protein L16
TraesCS1D01G295200	TraesCS1D01G127700	0.021929708	Photosystem II reaction center protein L
TraesCS4A01G074600	TraesCS1D01G127700	0.021860372	sulfate transporter 91
TraesCS7A01G164400	TraesCS1A01G125800	0.021752021	NADH dehydrogenase subunit 2
TraesCS7A01G260600	TraesCS1A01G125800	0.021746896	MADS-box transcription factor
TraesCS7A01G260600	TraesCS1A01G125800	0.021746896	MADS-box transcription factor
TraesCS5D01G220800	TraesCS1A01G125800	0.021723026	Photosystem II reaction center protein L
TraesCS6B01G208100	TraesCS1D01G127700	0.021673689	Photosystem I P700 chlorophyll a apoprotein A1

Appendix

Table S12: Co-express genes for only TaMADS13 (TraesCS5B01G115100, TraesCS5D01G118200 and TraesCS5A01G117500). Top 50 genes.

downstream.gene	TF	directed_edge_weight	Description
TraesCS6B01G052600	TraesCS5B01G115100	0.022255	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain-containing protein
TraesCS3A01G288600	TraesCS5D01G118200	0.021928	NPK1-related protein kinase 1
TraesCS6D01G264700	TraesCS5A01G117500	0.021471	Ubiquitin-conjugating enzyme E2
TraesCS1D01G308400	TraesCS5D01G118200	0.02135	Leucine-rich repeat receptor-like protein kinase family
TraesCS7A01G391400	TraesCS5D01G118200	0.02068	Alpha-L-fucosidase 2
TraesCS2B01G096400	TraesCS5A01G117500	0.020462	Tapetum determinant 1
TraesCS3B01G140600	TraesCS5D01G118200	0.020431	Eukaryotic aspartyl protease family protein
TraesCS2D01G052500	TraesCS5B01G115100	0.020189	FAD-binding Berberine family protein
TraesCS1B01G129100	TraesCS5B01G115100	0.019984	G-type lectin S-receptor-like Serine/Threonine-kinase
TraesCS6D01G264400	TraesCS5A01G117500	0.019678	Ubiquitin-conjugating enzyme E2
TraesCS1B01G479900	TraesCS5B01G115100	0.019315	Response regulator
TraesCS3B01G420700	TraesCS5A01G117500	0.01901	Amino acid permease
TraesCS3B01G170100	TraesCS5B01G115100	0.018848	Defensin
TraesCS5B01G115100	TraesCS5A01G117500	0.01884	AGAMOUS MADS box factor transcription factor
TraesCS5B01G115100	TraesCS5A01G117500	0.01884	AGAMOUS MADS box factor transcription factor
TraesCS1D01G308300	TraesCS5D01G118200	0.018795	Leucine-rich repeat receptor-like protein kinase family
TraesCS6B01G157900	TraesCS5B01G115100	0.018723	BURP domain RD22-like protein
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G100600	TraesCS5B01G115100	0.018483	BURP domain protein RD22
TraesCS6A01G416900	TraesCS5B01G115100	0.017773	F-box/RNI-like/FBD-like domains-containing protein
TraesCS7B01G362300	TraesCS5D01G118200	0.017637	Protein kinase-like protein
TraesCS7A01G536600	TraesCS5B01G115100	0.017587	ARM repeat superfamily protein
TraesCS3A01G141800	TraesCS5A01G117500	0.017567	Peroxidase
TraesCS6A01G158100	TraesCS5D01G118200	0.017245	MADS-box transcription factor
TraesCSU01G124200	TraesCS5D01G118200	0.017156	DNA (Cytosine-5-)-methyltransferase
TraesCS7D01G046600	TraesCS5B01G115100	0.017095	Alpha/beta-Hydrolases superfamily protein, putative
TraesCS5D01G118200	TraesCS5A01G117500	0.017065	AGAMOUS MADS box factor transcription factor
TraesCS5D01G118200	TraesCS5A01G117500	0.017065	AGAMOUS MADS box factor transcription factor
TraesCS4B01G333400	TraesCS5D01G118200	0.017012	carboxyl-terminal peptidase, putative (DUF239)
TraesCS5D01G013100	TraesCS5D01G118200	0.016961	Plasma membrane ATPase
TraesCS5D01G013100	TraesCS5D01G118200	0.016961	Plasma membrane ATPase
TraesCS3A01G164300	TraesCS5A01G117500	0.016902	MFS18 protein
TraesCS6B01G148600	TraesCS5D01G118200	0.016672	Retrovirus-related Pol polyprotein from transposon TNT 1-94
TraesCS6B01G052600	TraesCS5D01G118200	0.016597	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain-containing protein
TraesCS4D01G127100	TraesCS5D01G118200	0.01654	Alpha-L-arabinofuranosidase 1
TraesCS4D01G127100	TraesCS5D01G118200	0.01654	Alpha-L-arabinofuranosidase 1
TraesCS4D01G127100	TraesCS5D01G118200	0.01654	Alpha-L-arabinofuranosidase 1
TraesCS2B01G486800	TraesCS5D01G118200	0.016138	Glycine-rich cell wall structural protein
TraesCS2A01G120000	TraesCS5D01G118200	0.016086	Ypt/Rab-GAP domain of gyp1p superfamily protein
TraesCS2B01G580200	TraesCS5D01G118200	0.016028	Alpha-glucosidase
TraesCS7A01G460900	TraesCS5A01G117500	0.016022	Protein kinase-like protein
TraesCS1D01G308300	TraesCS5A01G117500	0.015977	Leucine-rich repeat receptor-like protein kinase family
TraesCS6D01G396200	TraesCS5B01G115100	0.015852	tRNA-dihydrouridine(47) synthase [NAD(P)(+)]
TraesCS5A01G117500	TraesCS5D01G118200	0.015829	AGAMOUS MADS box factor transcription factor
TraesCS5A01G117500	TraesCS5D01G118200	0.015829	AGAMOUS MADS box factor transcription factor
TraesCS5D01G071500	TraesCS5A01G117500	0.015775	WAT1-related protein

