Transcriptome Profiling of Triticale Reproductive Tissues

by

Frances Tran

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Master of Science

in

Biology

Carleton University
Ottawa, Ontario

© 2011

Frances Tran

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-81681-3
Our file *Notre référence*
ISBN: 978-0-494-81681-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Successful plant reproduction in cereal crops involves a variety of highly specialized processes including male and female gametophyte production and pollen-pistil interactions such as pollen adhesion, recognition, hydration and germination, pollen tube development, and fertilization. Using the Affymetrix 55K Wheat GeneChip[®] microarray, we have generated transcriptome profiles for each triticales reproductive organ (anther, ovary, stigma) at four consecutive stages of development, as well as for mature pollen grain. The triticales reproductive transcriptomes varied considerably within each organ over the course of development showing large fluctuations in the number of differentially expressed genes. Globally, distinct trends were shown among the different male and female tissues with more than 3,000 genes showing preferred expression in one of the tissues and nearly one third showed tissue-specific expression. In total, we identified 778 probe sets specific to the anther or mature pollen, 112 in the ovary and 117 for the stigma. Functional categorization of these genes were in line with biological activities within the male and female tissues reflecting their role in pollen cell wall synthesis, maturation, and germination as well as ovary macromolecule biosynthesis and stigma cuticular wax production and organogenesis. Nearly half of these genes had unknown functions, many of which are novel and require further investigation. Future functional analysis of the genes expressing at different developmental stages and among the different organs will undoubtedly provide further insight into the molecular mechanisms governing plant reproduction.

Acknowledgements

I would first like to express my gratitude to my supervisor Dr. Laurian Robert for giving me the opportunity to work in his lab. I have learned a tremendous amount from him and have become a confident researcher as a result of his wonderful guidance, leadership, encouragement and especially patience. I would also like to thank my other brilliant supervisors Dr. Owen Rowland and Dr. Linda Bonen for their great advice and continued support throughout the entire process. To my wonderful lab members: Madeleine Lévesque-Lemay, Mohsin Zaidi, John Chan, Nazila Nazemof, and Zongbin Wu whom were all a pleasure to work with, and made the time at the lab very enjoyable. I would also like to extend my appreciation to Ann-Fook Yang for donating his time with the SEM, Doug Miller for helping with the tissue collection, and to the bioinformatics group: Carolyn Penniket, André Laroche, Nick Tinker, Youlian Pan, Yong Xu and Philippe Couroux, all your help, input, and ability to solve different types of bioinformatic-related issues is greatly appreciated. Lastly, thank you to my close friends and family for their continued love and support.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
List of Tables.....	ix
List of Figures.....	x
List of Abbreviations.....	xiii

1. Introduction

1.1 Triticale- a crop with great potential.....	1
1.2 Anther development.....	3
1.3 Pollen development and fertilization.....	6
1.4 Pistil development.....	8
1.5 Pollen-pistil interaction.....	12
1.6 Characterization of the anther and mature pollen transcriptome.....	15
1.7 Characterization of the pistil transcriptome.....	22
1.8 Objectives.....	26

2. Materials and Methods

2.1 Plant material and growth conditions.....	28
2.2 Microscopy.....	28
2.3 RNA isolation and Affymetrix GeneChip® hybridization.....	30
2.4 Data analysis and experimental design.....	34
2.5 Gene ontology.....	40

3. Quality Control and Data Statistical Analysis

3.1 Results.....	42
------------------	----

3.1.1 Pearson correlation coefficient.....	42
3.1.2 Expression value and detection P-value call.....	42
4. Transcriptome Analysis of Anther Development and Mature Pollen	
4.1 Results.....	47
4.1.1 Microscopic analysis of triticales anther and pollen development.....	47
4.1.2 Expression profiling throughout anther development and in mature pollen.....	49
4.1.3 Co-regulated gene expression in the developing anther and mature pollen.....	54
4.1.4 Identification of differentially expressed genes in the developing anther and mature pollen.....	56
4.1.5 Transcripts enriched and specific to the developing anther and mature pollen.....	60
4.1.6 Characterization of anther- and pollen-specific genes.....	65
4.2 Discussion.....	76
4.2.1 Anther and pollen demonstrate distinctive gene expression patterns.....	76
4.2.2 Anther-only gene expression.....	81
4.2.3 Anther-only transport proteins.....	82
4.2.4 Anther-only protein metabolism.....	83
4.2.5 Anther-only cell wall biogenesis and cell wall modification.....	84
4.2.6 Anther-only transcription factors.....	85
4.2.7 Anther-only protein phosphorylation.....	87
4.2.8 Anther + pollen expression.....	88

4.2.9 Anther + pollen cell wall modification and carbohydrate metabolism.....	88
4.2.10 Anther + pollen metabolism.....	89
4.2.11 Anther + pollen regulatory factors.....	89
4.2.12 Mature pollen-only gene expression.....	90
4.2.13 Mature pollen-only metabolism.....	91
4.2.14 Mature pollen-only defense and stress response.....	91
4.3 Conclusion.....	92

5. Transcriptome Analysis of Triticale Ovary Development

5.1 Results.....	93
5.1.1 Microscopic analysis of ovary development.....	93
5.1.2 Gene expression profiling throughout ovary development...	93
5.1.3 Co-regulated gene expression in the developing ovary.....	99
5.1.4 Identification of differentially expressed genes in the developing ovary.....	102
5.1.5 Transcripts enriched and specific to the developing ovary....	104
5.1.6 Characterization of the ovary-specific genes.....	106
5.2 Discussion	111
5.2.1 Developmental changes in the ovary transcriptome.....	111
5.2.2 Ovary-specific phosphoregulation.....	114
5.2.3 Ovary-specific proteolysis.....	115
5.2.4 Ovary-specific transcription factor.....	116
5.2.5 Ovary-specific genes.....	117
5.3 Conclusion.....	118

6. Transcriptome Analysis of Triticale Stigma Development

6.1 Results.....	119
6.1.1 Microscopic analysis of stigma development.....	119
6.1.2 Gene expression profiling throughout stigma development..	121
6.1.3 Co-regulated gene expression in the developing stigma.....	125
6.1.4 Identification of differentially expressed genes in the developing stigma.....	128
6.1.5 Transcripts enriched and specific to the developing stigma..	132
6.1.6 Characterization of the stigma-specific genes.....	134
6.2 Discussion	137
6.2.1 Developmental changes in the stigma transcriptome.....	137
6.2.2 Stigma-specific transcription factors.....	140
6.2.3 Stigma-specific lipid metabolism.....	141
6.2.4 Stigma-specific signaling proteins.....	142
6.2.5 Stigma-specific defense response.....	143
6.3 Conclusion.....	143

7. Global Transcriptome Analysis of Triticale Reproductive Development

7.1 Results.....	145
7.1.1 Gene expression across reproductive tissues.....	145
7.1.2 Co-regulation of gene expression within male and female tissues.....	145
7.1.3 Differentially expressed probe sets show similar tendencies in reproductive tissues.....	156
7.1.4 Distribution of enriched and specific transcripts throughout reproductive development.....	160

7.1.5 Reproductive tissue-specific genes.....	160
7.2 Discussion.....	164
7.2.1 Male and female reproductive tissues display convergent and divergent expression patterns throughout development.	164
7.2.2 Reproductive tissue-specific genes.....	167
7.3 Conclusion.....	168
8. References.....	169

List of Tables

Table 3-1 Pearson correlation coefficient summary among biological replicates for different microarray experiments.....	43
Table 3-2 Number of expressed probe sets for each biological replicate.....	44
Table 3-3 Summary of ‘Present’ calls at each stage of reproductive development.....	46
Table 4-1 Top 50 candidate anther-only specific probe sets.....	69
Table 4-2 Top 50 candidate anther + pollen specific probe sets.....	73
Table 4-3 Top 50 candidate pollen-only specific probe sets.....	75
Table 5-1 Top 50 candidate ovary-specific probe sets.....	110
Table 6-1 Top 50 candidate stigma-specific probe sets.....	136
Table 7-1 Summary of the distribution of reproductive tissue-specific probe sets.....	163

List of Figures

Figure 1-1 Schematic diagram of the grass flower.....	4
Figure 1-2 Transverse section of a wheat anther at meiosis.....	5
Figure 1-3 Summary of pollen development in wheat anthers including microsporogenesis and microgametogenesis.....	7
Figure 1-4 Longitudinal section of a wheat (<i>Triticum aestivum</i>) pistil showing pollen tube growth into the embryo sac.....	9
Figure 1-5 Structure of an anatropous ovule.....	11
Figure 1-6 The pollen-stigma interaction.....	13
Figure 2-1 Experimental design and workflow summary for Affymetrix wheat GeneChip® microarray analysis.....	36
Figure 4-1 Microscopy of triticales reproductive tissues at four stages of pollen development.....	48
Figure 4-2 Schematic diagram of microarray analysis for anther development and mature pollen grain.....	50
Figure 4-3 Number of expressed probe sets at each stage of anther development and mature pollen.....	51
Figure 4-4 Dartboard model of anther development and mature pollen.....	53
Figure 4-5 Scatter plots comparing co-expression between anther developmental stages and/or mature pollen.....	55
Figure 4-6 Cluster analysis of anther and pollen expressed probe sets.....	57
Figure 4-7 Differentially up-regulated anther and pollen probe sets.....	59
Figure 4-8 Differentially down-regulated anther and pollen probe sets.....	61
Figure 4-9 Anther and pollen-expressed, -enriched, -stage-enriched, -specific, and -stage-specific probe sets.....	63
Figure 4-10 Distribution of biological process for anther-only specific probe sets.....	66