Appendix

Table S13: Co-express genes for only TaMADS21 (TraesCS1D01G262700, TraesCS1B01G273300 and TraesCS1A01G262700). Top 50 genes.

downstream.gene	TF	directed_edge_weight	Description
TraesCS5D01G103200	TraesCS1D01G262700	0.027931327	Beta-1,3-N-Acetylglucosaminyltransferase family protein, putative
TraesCS3D01G257800	TraesCS1B01G273300	0.026358606	Myb transcription factor
TraesCS1B01G334900	TraesCS1A01G262700	0.025415391	GDSL esterase/lipase
TraesCS5D01G103200	TraesCS1B01G273300	0.025160357	Beta-1,3-N-Acetylglucosaminyltransferase family protein, putative
TraesCS6A01G158100	TraesCS1A01G262700	0.023212104	MADS-box transcription factor
TraesCS3D01G468400	TraesCS1D01G262700	0.022585655	Myb transcription factor
TraesCS3D01G257800	TraesCS1D01G262700	0.022551185	Myb transcription factor
TraesCS5B01G034700	TraesCS1D01G262700	0.02246608	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
TraesCS3D01G318800	TraesCS1D01G262700	0.021865014	Pectinesterase
TraesCS6A01G320900	TraesCS1D01G262700	0.02157475	WAT1-related protein
TraesCS4D01G119000	TraesCS1B01G273300	0.021229732	Non-specific lipid-transfer protein
TraesCS5A01G117500	TraesCS1D01G262700	0.021176604	AGAMOUS MADS box factor transcription factor
TraesCS5A01G117500	TraesCS1D01G262700	0.021176604	AGAMOUS MADS box factor transcription factor
TraesCS4B01G075800	TraesCS1B01G273300	0.020957584	tapetum determinant 1
TraesCS4B01G075800	TraesCS1B01G273300	0.020957584	tapetum determinant 1
TraesCS2A01G543300	TraesCS1D01G262700	0.020914967	BTB/POZ/MATH-domain protein
TraesCS5A01G499200	TraesCS1B01G273300	0.020694721	Expansin-like protein
TraesCS5D01G118200	TraesCS1D01G262700	0.020566693	AGAMOUS MADS box factor transcription factor
TraesCS5D01G118200	TraesCS1D01G262700	0.020566693	AGAMOUS MADS box factor transcription factor
TraesCS7B01G345100	TraesCS1D01G262700	0.020040266	Ras family protein
TraesCS1B01G334900	TraesCS1D01G262700	0.019899552	GDSL esterase/lipase
TraesCS4B01G393200	TraesCS1A01G262700	0.01948406	Lipid transfer protein
TraesCS2A01G338200	TraesCS1B01G273300	0.019407502	MYB-related transcription factor
TraesCS2A01G060500	TraesCS1B01G273300	0.019314965	2-oxoglutarate-dependent dioxygenase-related family protein
TraesCS4B01G229500	TraesCS1A01G262700	0.019310194	Xyloglucan endotransglucosylase/hydrolase
TraesCS2A01G385900	TraesCS1D01G262700	0.019301898	Vacuolar-processing enzyme
TraesCS2A01G385900	TraesCS1D01G262700	0.019301898	Vacuolar-processing enzyme
TraesCS7A01G053200	TraesCS1D01G262700	0.019027686	eukaryotic translation initiation factor 2 gamma subunit
TraesCS1B01G113100	TraesCS1A01G262700	0.019006745	Transcription factor basic helix-loop-helix
TraesCS1B01G113100	TraesCS1A01G262700	0.019006745	Transcription factor basic helix-loop-helix
TraesCS1B01G113100	TraesCS1A01G262700	0.019006745	Transcription factor basic helix-loop-helix
TraesCS1B01G113100	TraesCS1A01G262700	0.019006745	Transcription factor basic helix-loop-helix
TraesCS1B01G113100	TraesCS1A01G262700	0.019006745	Transcription factor basic helix-loop-helix
TraesCS3D01G491700	TraesCS1B01G273300	0.018703088	Non-specific serine/threonine protein kinase
TraesCS7D01G522000	TraesCS1B01G273300	0.018547372	RING/U-box superfamily protein
TraesCS5A01G090900	TraesCS1B01G273300	0.018544156	Beta-1,3-N-Acetylglucosaminyltransferase family protein, putative
TraesCS7A01G314800	TraesCS1A01G262700	0.018322734	F-box/RNI-like/FBD-like domains-containing protein
TraesCS5B01G035800	TraesCS1B01G273300	0.018317969	Rhodopsin
TraesCS2D01G173000	TraesCS1B01G273300	0.018270807	Fasciclin-like arabinogalactan-protein-like
TraesCS5A01G034000	TraesCS1D01G262700	0.018234092	Potassium-transporting ATPase ATP-binding subunit
TraesCS2B01G075700	TraesCS1D01G262700	0.018203874	Proline-rich protein, putative, expressed
TraesCS3D01G468400	TraesCS1B01G273300	0.018166122	Myb transcription factor
TraesCS3D01G532800	TraesCS1B01G273300	0.018111815	WRKY transcription factor
TraesCS3A01G262900	TraesCS1A01G262700	0.018100836	70-kDa heat shock protein
TraesCS1B01G137100	TraesCS1D01G262700	0.017935843	Histone-lysine N-methyltransferase, H3 lysine-9 specific
TraesCS5B01G436700	TraesCS1D01G262700	0.017872417	Lipid transfer protein
TraesCS5B01G040500	TraesCS1B01G273300	0.01785133	Chloroplast ATP synthase delta subunit
TraesCS5B01G067300	TraesCS1D01G262700	0.017843085	WAT1-related protein
TraesCS2D01G383400	TraesCS1B01G273300	0.017610187	Vacuolar-processing enzyme
TraesCS7A01G551200	TraesCS1D01G262700	0.017601466	Ser/Thr-rich protein T10 in DGCR region

Appendix

TraesCS4B01G203700

TraesCS1D01G262700

0.017579301

basic helix-loop-helix (bHLH) DNA-binding superfamily protein

Table S14: UP-regulated genes list in only Pre-anthesis carpel sample.