Figure 4-11 Distribution of biological process for anther + pollen specific probe sets.....	71
Figure 4-12 Distribution of biological process for pollen-specific probe sets.....	74
Figure 5-1 Microscopy of triticales reproductive tissues at four stages of pollen development.....	94
Figure 5-2 Schematic diagram of microarray analysis for ovary development.....	95
Figure 5-3 Number of expressed probe sets at each stage of ovary development...	96
Figure 5-4 Dartboard model of ovary development.....	98
Figure 5-5 Scatter plots comparing co-expression between ovary developmental stages.....	100
Figure 5-6 Cluster analysis of stigma expressed probe sets.....	101
Figure 5-7 Differentially up-regulated ovary probe sets.....	103
Figure 5-8 Differentially down-regulated ovary probe sets.....	105
Figure 5-9 Ovary-expressed, -enriched, -stage-enriched, -specific, and -stage-specific probe sets.....	107
Figure 5-10 Distribution of biological process for ovary-specific probe sets.....	109
Figure 6-1 Microscopy of triticales reproductive tissues at four stages of pollen development.....	120
Figure 6-2 Schematic diagram of microarray analysis for stigma development.....	122
Figure 6-3 Number of expressed probe sets at each stage of stigma Development.....	123
Figure 6-4 Dartboard model of stigma development.....	124
Figure 6-5 Scatter plots comparing co-expression between stigma developmental stages.....	126
Figure 6-6 Cluster analysis of stigma expressed probe sets.....	127
Figure 6-7 Differentially up-regulated stigma probe sets.....	129
Figure 6-8 Differentially down-regulated stigma probe sets.....	131

Figure 6-9 Stigma-expressed, -enriched, -stage-enriched, -specific, and –stage-specific probe sets.....	133
Figure 6-10 Distribution of biological process for stigma-specific probe sets.....	135
Figure 7-1 Number of expressed probe sets at each stage of reproductive development.....	146
Figure 7-2 Dartboard model comparisons of reproductive development.....	148
Figure 7-3 Pearson correlation coefficient of co-expressed probesets during different stages of reproductive development.....	150
Figure 7-4 Venn diagram between different reproductive tissues.....	152
Figure 7-5 Venn diagram at different stages of reproductive development.....	154
Figure 7-6 Principal component analysis of reproductive development.....	155
Figure 7-7 Comparison between differentially up-regulated probe sets in reproductive development.....	157
Figure 7-8 Comparison between differentially down-regulated probe sets In reproductive development.....	159
Figure 7-9 Comparison of tissue- enriched, -stage-enriched, -specific, and –stage-specific probe sets throughout reproductive development.....	161

List Of Abbreviations

BCP: Bi-cellular pollen
bHLH: Basic helix-loop-helix
CER: Eceriferum
DEPs: Differentially expressed probe sets
DNA: Deoxyribonucleic acid
EST: Expressed sequence tag
EXS/EMS1: Extra sporogenous cells/excess microsporocytes1
GO: Gene ontology
Gp: Germinating pore
GPG: Germinated pollen grain
GUS: *β -glucuronidase*
LM: Laser-microdissection
LTP: Lipid transfer protein
LRR-RK: Leucine rich repeat-receptor kinase
MMC: Microspore mother cell
MPG: Mature pollen grain
MSP1: Multiple sporocyte1
RNA: Ribonucleic acid
PMC: Pollen mother cell
PM I: Pollen mitosis I
PM II: Pollen mitosis II
SEM: Scanning electron microscope
TCP: Tri-cellular pollen
Udt 1: Undeveloped tapetum 1
UNM: Uninucleate microspore

1. Introduction

Cereal grasses such as rice, maize, and wheat are members of the monocot family Poaceae and contribute the largest production of staple food grain worldwide. Cereal reproduction is a critical process not only because it provides the seeds required to grow more crops, but also because it ultimately yields the grain that serves as food and feed globally. In addition to generating the male and female gametophytes, reproduction in cereals involves a variety of highly specialized processes including pollen-pistil interaction, pollen germination and tube development, and fertilization (Swanson et al., 2004). Failure or inefficiency of any of these reproductive processes will have a negative impact on crop production. A thorough understanding of the molecular mechanisms underlying plant reproduction is crucial for efficient crop fertility management whether aimed at preventing transgene flow, sustaining production under changing climate conditions, or for practical purposes such as hybrid seed production.

1.1 Triticale – a crop with great potential

Triticale (*xTriticosecale* Wittmack) is a human made hybrid crop that represents an interesting biological system since it is made up of three genomes from two different Triticeae members of the grass family, wheat (*Triticum*) and rye (*Secale*), which are both relevant to crop production. Hexaploid triticale ($2n = 6x = 42$) contains the wheat AABB genome and rye RR genome, and based on next-generation sequencing data approximately 50,000 genes have been reported thus far (Xu et al., 2011). Triticale is a largely self-pollinating plant that is currently being produced in over 28 countries including Canada (FAO, 2009) and is of particular interest because it was bred to contain

favourable characteristics derived from both parents. The interaction between the two parental genomes have resulted in induced gene silencing and genomic sequence elimination targeting mainly the rye genome due to transcriptional interference (Ma et al., 2004; Kashkush et al., 2003).

Triticale has the capability to adapt to harsh environmental conditions such as high elevation, drought, and poor soil conditions that include sandy, saline, acidic, or showing metal toxicity. As a livestock feed it also contains higher protein content relative to other grains (Coffey & Gerrits, 1988). Therefore, it is useful for growth in a wide range of terrains as well as a more nutritional and cost-effective alternative feedstock (for review see Oettler, 2005). Furthermore, it has good disease resistance making it a valuable genetic resource (Kuleung et al., 2004). Recently, triticale has been proposed as a bio-industrial platform in Canada (<http://ctbi.ca/>) for the production of various biopolymers (e.g. bioplastics, BioSteel), adhesives, industrial enzymes, and by-products such as ethanol, thus opening a completely new range of possibilities for its use.

Triticale research, as with wheat, can be challenging due to the complexity of polyploid genomes and also because of the difficulty to generate transgenic plants. However, the recent sequencing of the wheat genome will help facilitate genetic characterization (<http://www.wheatgenome.org/>). Furthermore, triticale transformation is now feasible (Greer et al., 2009) and transforming *Brachypodium distachyon*, the model system for monocots with its recently sequenced genome, is currently a possible alternative to study Triticeae gene function (Alves et al., 2009; Vogel & Hill, 2008). The adoption in the field of bio-factories such as transgenic triticale may require mechanisms to keep these modified cereal plants genetically isolated from traditionally bred varieties.

This is another reason to understand the fundamental processes involved in its reproductive development.

1.2 Anther development

Reproductive structures are evolutionary diverse and vary across plant species, but are well conserved among members of the Triticeae. In triticale, the stamen is the male reproductive organ consisting of a single-veined filament connected to a four-loculed anther in which pollen develops (Figure 1-1). The stalk-like filament not only anchors the stamen to the flower, but also aids in transmitting water and nutrients to the anther and in positioning the organ at the time of pollen dispersal (Goldberg et al., 1993). Anther development can be divided in two phases. Early in development, the first phase involves arrangement of the different anther cell layers and the completion of meiosis of the pollen mother cells (PMCs). The second phase involves microspore development and anther dehiscence releasing the male gametophyte, the mature pollen (McCormick, 2004). Anther and pollen development are intimately linked.

Anther tissue consists of three germ layers: L1, L2, and L3, all with divergent cell differentiation fates. The cell fate of the different layers produces four concentric rings surrounding the developing sporogenous cells (Figure 1-2). The lineage of the L1 layer forms the epidermis and stonium, while L2 produces archesporial cells that give rise to sporogenous and parietal cells (Goldberg et al., 1993). The sporogenous cells undergo several generations of division leading to free microspores from which mature pollen develops. In the case of the parietal cells, they divide periclinally to give rise to the endothecium, middle wall, and outer tapetum layers (Ma et al., 2007). Lastly, L3 differentiates into the inner tapetum and vascular bundle (Goldberg et al., 1993).

Figure 1-1. Schematic diagram of the grass flower A) Partly dissected grass flower at anthesis, B) longitudinal, and C) transverse diagram of the flower, and D) spikelet (Johnson, 1931)

Figure 1-2. Transverse section of a wheat anther at meiosis. Close up of one lobe showing four cell lineages derived from L1, L2, and L3 anther germ layers, and starch grains (arrowhead). Ep, epidermis; En, endothecium; ML, middle layer; Ta, tapetum; Me, meiocyte. Scale Bar = 20 μ m. (Lalonde et al., 1997, copyright © www.springerlink.com)

1.3 Pollen development and fertilization

The development of a functional pollen grain is divided into two phases: microsporogenesis and microgametogenesis as shown in Figure 1-3. The sporogenous cells inside the anther multiply mitotically to form microspore mother cells (MMC's) which subsequently enter meiotic division. In triticales anthers, meiosis begins at the base of the pollen sac and proceeds towards the tip (Bennett et al., 1971).

At the onset of meiosis, the MMC begins to secrete callose, a secondary wall material composed of β (1-3) glucan polymers which accumulates around the forming tetrads (Figure 1-3A). The callose wall allows nutrients, such as soluble carbohydrates including sucrose, to reach the developing microspores (Barskaya & Balina, 1971). At the end of microsporogenesis a tetrad of microspores is produced (Koonjul et al., 2005). Microgametogenesis begins when microspores are freed from each tetrad when the callose wall is degraded by callase, an enzyme produced within a cell layer lining the inside of the anther known as the tapetum (Teng et al., 2005). At this stage, each microspore contains a prominent vacuole and is in direct contact with the tapetum via the germinating pore (Koonjul et al., 2005; Christensen & Horner, 1974). The microspore wall begins to form as tapetum degeneration is initiated. During microgametogenesis, a unicellular microspore undergoes two asymmetric mitotic divisions. Pollen mitosis I (PM I) produces bi-cellular pollen comprised of a vegetative and generative cell. At this point, the pollen grains begin accumulating starch while rapid pollen wall formation takes place (Koonjul et al., 2005). After pollen mitosis II (PM II), a three-celled body is

Figure 1-3. Summary of pollen development in wheat anthers including microsporogenesis and microgametogenesis. (A) Microsporocytes during microsporogenesis undergo two nuclear meiotic divisions producing a tetrad of four haploid microspores. (B-E) Microgametogenesis, involving two mitotic divisions to produce tri-cellular pollen in angiosperms. Aw, anther wall; Ca, callose; En, endothecium; Ep, epidermis; Ex, exine; Gn, generative nucleus; In, intine; Inf, inflorescence; Mi, microspore; MI, middle layer; MMC, microspore mother cell; Mn, microspore nucleus; Op, operculum; Po, pore; Sp, sperm; Tp, tapetum; Va, vacuole; Vn, vegetative nucleus. Modified from (Koonjul et al., 2005, copyright <http://jxb.oxfordjournals.org/>)

produced comprised of two sperm cells derived from the generative cell, and the remaining larger vegetative cell which will eventually form the pollen tube (McCormick, 2004). Lastly, the pollen coat is formed, most of the pollen grain cytoplasm is occupied by starch and it is ready for shedding (Koonjul et al., 2005). Triticale produces tricellular pollen for double fertilization (Lersten, 2004). One sperm cell combines with the egg cell, while the second cell is required to initiate endosperm formation, an essential tissue to angiosperms (Dresselhaus, 2006). The pollen grain is released in a partially desiccated state and begins to hydrate upon landing on a compatible stigma. A successful pollen-stigma interaction leads to a germinating pollen tube that elongates and penetrates through the style and ovary and eventually deposits the sperm cells into the embryo sac in which fertilization occurs. In triticale, fertilization was observed to occur within 40-60 minutes after pollination (Bennett et al., 1975).