Gene ID	logFC	PValue	Description
Bradi1g10567	15.12896	2.78E-07	PTHR11709//PTHR11709:SF86 - MULTI-COPPER OXIDASE // SUBFAMILY NOT NAMED (1 of 3)
Bradi4g44013	14.42651	1.18E-06	
Bradi3g00650	14.31994	3.92E-16	Glycosyl hydrolase (GH)
Bradi3g17997	13.16729	1.06E-38	KOG0157 - Cytochrome P450 CYP4/CYP19/CYP26 subfamilies (1 of 103)
Bradi2g39260	12.35251	2.92E-06	serine-type endopeptidase inhibitor activity (Blast2GO)
Bradi1g59440	11.89554	0.000145	PTHR30546//PTHR30546:SF4 - FLAVODOXIN-RELATED PROTEIN WRBA-RELATED // SUBFAMILY NOT NAMED (1 of 2)
Bradi1g77800	11.8908	1.41E-08	PTHR22814//PTHR22814:SF52 - COPPER TRANSPORT PROTEIN ATOX1-RELATED // SUBFAMILY NOT NAMED (1 of 3)
Bradi1g49200	11.45736	3.01E-50	IgE binding (Blast2GO)
Bradi1g47300	11.30894	7.77E-12	Cinnamoyl-CoA reductase (CCR)
Bradi2g37430	11.1717	1.87E-19	zinc ion binding (Blast2GO)
Bradi1g13040	11.06724	8.85E-09	globulin
Bradi1g70390	10.81292	4.18E-13	K17991 - peroxygenase (PXG) (1 of 8)
Bradi2g11320	10.68811	7.80E-25	1.11.1.7 - Peroxidase / Lactoperoxidase (1 of 144)
Bradi2g12880	10.64932	8.83E-07	acyltransferase activity (Blast2GO)
Bradi3g18027	10.57252	1.41E-05	KOG0157 - Cytochrome P450 CYP4/CYP19/CYP26 subfamilies (1 of 103)
Bradi3g36330	10.27681	2.67E-20	1.14.13.138 - Indolin-2-one monooxygenase / CYP71C2 (1 of 4)
Bradi4g23035	10.17226	2.63E-13	2.3.1.199 - Very-long-chain 3-oxoacyl-CoA synthase / Very-long-chain beta-ketoacyl-CoA synthase (1 of 29)
Bradi2g39272	10.07687	0.000982	PF00280 - Potato inhibitor I family (potato_inhibit) (1 of 31)
Bradi1g50927	9.81397	5.82E-09	PTHR22835//PTHR22835:SF134 - ZINC FINGER FYVE DOMAIN CONTAINING PROTEIN // SUBFAMILY NOT NAMED (1 of 2)
Bradi2g08300	9.720972	4.85E-26	aminopeptidase activity
Bradi1g61470	9.643961	3.09E-16	PTHR31621:SF1 - DUF679 DOMAIN MEMBRANE PROTEIN 1 (1 of 1)
Bradi1g62680	9.512404	4.42E-22	PF13668 - Ferritin-like domain (Ferritin_2) (1 of 3)
Bradi2g39265	9.452091	2.85E-05	trypsin inhibitor activity (Blast2GO)
Bradi2g47400	9.117347	1.66E-10	PGP-like phosphoglycoprotein auxin transporter
Bradi1g75120	9.083483	7.77E-26	PTHR31529:SF4 - LOB DOMAIN-CONTAINING PROTEIN 14-RELATED (1 of 1)
Bradi2g45030	8.858101	6.18E-07	
Bradi1g36750	8.68937	0.000221	K16241 - transcription factor HY5 (HY5) (1 of 3)
Bradi3g29310	8.676994	2.31E-09	PTHR24296:SF3 - CYTOCHROME P450 86B1-RELATED (1 of 1)
Bradi4g09257	8.479479	9.68E-15	PF01439 - Metallothionein (Metallothio_2) (1 of 7)
Bradi2g50480	8.473732	1.20E-21	cysteine-type endopeptidase activity (Blast2GO)
Bradi4g14620	8.459805	6.90E-14	PF04398 - Protein of unknown function
Bradi1g59450	8.439332	0.000433	PTHR30546//PTHR30546:SF4 - FLAVODOXIN-RELATED PROTEIN WRBA-RELATED // SUBFAMILY NOT NAMED (1 of 2)
Bradi1g61540	8.401381	1.42E-06	1.11.1.7 - Peroxidase / Lactoperoxidase (1 of 144)