1.4 Pistil development

The pistil is the female counterpart involved in plant reproduction and is made up of a stigma, style, and ovary (Figure 1-4). In grass species, the stigma is a double branched plumose structure located at the uppermost part of the pistil. Secondary receptive multiseriate hairs called papillae protrude from the main branches (Ciampolini et al., 2001). As in all dry stigma types, the papilla is covered by a thin discontinuous proteinaceous waxy cuticle layer called the pellicle (Vishnyakova & Willemse, 1994). An intact cutin pellicle layer is essential for successful pollen-stigma interactions.

Figure 1-4. Longitudinal section of a wheat (*Triticum aestivum*) pistil showing pollen tube growth into the embryo sac. Modified from (Vishnyakova & Willemse, 1994).

Removal of the pellicle with proteases and/or detergents can prevent pollen tube entry into the stigma, which suggests a role in modulating water passage from the papillae for pollen hydration (Heslop-Harrison & Heslop-Harrison, 1975). The components that make up the pellicle are poorly understood however, based on the limited studies performed, esterases and glycoproteins seem to be constituents (Heslop-Harrison & Heslop-Harrison, 1980; Vishnyakova & Willemse, 1994). The cellular functions of proteins involved in the pollen-stigma interaction are still unknown and in fact the stigma itself remains poorly understood.

Once the pollen has germinated on the stigma, the pollen tube grows through the style to reach the ovary which contains the female gametophyte and is the precursor of the seed. The ovule is nested in the ovary and develops from the placenta. It comprises two concentric outer layers of integuments, the nucellus, the embryo sac (female gametophyte), and the funiculus stalk which attaches the ovule to the placenta as shown in Figure 1-5 (Esau, 1965). The nucellus is encased by two layers of integuments except at the apex where a small opening is formed known as the micropyle. The micropyle allows the pollen tube to enter the ovule and into the embryo sac where fertilization of the egg cell ultimately occurs. The embryo sac which arose from the megasporangium via meiosis enters three rounds of mitosis to divide into eight nuclei which include the egg cell and a binucleate central cell that eventually become the zygote and triploid endosperm respectively (Robinson-Beers et al., 1992). The two synergid cells on either side of the egg cell are believed to attract and guide the pollen tube (Chaubal & Reger, 1990), similar roles have been attributed to the integuments as well (Palanivelu et al., 2003). Ovule development have been well reviewed in Arabidopsis

Figure 1-5. Structure of an anatropous ovule (modified from www.2classnotes.com)

(Robinson-Beers et al., 1992; Schneitz et al., 1995; Berger et al., 2008), yet the structures are nonetheless highly conserved among flowering plants.

1.5 Pollen-pistil interaction

The initial interaction between the pollen grain and the stigma portion of the pistil is the interface where the recognition and signals regulating pollen germination occurs and ultimately, leads to fertilization. Presently, little is known about the molecular mechanisms involved in the pollen-stigma interaction and this is especially true in cereals. Understanding this interaction is essential as it is recognized as the main barrier to fertilization in cereals. The major biological events that occur in this interaction include: pollen adhesion, recognition, hydration, germination, and pollen tube penetration through the stigma as shown in Figure 1-6 (Heslop-Harrison & Heslop-Harrison, 1980; 1981).

Virtually nothing is known of the initial recognition phase of pollen/stigma interactions in cereals, although self-incompatibility responses have been investigated (Langridge & Bauman, 2008). On the other hand, pollen adhesion was shown to be largely dependent on the pollen exine wall in *Arabidopsis*. Supporting evidence was shown with the *Arabidopsis ECERIFERUM* (*CER*) mutant *cer 6-2* that lacks a pollen coat but can still adhere to the stigma surface, indicating adhesive properties present in the sporopollenin exine wall (Zinkl et al., 1999). In the *Brassicaceae*, following the initial pollen capture, the lipidic pollen coat flows from the exine cavities onto the papilla surface forming the so-called ‘foot’ layer (Figure 1-6B). Important protein-protein interactions are believed to occur within the foot layer. Although such a foot layer has

Figure 1-6. The pollen-stigma interaction. (A) Events of the pollen-stigma interaction include: 1) pollen adhesion, 2) pollen hydration, 3) pollen tube germination, and 4) pollen tube penetration of the papillae and style. (B) Transmission electron micrograph of pollen-stigma interaction in *Arabidopsis*, arrows indicating the foot layer. (Edlund et al., 2004; www.plantcell.org, Copyright American Society of Plant Biologists)

not been reported in the grasses, a thick mucilaginous-like material including stigmatic surface secretions and pollen exudates has been observed (Heslop-Harrison & Heslop-Harrison, 1981). Esterase activity have also been found to localize at the site of contact between the pollen exine outer layer and the stigma pellicle providing further insights as to how the pollen is able to adhere to the stigma surface. Earlier work attempting to isolate and characterize proteins in the rye stigma pellicle by iso-electric focusing also revealed the presence of esterase isozymes, as well as glycoproteins (Heslop-Harrison, 1978). The molecular constituents and the chemical interactions that occur between the pollen and stigma upon initial contact, however, still remain unclear.

At dehiscence, pollen grains from rye release from the anthers in a desiccated state dropping down to 15-30% water content (Heslop-Harrison, 1979a). Activation of the grain is dependent on the inflow of water from the stigma for pollen rehydration (Heslop-Harrison, 1978). Pollen hydration will cause polarization and activate metabolism in the vegetative cell for pollen germination (Heslop-Harrison, 1978). In cereals, germination occur quite rapidly taking place in less than five minutes for triticales (Heslop-Harrison, 1979b). The signalling for germination upon cell polarization still remains unclear but studies do show that a water gradient does act as a cue for guiding the pollen tube growth towards the stigma. In the pollen coat deficient *Arabidopsis cer* mutants, when no water potential was established, the pollen tubes emerged randomly and were unable to perceive the presence of the stigma (Preuss et al., 1993).

The last step involved in the pollen-stigma interaction is pollen germination and tube entry into the stigma. In rye, pollen tube emergence is strictly restricted to the aperture/pore and does not break through the exine wall (Heslop-Harrison, 1979b). The

presence of a cuticular layer covering the stigma, a characteristic among dry type stigmas, acts as a physical barrier preventing pollen tube entry. In order for the tube to penetrate through the stigma, it must first breach the cuticle layer of the pellicle. In the Poaceae family, pollen cutinase activity was detected during pollen tube penetration (Heslop-Harrison, 1979b). Wall softening enzymes found at the tip of the pollen tube including β -(1-3)-, β -(1-4)-glucanase and pectinases were also present to facilitate penetration of the stigma (Heslop-Harrison, 1979b; Pressey & Reger, 1989). Furthermore, group I allergens in grass pollen, such as β -expansins were also identified as mediating tissue weakening of the grass stigma and style for pollen tube invasion (Cosgrove et al., 1997).

1.6 Characterization of the anther and mature pollen transcriptomes

Comprehensive knowledge of the role of both the female and male counterparts must be attained in order to understand the processes involved in plant reproduction. The introduction of DNA microarray technology has now made it possible for rapid high-throughput gene expression analysis of tens of thousands of genes simultaneously. With the ability to measure individual transcript levels on a larger genome-wide scale, the field of genomics have now become established for surveying gene expression signatures associated with abiotic stress responses, pathogen responses, or the full transcriptomes of various tissue types. Over the years, microarrays have evolved from cDNA spotted arrays to more comprehensive platforms such as the Affymetrix GeneChip® Genome Array. The Affymetrix Wheat GeneChip® Genome Array represents the most comprehensive wheat microarray available to date and it contains 61,290 probe sets

potentially representing 55,052 wheat transcripts and including 127 internal control probe sets. The oligonucleotides on the array were derived from expressed sequenced tags (ESTs) and full length mRNAs from GenBank[®] (dbEST and UniGene) from *Triticum aestivum*, *Triticum monococcum*, *Triticum turgidum*, and *Aegilops tauschii*.

In the past decade, substantial advances have been made in the study of the stamen (see reviews Scott et al., 2004; McCormick, 2004; Grennan, 2007) and emerging microarray-based transcriptome studies will help to further elucidate the biological processes involved in stamen development. Accordingly, there has been a large variety of microarray studies performed across different plant species. Global expression profiling has been performed on anther tissues and/or pollen throughout development. Others took a step further and examined the male gametophyte transcriptome by specifically isolating pollen at landmark stages of development. Recently, laser microdissection (LM) has been combined with microarray analyses to study the transcriptomes of microspore/pollen cells or anther cell types separately. This has proven to be reliable in revealing genes that are derived from the male gametophyte or specific to a particular anther cell type such as the tapetum which previously was very difficult to differentiate (Suwabe et al., 2008).