Appendix

Bradi5g11690	8.346526	1.67E-10	aspartic-type endopeptidase activity (Blast2GO)
Bradi1g63616	7.816218	6.27E-51	
Bradi3g38169	7.719833	6.98E-09	
Bradi2g16261	7.549793	8.17E-09	
Bradi2g27430	7.406456	1.69E-06	PTHR10992:SF56 - METHYLESTERASE 4-RELATED (1 of 2)
Bradi2g30390	7.286458	7.95E-11	PTHR11615//PTHR11615:SF144 - NITRATE
Bradi5g06660	7.234381	4.42E-11	cysteine protease inhibitor activity (Blast2GO)
Bradi2g18930	7.216234	5.43E-07	PF04520 - Senescence regulator (Senescence_reg) (1 of 17)
Bradi2g42570	7.179679	2.54E-10	PTHR10992//PTHR10992:SF841 - ALPHA/BETA HYDROLASE FOLD-CONTAINING PROTEIN // SUBFAMILY NOT NAMED (1 of 1)
Bradi2g05580	7.015079	9.58E-06	FAD binding (Blast2GO)
Bradi1g68850	6.83129	9.54E-10	dipeptide transporter activity
Bradi1g66330	6.821815	2.48E-07	zinc ion binding (Blast2GO)
Bradi3g15740	6.791721	3.45E-05	PF02458 - Transferase family (Transferase) (1 of 99)
Bradi2g21951	6.736054	7.35E-30	
Bradi5g03632	6.701569	5.48E-25	PF07714//PF13947 - Protein tyrosine kinase (Pkinase_Tyr) // Wall-associated receptor kinase galacturonan-binding (GUB_WAK_bind) (1 of 20)
Bradi1g49190	6.650023	4.73E-44	PTHR22835//PTHR22835:SF252 - ZINC FINGER FYVE DOMAIN CONTAINING PROTEIN // SUBFAMILY NOT NAMED (1 of 3)
Bradi3g21560	6.587011	9.14E-08	serine-type carboxypeptidase activity (Blast2GO)
Bradi2g41817	6.505232	0.000173	
Bradi2g56037	6.427918	2.28E-05	PTHR12231:SF112 - RHODANESE-LIKE DOMAIN-CONTAINING PROTEIN 14
Bradi1g63720	6.343836	6.42E-09	RING
Bradi2g34930	6.335165	0.00148	hydrogen-translocating pyrophosphatase activity
Bradi2g57340	6.299356	5.49E-05	
Bradi3g36347	6.281134	1.83E-11	1.14.13.139 - 3-hydroxyindolin-2-one monooxygenase / CYP71C1 (1 of 2)
Bradi2g11420	6.204314	0.000507	PF00892 - EamA-like transporter family (EamA) (1 of 57)
Bradi2g27140	6.001919	0.001261	Glycosyl hydrolase (GH)
Bradi5g11900	5.869603	5.63E-10	
Bradi4g43860	5.848971	2.24E-07	iron ion binding (Blast2GO)
Bradi3g19240	5.815227	2.40E-13	KOG0157 - Cytochrome P450 CYP4/CYP19/CYP26 subfamilies (1 of 103)
Bradi1g06030	5.783246	3.09E-13	arachidonic acid binding (Blast2GO)
Bradi1g71280	5.779389	3.21E-07	PF00046 - Homeobox domain (Homeobox) (1 of 64)
Bradi5g15418	5.757816	7.97E-08	K17991 - peroxygenase (PXG) (1 of 8)
Bradi4g30750	5.736902	0.00014	PTHR11863//PTHR11863:SF46 - STEROL DESATURASE // SUBFAMILY NOT NAMED (1 of 1)
Bradi3g50250	5.659545	1.08E-26	aldo-keto reductase activity
Bradi1g01150	5.65815	0.001325	Glycosyl hydrolase (GH)
Bradi1g46300	5.557769	1.44E-12	
Bradi1g37870	5.491563	0.000115	aldo-keto reductase activity
Bradi2g56650	5.49088	2.55E-09	F-Box

Table S15: Down-regulated genes list in the Pre-anthesis carpel and their fold-change.

Gene ID	logFC	PValue	Description
BRADI_5g26380	-6.058	1.06E-13	
BRADI_3g08301	-5.343	7.27E-43	PTHR24031//PTHR24031:SF259 - RNA HELICASE // SUBFAMILY NOT NAMED
BRADI_5g01574	-4.170	9.10E-05	
BRADI_3g47290	-3.958	9.98E-17	
BRADI_3g28100	-3.836	5.00E-18	PTHR22835:SF160 - GDSL ESTERASE/LIPASE LTL1 (1 of 1)
BRADI_5g21770	-3.806	1.31E-05	histidine decarboxylase activity aromatic-L-amino-acid decarboxylase activity
BRADI_4g13987	-3.413	3.65E-35	
BRADI_1g12900	-3.411	1.18E-06	THX transcription factor
BRADI_1g37742	-3.297	0.000794	PF12023 - Domain of unknown function (DUF3511) (DUF3511) (1 of 8) sulfite oxidase activity NADPH-hemoprotein reductase activity L-lactate dehydrogenase (cytochrome) activity stearyl-CoA 9-desatur
BRADI_2g23350	-3.274	6.03E-07	di- tri-valent inorganic cation transmembrane transporter activity (Blast2GO)
BRADI_2g40632	-3.161	4.41E-06	
BRADI_4g30523	-3.115	1.77E-09	PF12874 - Zinc-finger of C2H2 type (zf-met) (1 of 15)
BRADI_1g23450	-3.016	5.63E-06	D-glucose transmembrane transporter activity (Blast2GO)
BRADI_2g18050	-2.914	1.26E-25	Glycosyl hydrolase (GH) subfamily GH28
BRADI_1g21455	-2.901	0.000699	
BRADI_1g68102	-2.818	5.20E-06	PTHR22904//PTHR22904:SF326 - TPR REPEAT CONTAINING PROTEIN // SUBFAMILY NOT NAMED (1 of 4)
BRADI_3g50050	-2.774	1.59E-09	YABBY transcription factor
BRADI_2g17300	-2.738	9.92E-08	
BRADI_1g63370	-2.687	1.84E-08	PF01190 - Pollen proteins Ole e l like (Pollen_Ole_e_l) (1 of 27)
BRADI_1g00770	-2.642	0.000992	copper-exporting ATPase activity (Blast2GO)
BRADI_5g08740	-2.613	5.93E-07	protein disulfide oxidoreductase activity (Blast2GO)
BRADI_1g17910	-2.606	0.000416	
BRADI_3g49440	-2.548	5.58E-12	3' 5'-cyclic-GMP phosphodiesterase activity 3' 5'-cyclic-AMP phosphodiesterase activity (Blast2GO)
BRADI_1g77710	-2.542	5.55E-07	nucleoside-triphosphatase activity ATP binding (Blast2GO)
BRADI_4g17900	-2.520	4.95E-19	serine-type carboxypeptidase activity (Blast2GO)
BRADI_2g37650	-2.403	1.28E-08	PTHR24326:SF115 - WUSCHEL-RELATED HOMEODOMAIN 3 (1 of 2)
BRADI_1g51910	-2.383	0.001484	PF00892 - EamA-like transporter family (EamA) (1 of 57)
Bradi4g04600	-2.375	6.56E-06	
Bradi2g22380	-2.365	1.85E-07	
Bradi4g38965	-2.363	1.97E-05	PTHR10641//PTHR10641:SF506 - MYB-LIKE DNA-BINDING PROTEIN MYB
Bradi3g54935	-2.264	4.13E-06	
Bradi3g38990	-2.263	1.15E-16	single-stranded DNA binding (Blast2GO)
Bradi2g38552	-2.259	5.70E-10	
Bradi2g07820	-2.230	3.07E-07	extra sex combs like (ESCL)
Bradi3g27430	-2.223	1.94E-05	RING subfamily zinc finger C3HC4 type domain containing protein
Bradi5g01546	-2.169	2.03E-07	PTHR23155:SF402 - DISEASE RESISTANCE PROTEIN RPP13-RELATED (1 of 16)
Bradi3g38890	-2.168	1.22E-05	bHLH transcription factor
Bradi3g41510	-2.068	7.32E-06	bHLH transcription factor
Bradi1g61690	-2.056	7.90E-06	protein kinase family protein putative expressed subfamily RLCK-VIIa
Bradi1g33001	-2.009	0.001374	
Bradi1g28726	-2.001	0.000102	PTHR26402//PTHR26402:SF544 - RESPONSE REGULATOR OF TWO-COMPONENT SYSTEM
Bradi1g46340	-2.001	0.000113	zinc ion binding (Blast2GO)