Microarray studies involving stamen development have been performed in *Lilium longiflorum* (lily, Okada et al., 2006); *Glycine max* (soybean, Haerizadeh et al., 2009); *Zea mays* (maize, Ma, et al., 2008); *Arabidopsis thaliana* (Arabidopsis, Honys & Twell, 2004; Pina et al., 2005; Borges et al., 2008); *Oryza sativa* (rice, Hobo et al., 2008; Wei et al., 2010); *Hordeum vulgare* (barley, Druka et al., 2006) and *Triticum aestivum* (wheat, Crismani et al., 2006; Schreiber et al., 2009). Transcriptome studies on the whole anther

in cereals have been performed mostly in rice and the most recent extensive studies involve high density arrays. Expression profiling of rice anthers at five developmental stages: meiosis, tetrad, uninuclear microspore, bi-cellular pollen, and tri-cellular pollen were performed using the Agilent rice 44K oligo microarray (Suwabe et al., 2008). In addition, the tapetum was isolated from the microspore/pollen at meiosis, tetrad, and uninuclear microspore stage using laser capture in order to characterize the transcriptomes separately. In parallel, Hobo et al. (2008) performed an extensive cluster analysis on spatiotemporal expression patterns of the same samples and classified each cluster according to gene ontology categories that represent biological events related to anther development. Predominant processes represented during early anther development include pollen wall formation and fatty acid synthesis. Several MADS box and basic helix-loop-helix (bHLH) zinc-finger-like transcription factors were shown to be essential in anther cell specification, a finding reported in other plant species (Ma et al., 2007; 2008 Skibbe et al., 2008).

Several key regulators involved in anther development have been identified using mutant analysis and the expression profiles of these mutants have been compared to their wild-type counterparts. The *MSPI* (*MULTIPLE SPOROCTE1*) gene is involved in early sporogenic development and tapetal cell fate determination in rice and Arabidopsis (Nonomura et al., 2003; Alves-Ferreira et al., 2007). It encodes a putative leucine-rich repeat receptor kinase (LRR-RK) orthologous to the Arabidopsis EXS/EMS1 (EXTRASPOROGENOUS CELLS/EXCESS MICROSPORCYTES1) LRR-RK (Zhao et al., 2002; Wijeratne et al., 2007). The Undeveloped Tapetum1 (Udt1) is a rice bHLH transcription factor with a putative role in tapetal development and PMC meiosis as

identified using the 60K Rice Whole Genome Microarray (Jung et al., 2005). Anther transcriptome data has also been produced using ultra-high throughput platforms such as next generation sequencing. Transcriptome profiling of anther tissues at 6 developmental stages and mature pollen has been generated in rice using sequencing-by-synthesis technology yielding more than 1 million sequenced signatures (Huang et al., 2009). A total of 18,267 genes were identified in the anther transcriptomes of which approximately 12-25% were natural occurring antisense transcripts whose functions remain to be elucidated. The ability to generate millions of sequences will be a valuable tool for deep transcriptome profiling and should accelerate investigations in reproductive development.

Most of the information provided for the pollen transcriptome focus on rice and Arabidopsis. For example, using the Affymetrix GeneChip® Arabidopsis ATH1 Genome Array representing approximately 24,000 genes, Pina et al. (2005) found 6,587 genes expressed in pollen with a large proportion (11%) selectively expressed in pollen over other vegetative tissues, whereas 26% were pollen-enriched. Gene ontology revealed a functional skew towards cell cycle and signalling, vesicle transport, and cell wall metabolism. Notably, a transcriptomic study by Honys and Twell (2004) focusing on Arabidopsis microsporogenesis was performed using the same chip and using isolated microspores/pollen at four developmental stages: uninuclear microspore (UNM), bi-cellular pollen (BCP), tri-cellular pollen (TCP), and mature pollen grain (MPG). Of the 13,977 pollen expressed genes, an increase in both pollen-specific and low expressing transcripts was observed over development. They also reported the functional skew described by Pina et al. (2005). A transcriptome study of soybean pollen using the Affymetrix Soybean GeneChip® has also reported similar enrichment of transcripts in

addition to transporter and heat shock-related genes (Haerizadeh et al., 2009). Lastly, the ATH1 array was also used to reveal the 5,829 genes expressed in the sperm cells (Borges et al., 2008). A majority of these transcripts (65%) were also detected in the pollen. Overrepresented sperm-enriched transcripts were involved in DNA repair, ubiquitin-mediated proteolysis, and cell cycle progression.

In cereals, comparison of expression profiles of isolated rice pollen at four sequential developmental stages (UNM, BCP, TCP, MPG) as well as germinated pollen grain (GPG) using the Affymetrix GeneChip[®] Rice Genome Array also revealed similar pollen-preferential expression (Wei et al., 2010). Among the 25,062 pollen-expressed transcripts, 2,203 were developmentally stage-enriched, with 660 found at the UNM stage (Wei et al., 2010). Analysis of their functional features demonstrated dynamic changes at distinct stages. For example, transcripts overrepresented in UNM included protein degradation as well as transcription and RNA processes, whereas metabolism was overrepresented in BCP. Transporter and channel activity was predominant in TCP, while metabolism and defense/stress response, and transcription/RNA processes were most prevalent in mature pollen and germinated pollen respectively. Based on the different functional groups represented throughout pollen development, it is evident that different genes are required at particular stages for proper development and function specification. More importantly, among the many identified transcripts enriched or specific to the pollen, a vast majority still remain unknown with no putative annotation. RNA-seq profiling data was also obtained for mature rice pollen as part of an anther development study mentioned above (Huang et al., 2009).

Generally, transcriptome studies of the early development of cereal pollen have been limited. Recently, a step towards characterizing the premeiotic transcriptome was performed by isolating a population of pollen mother cells (PMCs) of rice using laser-capture microdissection and the Agilent 60-mer oligo-microarray (Tang et al., 2010). Isolated PMCs were compared to tri-cellular pollen and seedling tissue. A total of 1,158 genes were identified as preferentially expressed in the PMC, demonstrating at least a 4-fold higher expression. Markedly, all genes belong to the DNA replication and repair category which is indicative of roles in meiotic recombination and cell cycle regulatory events. Furthermore, 127 of those genes were expressed exclusively in PMCs and not in the other two tissues.

Two studies reported the anther transcriptome of the Triticeae. The first study surveys the expression profiles of 21,439 barley genes in 15 different tissue samples, including anthers prior to anthesis, using the Affymetrix GeneChip® Barley1 Genome Array (Druka et al., 2006). The purpose of this study was to build an expression profile database of barley genes for comparative analyses. The smallest number of expressed probe sets was found in the mature anther in comparison to the other tissues. However, of the 650 probe sets identified as tissue-specific, the highest number (251) was found in the anther. Furthermore, of the known 1,059 probe sets representing barley regulatory factors, such as transcription factors and protein kinases, the anther contained the highest number (135). This is consistent with the fact that the anther possesses highly specialized functions required for floral development. In the second study, gene expression profiling in wheat was performed using the Affymetrix Wheat GeneChip® on the same tissue series used for barley above (Schreiber et al., 2009). This is considered the first exclusive

comparative analysis of key stages of plant development between two important crops that are evolutionary closely related. Putative homologous genes identified between wheat and barley can be compared using the web based program *WebComparator*: <http://contigcomp.acpfg.com.au/>. For example, hierarchical clustering of 1,875 'high quality' orthologous probe sets demonstrated a gene expression pattern that is highly conserved among the two species, with expression in barley generally higher than in wheat (Schreiber et al., 2009). The ability to compare transcript levels across different tissues in different species is definitely a significant advancement in Triticeae transcriptomics.

Only one Triticeae developmental anther transcriptome study was found and it was restricted to young anthers during meiosis (Crismani et al., 2006). To identify meiotically regulated genes, transcriptome data was generated for 7 stages of anther development including five meiotic stages: pre-meiosis, leptotene to pachytene, diplotene to anaphase I, telophase I to telophase II, tetrad, as well as immature pollen and mature anther. Upon comparison across the 7 stages, the most variability in expression patterns was observed in mature anthers, followed by immature pollen. Cluster analysis revealed 350 meiotically-regulated transcripts with the largest proportion (60, 17.1%) falling into the functional classification of meiotic/and or cell division. These transcripts could be further sub-divided into meiotic/cell division categories with the majority predicted to encode histone-related (25%) and putative cell cycle proteins (23.3%), followed by cytoskeletal- (16.7%), and chromosome-associated (16.7%) proteins. So far, this is the most detailed analysis of plant male meiosis.

In summary, considerable research has been devoted to the characterization of the *Arabidopsis* and rice anther/pollen transcriptomes, and less so within species of the *Triticeae*.

1.7 Characterization of the pistil transcriptome

The pistil transcriptome has not yet been as extensively studied compared to the stamen transcriptome. There have been no microarray studies of the entire pistil transcriptome in *Arabidopsis* or rice. However, a comparison of various re-constituted pistil microarray analyses and EST datasets was reported for plants with ‘wet’, ‘dry’, or ‘semidry’ type stigmas (*Crocus sativus*, *Oryza sativa*, *Arabidopsis thaliana*, *Nicotiana tabacum* and *Senecio squalidus*) (Allen et al., 2010). Most differences were observed between the wet (tobacco) and dry (*Arabidopsis*, rice, crocus) stigma-types, while the semidry (*squalidus*) shared more genes with the dry type (Allen et al., 2010). Genes preferentially expressed in the pistils of all four different species encode, for example, cytochrome P450, lipid transfer proteins (LTPs), and receptor protein kinase.

Several whole genome transcription profiles have been performed on separated ovary and stigma tissues, and a significant amount of literature is based primarily on research done in *Arabidopsis*. A fair portion consists of expression analysis of the ovule as opposed to the whole ovary. Gene expression profiles of whole ovary tissues have been produced, but they mainly served for comparison purposes in analyses of the stigma transcriptome of *Arabidopsis* (Swanson et al., 2004; Tung et al., 2005) or rice (Li et al., 2007).

As previously mentioned, the ovule consists of four parts: the embryo sac, the nucellus, the integuments, and the funiculus stalk that connects the ovule to the placenta. Many genes controlling ovule identity have been identified through mutant analysis. For example, microarray analysis on *Arabidopsis* pistils comparing wild type to mutants *ino-1*, which lacks outer integument, and *ant-4*, which lacks both inner and outer integuments, was performed to identify genes regulating integument ontogeny in the ovule (Skinner & Gasser, 2009). Comparison of both mutants to the wild type revealed 132 genes expressed within the inner integument, 50 in the outer integument, and 25 that showed preferential expression in both. Among them, the most highly represented functional classes included proteins involved in metabolism, transcriptional regulators and DNA-binding proteins, as well as many genes of unknown function (especially inner integument genes). The overabundance of transcription factors (20%) in the dataset can be compared to that of the whole genome which consists of an estimated 6-7% only (Skinner & Gasser, 2009). Other gene expression analyses focused specifically on the female gametophyte, or embryo sac, and demonstrated a distinct transcriptome. The female gametophyte consists of the egg cell, central cell, two synergid cells, and three antipodal cells (Jones-Rhoades et al., 2007). Studies have typically used the mutant *sporocyteless/nozzle* (*spl/nzz*) because it encodes a transcription regulator of sporogenesis required in both sexes and produces ovules lacking an embryo sac. Yu et al. (2005) identified 225 female gametophyte-specific genes by performing microarray analysis on *spl* mutant ovules. Validation of six genes was performed using promoter:: β -*glucuronidase* (GUS) fusions and all showed embryo sac-specific expression with three genes expressed exclusively in the antipodal cells. Similarly, comparison between

mutants *male sterility* (*ms1*) (normal ovules) and *determinant infertile* (*dif1*) (ovules lacking female gametophyte) identified a collection of genes expressed in specific cells of the female gametophyte (Steffen et al., 2007). A differential expression screen was performed identifying 71 genes displaying reduced expression in the *dif1* ovules. Based on promoter::GUS fusion assays, 40 genes were found exclusively or predominantly expressed in one cell type of the ovule. Most recently, cell-type specific transcriptomes were generated for the Arabidopsis female gametophyte using laser-assisted microdissection (Wuest et al., 2010). This method is a rapid and efficient way of screening for potential candidate genes involved in development and regulatory networks in the female gametophyte. Posttranscriptional regulatory mechanisms showed significant differential expression among distinct cell types (Wuest et al., 2010). When compared to 59 other Arabidopsis tissues, transcription factor families containing type1 MADS and RWP-RK domains were highly expressed. Furthermore, double-stranded RNA binding factors were abundant in the egg or central cell suggesting involvement in gene silencing pathways.

A recent microarray study based on the Affymetrix platform profiled the rice reproductive process from anther development to embryogenesis (Fujita et al., 2010). This is the most extensive comparative analysis of gene expression in male and female tissues to date that includes anther tissue from pre-meiosis to TCP and ovary and stigma at the mature stage and post pollination.

As mentioned previously, the stigma is very important in plant reproduction. However, data characterizing its underlying molecular mechanisms and functions are limited. Stigma transcriptome studies using ESTs have been performed on saffron

(D'Agostino et al., 2007) and tobacco (Quiapim et al., 2009). To our knowledge, there are only two stigma microarray gene expression studies and they relied on the *Arabidopsis* ATH1 GeneChip[®] (Swanson et al., 2005; Tung et al., 2005). After comparison of transcription profiles of isolated stigmas, ovaries, as well as pistils from transgenic plants where the stigma epidermis was specifically ablated by a cellular toxin, Tung et al. (2005) identified 115 genes expressed exclusively in the stigma. Similarly, Swanson et al. (2005) compared stigma, ovary, and seedling expression profiles using the same chip and in combination with cDNA library subtraction, built a profile of 317 candidate stigma-expressed genes. In both cases, many of the highly expressed genes in the stigma involved cell signalling and lipid metabolism, suggesting roles in pollen adhesion, tube growth and penetration.

Within cereals, the only stigma transcript profiling study was performed on the mature rice stigma and its transcriptome was compared to 7 different tissues, including mature anther and ovary (Li et al., 2007). They identified 548 genes preferentially expressed in the stigma and several of these belonged to groups having conserved functions in both rice and *Arabidopsis*. The majority of these roles were cell wall- and signal transduction-related, which are functions likely involved in the pollen-pistil interaction. Most notably, of the limited stigma gene expression experiments performed, all reported very similar functional categories demonstrating conserved functions across taxa.

Markedly, studies of Triticeae female tissues are very limited and only include expressed sequence tags (ESTs) generated from cDNA libraries of egg cells or two cell pro-embryos isolated from wheat using LM (Sprunck et al., 2005). Transcripts specific

to the egg cell were highly represented in functional categories related to primary and secondary metabolism followed by ribosomal proteins, with the majority of transcripts having an unknown function. As mentioned above, there have been transcriptome studies on the same collection of tissues from barley and wheat and these included the mature pistil (Schreiber et al., 2009; Druka et al., 2006). Since the main goal of these studies was to generate global transcriptome resources for these cereals, the pistil transcriptome was not characterized in detail. However, Druka et al. (2006) reported 13,817 barley pistil-expressed probe sets with very few pistil-specific genes (4).

In summary, as with the anther/pollen, most studies on the female reproductive organ have been performed in rice and Arabidopsis. The few studies carried out in the Triticeae did not cover different phases of development or focus on the stigma or ovule.

1.8 Objectives

Reproductive management systems can have a considerable effect on crop production as evidenced by the impact of hybrid crops in agriculture. As previously mentioned, cereals are of immense agricultural importance worldwide. To properly manage reproduction in the Triticeae, a thorough knowledge of the molecular mechanisms involved in cereal reproductive development is required and in this study triticales will be used as a representative of the Triticeae. Presently, no known study encompasses parallel transcript profiling of male and female reproductive tissues throughout development.

Therefore, the main objective of this study was to use the most comprehensive wheat microarray available, the Affymetrix 55K Wheat Genechip[®], to build the most complete Triticeae reproductive transcriptome profile resource. This research not only

provides a transcriptome database useful to all researchers interested in Triticeae reproduction, but also provides a better understanding of pollen/pistil interactions in order to increase seed production (eg. hybrid production) or help prevent outcrossing and gene flow from transgenic crops.

2. Materials and Methods

2.1 Plant material and growth conditions

Triticale (x *Triticosecale* Wittmack cv. Alta) plants were grown in 15 cm diameter plastic pots containing soil in a temperature-controlled growth chamber maintained at 20°C (day) and 18°C (night) under a photoperiod of 16 h light (250-275 $\mu\text{E m}^{-2} \text{s}^{-1}$) provided by fluorescent and incandescent light. Plants were held at a constant humidity of 70% and watered twice daily and fertilized bi-weekly with 20-20-20 (NPK) fertilizer. Plants were separated into three separate biological replicates comprising eight plants per replicate and flowers were produced after 2-3 months under these growth conditions.

2.2 Microscopy

Fluorescence Microscopy-DAPI Staining

Triticale anther squashes were prepared to determine the stages of pollen development used in this study: tetrad (TET), uninucleate microspore (UNM), bi-cellular pollen (BCP), tri-cellular pollen (TCP), and mature pollen grain (MPG). The corresponding Zadoks' scale (Zadoks et al., 1974) is as follows: Z41 (TET), Z49 (UNM), Z55/57 (BCP), Z59 (TCP), and Z61 (MPG). Using anther length as a guide, triticale anthers were dissected using a Leica Zoom 2000 dissecting microscope and fixed immediately in 3:1 (70% ethanol:acetic acid) for 30-90 min. Anthers at the tetrad stage were stained with 10 μl of a 1:1 solution of 0.1 $\mu\text{g/ml}$ 4',6-diamidino-2-phenylindole (DAPI):0.05% Aniline blue in 0.067M KPO_4 to observe the callose walls, whereas anthers at later stages were stained in 10 μl of 0.1 $\mu\text{g/ml}$ DAPI only. Prepared slides

were observed under a fluorescent microscope (Zeiss Axioplan 2 Imaging) and images were taken with a digital camera (Zeiss AxioCam) and the Axiovision software (Carl Zeiss Microimaging). Based on their cytological changes, anther lengths can be used as a reference point for staging (Bennett et al., 1973). The anther lengths corresponding to the four stages were defined as: 1.3-1.5 mm for tetrad (TET), 2.5-3.5 mm for uninucleate microspore (UNM), 4-4.5 mm for bicellular pollen (BCP), and 5.5 mm or yellow anthers for the tricellular pollen stage (TCP). These anther sizes were also used as reference points for staging of the female tissues (stigma and ovary).

Light microscopy

Fresh stamens and pistils were dissected at the TET, UNM, BCP, and TCP stages of pollen development. Phenotypic analysis for each sample was observed under a Zeiss Discovery V12 dissecting microscope.

Scanning electron microscopy (SEM)

For a more in-depth analysis, SEM was performed on each reproductive organ at the TET, UNM, BCP, and TCP stages of pollen development. Fresh stigmas or ovaries were dissected and placed onto a SEM stub layered with a carbon adhesive sticker and examined immediately using a Philips XL30 ESEM (FEI, USA) operated at 7.5kV under low vacuum at almost 0.1 Torr.

Cryo-SEM

Cryo-SEM was performed as follows: (1) a Philips XL30 SEM (FEI, USA) coupled with a Hexland Cryo-Trans CT 1000 (UK, England) cryogenic preparation unit was pre-cooled to at least -180°C using liquid nitrogen (N₂) and held under high vacuum (10⁻⁵ Torr). (2) Fresh developing whole anthers containing pollen were dissected and mounted on a 3-hole mounting stub. Each anther was placed in a separate hole filled with Tissue-Tek® OCT Compound (Sakura Finetek #4583), a glycol resin matrix glue where samples can be frozen off the stub to achieve fast freezing rates for cryostat sectioning. The stub was fixed to a custom made spring-loading plunger and immediately frozen in liquid N₂ slush within the freezing chamber. The sample was frozen within 1 min and then transferred to a pre-cooled preparation unit where they can be observed. (3) The anthers were fractured with a scalpel and subsequently sputter coated with 2 nm of gold using argon gas under vacuum conditions. (4) The sample was then transferred to the pre-cooled cryo-stage within the SEM chamber for observation. All micrographs were taken under 7.5kV accelerating voltage.

2.3 RNA isolation and Affymetrix GeneChip® hybridization

RNA isolation

Anthers, ovaries, and stigmas were harvested at each of the four pollen developmental stages: TET, UNM, BCP, and TCP. Mature pollen was also collected for analysis. Freshly dissected tissues were immediately flash frozen in liquid nitrogen and stored at -80°C until use. Mature pollen was collected onto 3" x 3" weighing paper and flash frozen in a 1.5 ml microtube. For each biological replicate, total RNA was isolated

from anthers, ovaries, and stigmas at each of the four developmental stages, and from mature pollen. To prevent RNA degradation, all glassware was baked overnight at 250°C and all required solutions were made up in 0.1% v/v diethylpyrocarbonate (DEPC)-treated water and autoclaved. Solutions that included either sodium dodecyl sulfate (SDS) or Tris base were prepared with autoclaved DEPC-treated water instead.

Three different RNA extraction methods were used depending on availability and nature of the tissue sample.

(1) At the TET stage, total RNA was extracted from each of the following: 10 anthers, 5-6 ovaries and 10-20 stigmas using the PicoPure™ RNA isolation kit (Molecular Devices, Sunnyvale, CA) according to the manufacturer's instructions (including DNase treatment). This kit was used because tissue at the TET stage is very limited. Only nanograms of total RNA were obtained.

(2) Total RNA from frozen anthers, ovaries, and stigmas at the UNM and BCP stages was extracted using the RNeasy Plant mini kit (Qiagen, Mississauga, ON) following the manufacturer's instructions. RNA samples were subsequently cleaned up using the InviTrap® Spin Cell Plant RNA Mini Kit (Invitex, Berlin, Germany).

(3) For mature pollen and samples at the TCP stage, total RNA was isolated using a modified version of the guanidine isothiocyanate/cesium chloride ultracentrifugation method described by Wang et al. (2010). Whole frozen tissues were used as the starting material: 0.20 g of anthers, 0.10 g of stigmas, 0.20 g of ovaries, or 0.15 g of mature pollen, and ground with a mortar and pestle in liquid nitrogen. The ground tissue was immediately added to 3.3 ml of GIT buffer [4 M guanidine isothiocyanate, 25 mM sodium acetate pH 6.0, 0.8% β -mercaptoethanol] in a 4 ml polypropylene culture tube

(UltidentBrand, 12 x 75 mm). Samples were homogenized for at least 30 sec using a PRO 200 handheld homogenizer (PRO Scientific Inc, 5 mm x 75 mm generator). Samples were centrifuged at 5,000 rpm to sediment tissue debris. The supernatant was layered onto 2.2 ml of filter sterilized CsCl buffer [5.7 M cesium chloride, 25 mM sodium acetate pH 6.0] in a 5 ml ultracentrifugation tube (Beckman polyallomer tube, No. 326819, 13 x 51 mm). To pellet the RNA, samples were centrifuged at 35,000 rpm using a SW 55T swinging bucket rotor for 23 h at 20°C in a Beckman LB-70M ultracentrifuge. A sterile transfer pipette was used to carefully remove most of the GIT and CsCl buffers. The remaining CsCl layer was decanted gently and drained to isolate the pellet at the bottom of the tube. The centrifuge tube was cut down to approximately 2 cm from the bottom and placed on ice. The pellet was rinsed twice with 500 µl of 70% ethanol. After the last rinse, residual ethanol was removed by decanting the tube and placing it inverted on a Kimwipe tissue for 5-10 min. The pellet was subsequently dissolved in 50 µl of TES buffer [10 mM Tris-HCl pH 7.4, 5 mM EDTA, 1% SDS] and left on ice for 1 h to precipitate out the SDS. The 50 µl solution (pellet + TES) was then transferred to a clean 1.5 ml Eppendorf microtube on ice and 5 µl of 3 M sodium acetate pH 5.2 and 125 µl of cold 95% ethanol were added for an overnight precipitation at -20°C. The samples were then centrifuged at 14,000 rpm for 30 min at 4°C to pellet the RNA. The pellet was rinsed three times with 70% ethanol. After the last rinse, the sample was re-centrifuged at 14,000 rpm for 2 min, the remaining ethanol was removed, the pellet was air dried on ice (~ 25 min), dissolved in nuclease-free water (Ambion, Streetsville, ON) and stored at -80°C.

The yield and purity of all RNA samples were verified using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Ottawa, ON), an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA), as well as gel electrophoresis. The gel apparatus including tray, comb, and wedge supports were pre-treated for at least 1 h in a 200 ml solution of 0.1% SDS and 0.1 N NaOH. A 1 μ l aliquot of RNA was resuspended in 3.5 μ l RNA loading buffer [11% (v/v) 10X MOPS, 7.3% (v/v) 37% formaldehyde, 5.6% glycerol, 0.7 mM EDTA, 0.04 mg/ml ethidium bromide, 0.56% (v/v) deionized formamide, and 0.2% (w/v) Bromophenol Blue], denatured in a heating block at 65°C for 20 min, placed on ice, and resolved in a 1.3% agarose gel [10% (v/v) 10X MOPS, 1.3% (w/v) agarose, 11% (v/v) 37% formaldehyde]. The gel was run at 100V for 1 h in 1X MOPS running buffer [10 mM MOPS, 5 mM sodium acetate trihydrate, 10 mM EDTA pH 8].

Affymetrix GeneChip[®] hybridization

The Affymetrix GeneChip[®] Wheat Genome Array hybridizations including cDNA synthesis, target labeling, hybridization, washing, staining, and scanning were performed at the McGill University and Génome Québec Innovation Centre. Three different Affymetrix labeling protocols were used based upon sample size or because the one and two-cycle kits were discontinued during the course of the experiment. For samples at the TCP stage and mature pollen, the older version of the one-cycle target labeling protocol was used, starting with 5 μ g of total RNA. Poly(A) mRNA was converted to double stranded cDNA using an Affymetrix T7oligo(dT) primer for *in vitro* transcription (IVT) using the SuperScript Choice System (Invitrogen). Biotin-labeled

antisense cRNA targets were produced by incorporating biotin-ribonucleotides using T7 RNA polymerase (Enzo Life Sciences), and then purified and chemically fragmented as per the Affymetrix technical manual. Biotin-labeled cRNA (15 µg) were hybridized to the Affymetrix 55K Wheat GeneChip[®] array for 16 h at 45°C. Washing, staining, and scanning of the microarrays were performed as described on the Affymetrix website (<http://www.affymetrix.com/>). For TET samples, the older version of the two-cycle target labeling protocol was used for 50-100 ng samples of total RNA. The modified protocol included two rounds of cDNA synthesis and IVT reactions for target amplification. For UNM and BCP samples, the new generation 3'-IVT labeling kit was used which replaced the previous one-cycle and two-cycle target labeling assays (Enzo Life Sciences) and required 1 µg of total RNA starting material. The new optimized protocol is also available on the Affymetrix website.

Three biological replicates were performed for the anther, ovary, and stigma at four developmental stages, as well as for mature pollen, giving a total of 39 arrays. The different RNA isolation and labeling methods were compared in separate hybridization experiments and were found to have little impact on experimental results with the correlation coefficients ranging from 0.9502 to 0.9925 (data not shown).

2.4 Data analysis and experimental design

Scanned arrays were pre-processed using the Affymetrix Expression Console software. Microarray Suite 5.0 (MAS5) software was used to make an absolute detection call for each transcript: present (P), absent (A), or marginal (M) with a corresponding confidence value of $P < 0.05$. In subsequent analyses the marginal calls were considered

unreliable and omitted from the analysis. The probe cell intensity (CEL) files for all arrays were normalized using ArrayStar4 (DNASTAR, WI, USA) to produce $\log_2$ transformed values. The Robust Multichip Average (RMA) normalization algorithm was applied, which included background correction, quantile normalization, and a median polish summarization (Bolstad et al., 2003). Overall, a conservative approach was implemented whereby probe sets that have a present call in all three biological replicates (3P) were considered 'expressed'. The experimental design for each type of microarray analysis is described below and summarized in Figure 2-1.

Filtering of probe sets that show expression

Of the 61,290 probe sets on the Affymetrix GeneChip[®] Wheat Genome Array all internal control probe sets were removed, reducing the dataset to 61,115. Genes showing no expression in all tissues based on MAS5 ($P \geq 0.05$) were excluded from further analysis. Microarray analysis was performed using ArrayStar 4 (DNASTAR, WI, USA) software package, except for the cluster analysis in which a combination of Acuity 4.0 (Molecular Devices, CA, USA) software was used. Probe sets that were expressed in at least one reproductive tissue or developmental stage were included in the analysis, and consisted of a total of 31,550 probe sets. The analyses described below were performed on three different datasets, one for each reproductive tissue.

Figure 2-1. Experimental design and workflow summary of microarray analyses. Number in parenthesis indicates the number of different sample arrays used, and type of analysis in red. RMA, robust multichip average; P, present call; ANOVA, analysis of variance; 2-FC, 2-Fold-Change.

Scatter plot, principle component analysis, and dartboard models

To determine the relationship between transcriptomes of different developmental stages within a specific organ, pairwise transcriptome comparisons were made using scatter plots. Probe sets that co-express in the two stages being compared were plotted on a base 2-logarithmic scale ($\log_2$) and the linear correlation (R^2) was calculated. Principle component analysis was performed using the FlexArray software package (McGill University and Génome Québec Innovation Centre) to examine in more detail the variability among different samples. A dartboard analysis using ArrayStar was also performed to examine the extent of gene expression overlap between developmental stages in a particular organ. Probe sets that were expressed in at least one stage where included in the analysis.

Cluster analysis of filtered dataset

To identify expression trends, an unsupervised *K*-means clustering was performed for each dataset which included all probe sets that are expressed in at least one stage of development. To clarify the graphical outputs, probe sets that did not generate a statistically significant signal in all three biological replicates (and therefore considered 'absent') were given an expression signal value of '0' using Microsoft Excel. To find the optimal number of clusters, a gap statistic test was performed (Tibshirani et al., 2001). Analyses performed with varying numbers of clusters were evaluated manually and found to be in good agreement with the number specified by the gap statistic as optimum in partitioning the data.

Identification of differentially expressed probe sets (DEPs)

Differentially expressed probe sets (DEPs) throughout the different stages of development or among different tissues were analyzed using the ArrayStar4 software. To avoid false negatives or false positives in these analyses, the following approach was used. First, two different datasets were generated. The 3P dataset only considers probe sets with 3P calls as present (ie. values of '0' are given to 1A/2P, 2A/1P, or 3A). This is a very stringent approach and removes probe sets with absent calls that have relatively high expression values (these would hide 3P probe sets with low expression values and generate false negatives). The 2P dataset on the other hand, corresponds to an analysis of differential expression which includes values from probe sets having a 2P call. Giving a value of '0' to probe sets with 2P calls as done in the 3P dataset could lead to false positives. The intersection or common probe sets between the 2P and 3P dataset contains differentially expressed probe sets. Therefore, the probe sets remaining are 3P, but the 2P probe sets are considered and 2A probe sets are ignored, thus reducing both the number of false positives and negatives. This approach was used to generate the data for the up- and down-regulated, the enriched, and the specific probe sets described below.

Up- and down-regulated genes

An analysis was performed to identify probe sets that are up- or down-regulated during the development of a particular organ. First, RMA normalization was performed using the datasets of all the reproductive tissues. A one-way analysis of variance (ANOVA) statistical test including the Benjamini-Hochberg false discovery rate (FDR) multiple testing correction was carried out on one organ across the developmental stages

to identify probe sets whose level of expression shows statistically significant variation throughout development. Probe sets were considered statistically significant if they met a corresponding confidence cutoff value of $P < 0.05$. Both up- and down-regulated DEPs were identified. Probe sets that were considered differentially up-regulated in a particular stage met the following criteria: i) display a significance level of $P < 0.05$ and ii) demonstrate an up-regulation of at least 2-fold change (2-FC) above expression levels in the other stages. Probe sets that were considered differentially down-regulated met the following criteria: i) display a significance level of $P < 0.05$ and ii) at least a 2-FC down-regulation compared to the previous stage. Microarray data typically span a large numerical range contributed by signal noise, therefore the final DEPs datasets were further filtered removing signal intensity values less than or equal to 7 ($\log_2$). This $\log_2$ value corresponds to an absolute expression signal value of 128 and it was chosen because the majority of absent calls were predominantly below 7 and a similar cutoff value was used for other microarray studies pertaining to the Triticeae (Druka et al., 2006; Schreiber et al., 2009). The ≥ 7 dataset therefore represents DEPs which show expression at a more reliable detection range.

Identification of tissue-enriched genes

To identify genes truly enriched in a particular tissue, gene expression was compared across other reproductive tissues at all stages of development, and to a separate triticale Wheat GeneChip[®] microarray experiment comprising 16 vegetative tissues including: coleoptile, root, 5 seed tissues – embryo, epidermis, endosperm, pericarp, and crease, stem tissue throughout development based on the Zadoks' scale Z21, Z45, Z59,

Z75, and leaf tissue at seedling, tillering, early boot, early and late senescence, totaling 25 tissues being compared and representing most triticales tissues (Penniket et al., 2009). For this analysis, RMA normalization was performed across all 29 different tissues. An ANOVA statistical test including the Benjamini-Hochberg false discovery rate (FDR) multiple testing correction was also carried out on all 29 datasets to remove probe sets that do not show significant variation in expression across all samples. Only probe sets that show at least a 2-fold up-regulation over all other tissues being compared were considered tissue-enriched. Developmental stage-enriched genes within each reproductive organ were identified as probe sets having at least a 2-fold increase in expression over all other tissues including stages within that particular tissue.

Identification of tissue-specific genes

The enriched dataset for the reproductive tissues contains not only probe sets whose expression is enriched with respect to expression in other tissues, but it also includes probe sets that show expression in the reproductive tissues and absence of expression in all other tissues. Probe sets considered tissue-specific were present in at least one stage of development of a particular tissue but absent in all other tissues being compared. Likewise, tissue-stage-specific probe sets were identified using the same comparison, but were also absent in the other stages of that particular tissue.

2.5 Gene Ontology (GO)

The Affymetrix Wheat GeneChip[®] was annotated by using an automated algorithm developed by the Laroche lab (AAFC-Lethbridge, AB). Each EST sequence from which the probe set was derived from was annotated by first performing a BLASTn

(<http://blast.ncbi.nlm.nih.gov>) to retrieve a longer version of the sequence if available, followed by a BLASTx to retrieve the corresponding protein sequence which was then annotated using the Blast2Go software (www.blast2go.org) under default parameters. Genes of interest (pollen-, anther-, ovary- and stigma-specific probesets) were manually verified and annotated using the same method. For unknown proteins, a search was performed using InterPro Scan (<http://www.ebi.ac.uk/>) and Pfam (<http://pfam.sanger.ac.uk/>) databases to identify domains and gene ontology (GO) classifications. GO process groups were merged into 18 categories that are most representative of the anther/pollen-, ovary-, and stigma-specific dataset.

3. Quality Control and Data Statistical Analysis

3.1 Results

3.1.1 Pearson correlation coefficient

A global gene expression analysis was performed on the triticales anther, ovary, and stigma throughout development, as well as mature pollen. Experiments were performed with the most comprehensive wheat microarray available, the Affymetrix Wheat GeneChip[®], which represents 55,052 transcripts. The fact that triticales consists of a *Triticum* and a *Secale* genome allows the use of the wheat chip for triticales analysis, although, due to the high homology, the wheat and rye genomes cannot be distinguished based on hybridization. The four developmental stages chosen correspond to landmark stages in pollen development: tetrad (TET), uninucleate microspore (UNM), bi-cellular pollen (BCP), and tri-cellular pollen (TCP). Based on Pearson correlation coefficients, the correlation among the three biological replicates for each sample ranged from 0.9770-0.9957 (Table 3-1) indicating that the overall quality of the replicate data was high and exceptionally consistent.

3.1.2 Expression value and detection P-value call

From the expression values generated by MAS 5 (see Materials and Methods), a statistical P-value was generated suggesting the level of confidence for each qualitative measurement. Probe sets with a detection P-value < 0.05 were considered present (P). Table 3-2 displays the number of probe sets that were considered present for each individual biological replicate and as predicted by the high Pearson correlation

Sample	Pearson Correlation Coefficient		
	REP 1/2	REP 2/3	REP 1/3
Anther			
TET	0.9916	0.9895	0.9882
UNM	0.9900	0.9922	0.9858
BCP	0.9926	0.9858	0.9902
TCP	0.9816	0.9881	0.9804
Ovary			
TET	0.9930	0.9908	0.9883
UNM	0.9925	0.9894	0.9895
BCP	0.9947	0.9946	0.9957
TCP	0.9957	0.9937	0.9955
Stigma			
TET	0.9903	0.9927	0.9895
UNM	0.9925	0.9909	0.9948
BCP	0.9923	0.9877	0.9889
TCP	0.9907	0.9915	0.9935
Pollen			
MPG	0.9770	0.9782	0.9795

Table 3-1. Pearson correlation coefficient summary among biological replicates for different microarray experiments. REP, replicate; TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

	ANTHER			OVARY			STIGMA			POLLEN		
STAGE	<i>Rep 1</i>	<i>Rep 2</i>	<i>Rep 3</i>	<i>Rep 1</i>	<i>Rep 2</i>	<i>Rep 3</i>	<i>Rep 1</i>	<i>Rep 2</i>	<i>Rep 3</i>	<i>Rep 1</i>	<i>Rep 2</i>	<i>Rep 3</i>
TET	23,478	22,168	22,955	21,789	22,364	22,681	21,279	21,879	21,976	-	-	-
UNM	21,298	26,135	25,160	22,452	25,790	22,420	24,679	24,586	25,143	-	-	-
BCP	27,441	26,473	25,984	26,964	27,448	27,080	26,468	26,188	25,387	-	-	-
TCP	19,133	19,530	19,861	27,070	26,242	27,087	23,570	24,274	24,176	-	-	-
MPG	-	-	-	-	-	-	-	-	-	12,942	12,796	13,490

Table 3-2. Number of probe sets showing expression in individual biological replicates (Rep) for each reproductive organ (anther, stigma, ovary) at four stages of development and for mature pollen. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

coefficients, the number of expressed probe sets among replicates within a sample was very similar.

An individual probe set can be considered expressed within a microarray experiment if one (1P), two (2P), or three (3P) present calls are assigned to it for the three biological replicates. For example, the 1P criterion implies that for a particular probe set the data shows the expression to be absent (A) in two of the three replicates (ie. 1P/2A), this represents the least stringent criterion for establishing that this probe set is expressed. Table 3-3 shows a comparison between the results obtained when the 2P or 3P criteria (ie. majority of the replicates have a present call for a particular probe set) are applied. As expected, the number of probe sets obtained using the 2P criterion is higher than that for the more rigorous 3P requirement, however, the fact that the difference between these two datasets is relatively small also reflects the quality of the data for the individual replicates. The 2P and 3P data was further filtered to show probe sets having a $\log_2$ expression value ≥ 7 (Table 3-3), which is equivalent to a hybridization signal value of 128 and represents a proportion of the dataset whose low expression could be considered to have a certain degree of uncertainty. The overall average of the data (ie. 13 datasets) shows that 35% of the 2P and 27% of the 3P probe sets show a level of expression < 7 . The difference between the numbers of expressed probe sets when comparing the 2P and 3P data ≥ 7 is low (1.2-8.3 %), indicating probe sets expressing at this level tend to be reliable. In this study, a conservative approach was taken in order to limit the number of false positives, therefore most of the data presented corresponds to the 3P criterion.

Tissue	2P	2P ≥ 7	3P	3P ≥ 7
Anther				
TET	22,352	14,809	19,467	14,443
UNM	23,869	15,840	19,796	15,366
BCP	26,235	15,941	23,025	15,596
TCP	18,803	12,294	15,491	11,647
Pollen				
MPG	12,274	7,816	9,549	7,166
Ovary				
TET	21,737	14,728	19,081	14,387
UNM	22,921	16,203	19,875	15,808
BCP	26,773	16,380	23,788	16,181
TCP	26,376	16,207	23,488	15,933
Stigma				
TET	21,199	14,387	18,531	13,994
UNM	24,340	15,717	21,396	15,390
BCP	25,668	15,971	22,374	15,650
TCP	23,396	15,550	20,577	15,225

* "P" denotes 'present' call for expression

Table 3-3. Summary of all the probe sets expressed in at least two out of three (2P) or all three (3P) biological replicates, and the number of probe sets expressing at a cutoff value of ≥ 7 based on a $\log_2$ scale. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

4. Transcriptome Analysis of Triticale Anther Development and Mature Pollen

4.1 Results

4.1.1 Microscopic analysis of anther and pollen development

The four stages of pollen development were identified by staining the nuclei (Figure 4-1A) and these represent the tetrad stage (TET) where four microspores are enclosed within a callose wall, the uninucleate microspore (UNM) representing free microspores with a single nuclei, the bi-cellular pollen (BCP) corresponding to the pollen with a vegetative and generative cell following pollen mitosis I (PM I), and the tri-cellular pollen (TCP) containing the vegetative cell's nucleus and two sperm nuclei post pollen mitosis II (PM II). Light microscopy of the corresponding developmental stages of the floret is shown in Figure 4-1B.

Anther cross sections were also observed using cryo-scanning electron microscopy (SEM, Figure 4-1C). At the earliest stage, the anther locule consists of four distinct wall layers surrounding the developing pollen mother cells: from outer to inner layer these correspond to the epidermis, endothecium, middle layer, and tapetum (Figure 4-1C, TET). As the anther develops, the initial wall layers begin to rearrange. The parietal middle layer is frequently destroyed by compression at the free uninucleate microspore stage and is no longer visible (Figure 4-1C, UNM), while the tapetal cells begin to degrade after the first pollen mitosis (Bhandari & Khosla, 1981). The epidermis and endothecium continue to elongate and thin out up until pollen maturation, where the anther is bordered only by the two remaining layers (Figure 4-1C, BCP, TCP). These observations are in concordance with results from a previous study on triticale anther ontogeny (Bhandari & Khosla, 1981). Once released from the tetrads, the uninucleate

Figure 4-1. Microscopy of triticales reproductive tissues at four stages of pollen development. (A) Aniline blue and DAPI stained microspore/pollen; (B) light microscopy of developing floret; and (C) cryo-SEM sections of anthers during development with pollen exposed. Ep, epidermis; En, endothecium; ML, middle layer; T, tapetum; PMC, pollen mother cell; Gp, germinating pore; TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular; TCP, tri-cellular pollen; MPG, mature pollen grain. Scale bar = 10 μ m (top, panel A), 1 mm (middle, panel B).

microspores begin to enlarge and become round in shape, as well as highly vacuolated, causing the pollen grains to be appressed to the tapetum. This forms depressions into the tapetal orbicular wall and aligns the microspores as a single layer across the periphery of the anther locule (Figure 4-1C, UNM-TCP), an orientation predominant in the Poaceae family (Christensen & Horner, 1974; Kirpes et al., 1996). This particular arrangement causes lateral pressure to be exerted onto adjacent pollen grains, which gives them a flattened-obovate shape (Figure 4-1, BCP-TCP). The pollen germinating pore (Gp) was continuously oriented towards the tapetal surface and stays in direct contact with the tapetum up until maturity. Once mature, the round pollen grains are released upon anther dehiscence (Figure 4-1C, MPG).

4.1.2 Expression profiling throughout anther development and in mature pollen

A summary of the experimental design and scheme of the microarray data analysis is illustrated in Figure 4-2. After removal of reporter and control probe sets from the Wheat GeneChip[®], 61,115 probe sets remain representing 55,052 transcripts. Of these, 27,137 were expressed in at least one stage of anther development or the mature pollen grain covering 44.4% of the targets on the array. As shown in Figure 4-3, a greater number of transcripts were expressed early in anther development relative to stages following pollen mitosis II (PM II). The TET and UNM stages have a very similar number of expressed genes. The overall trend shows an increase in expressed probe sets from the TET stage to the BCP stage followed by a significant decrease at the TCP stage of anther development. Compared to the anther, the mature pollen grain showed a drastic reduction in expressed probe sets.

A

B

Figure 4-2. Schematic diagram of microarray data analysis for the developing anther and mature pollen grain. A) Workflow for finding differentially expressed probe sets (DEPs) and (B) tissue-enriched and -specific probe sets. Analysis performed for the 3P+2P (at least 2 Present calls) and 3P datasets, only intersecting probe sets are relevant as they represent 3P data, but also consider 2P values. Number in parentheses indicate the number of different sample arrays used. AN, anther at 4 different stages of development; MP, mature pollen; RMA, robust multichip average; ANOVA, analysis of variance; DEPs differentially expressed probe sets.

Figure 4-3. Number of probe sets expressed (present calls in all replicates, 3P) at each stage of anther development and in mature pollen. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

To further describe and understand the relationship between the five anther/pollen transcriptomes, a dartboard model was used to visualize the extent of expression overlap among these transcriptional profiles (Figure 4-4). As mentioned above, a total of 27,137 probe sets are expressed in at least one anther stage or the mature pollen. The largest number of overlapping probe sets is found in the region depicting probe sets shared by all samples, including MPG, and consisted of 6,921 probe sets. This indicates that 25.5% of all male expressed probe sets are found throughout the entire development of the pollen grain. When looking at the anther data only, it includes the next two largest regions of overlap, TET/UNM/BCP/TCP (4,585; 16.9%) and TET/UNM/BCP (4,237; 15.6%), which share the first three developmental stages of anther development. The latter 3-way comparison differs markedly from the other possible 3-way comparisons of developing anther stages, UNM/BCP/TCP (700; 2.6%), TET/BCP/TCP (238; 0.9%), and TET/UNM/TCP (36; 0.1%), demonstrating again the close relationship of the first three consecutive anther developmental stages, and the variability associated with TCP. The number of probe sets expressed in only two anther developmental stages ranges from 42 (UNM/TCP) to 1,367 (UNM/BCP). Groups with the lowest number of overlapping probe sets were TET/UNM/TCP/MPG (12), TET/UNM/MPG (12), UNM/MPG (14), and UNM/TCP/MPG (18) and were all characterized by shared samples from both early and late pollen development, and the absence of the BCP stage. There were a significant number of probe sets (17%) whose expression was limited to a single developmental anther stage or to the mature pollen grain (although this does not exclude expression in other plant tissues). The results were as follows: TET (1,100; 4%), UNM (705; 2.6%), BCP (1,940; 7%), TCP (635; 2.3%), and MPG (316; 1.2%), although they do not

Figure 4-4. Dartboard model showing the relationship between the 27,137 probe sets expressed during anther development and in mature pollen. Numbers shown represent the number of expressed probe sets in individual or overlapping samples. Total number of expressed probe sets at each stage: Te: 19,467, U: 19,796, B: 23,025, Tr: 15,491, P: 9,549. Te, tetrad; U, uninucleate microspore; B, bi-cellular pollen; Tr, tri-cellular pollen; P, mature pollen grain.

appear to show a particular tendency, except perhaps the numbers are to some extent smaller post PM II.

Pairwise comparisons of the co-expressed transcriptomes from different stages of anther development and the mature pollen grain are represented as scatter plots in Figure 4-3. As mentioned above, the TET and UNM have a similar number of expressed genes (Figure 4-3), and upon comparison of the two datasets a relatively modest relationship ($R^2 = 0.6318$) between the two was observed (Figure 4-5A). Despite the fact that considerably more probe sets are expressed at the BCP stage (Figure 4-3), the UNM and BCP were among the most highly correlated anther transcriptomes ($R^2 = 0.7906$, Figure 4-5E) and this is in close agreement with the overlapping probe sets data (Figure 4-4). The correlation between the TET transcriptome and that of later stages decreased sequentially as the anther matured. When comparing to the other stages of the developing anther only, the TCP transcriptome always shows the most variation (Figure 4-5C, F, H).

MPG shares the least similarity to the other transcriptomes as predicted by the low number of expressed genes. This is also in concordance with the dartboard data, where the least number of overlapping probe sets always included MPG (Figure 4-4). However, it does have a relatively good correlation with TCP ($R^2 = 0.6945$, Figure 4-4J) and this is attributed to the fact that TCP anther contains nearly mature pollen grains.

4.1.3 Co-regulated gene expression in the developing anther and mature pollen

Cluster analysis was performed to study the co-regulation of the 27,137 genes expressed in the developing anther and mature pollen grain using the Acuity 4.0

Figure 4-5. Scatter plots showing the correlation of co-expressed probe sets at different stages of anther development and/or pollen with the linear correlation (R^2) indicated. Pairwise transcriptome comparison between (A) TET versus UNM, (B) TET versus BCP, (C) TET versus TCP, (D) TET versus MPG, (E) UNM versus BCP, (F) UNM versus TCP, (G) UNM versus MPG, (H) BCP versus TCP, (I) BCP versus MPG, and (J) TCP versus MPG are shown. The linear regression is represented as a purple line. The middle green line represents probe sets expressed at the same level in both datasets whereas the two outer lines represent probe sets with at least a two-fold change in intensity value in one of the datasets. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

microarray analysis software package (see Materials and Methods). Based on gap statistical analysis (Tibshirani et al., 2001), the optimal number of clusters for this dataset is $k = 45$. The expression profiles are represented as color maps and the overall expression trend is presented as a line graph for each cluster (Figure 4-6). The results from this clustering analysis allowed the identification of specific gene expression patterns. For example, probe sets that express early in anther development (TET – BCP) were found in clusters 8, 29, 37, 45. Early expressing probe sets could be further subdivided into probe sets predominantly expressed in TET and UNM (cluster 33), TET and BCP (cluster 12, 25, 41), or UNM and BCP (clusters 34). The probe sets expressed post PM II, the so-called ‘late’ expressing genes (TCP – MPG), were found predominantly in cluster 32. Some clusters showed probe sets whose expression was primarily in a single anther stage or in mature pollen, for example, TET (clusters 4, 17), UNM (cluster 38), BCP (clusters 5, 18), TCP (cluster 44), and MPG (cluster 20). Other clusters showed patterns such as relatively high expression throughout all the stages (eg. clusters 1, 30) or expression in all stages except one (eg. clusters 6, 7, 10, 16, 19, 21, 27, 35, 39, 40, 43).

4.1.4 Identification of differentially expressed genes in the developing anther and mature pollen

An analysis of variance (ANOVA) statistical test was performed across all anther and pollen samples in order to identify differentially expressed probe sets (DEPs). This analysis reduced the original dataset of 27,137 probe sets expressed in the anther and pollen to 25,813 probe sets, indicating that the majority of probe sets display differential expression between at least two samples at an adjusted P value < 0.05 . In effect, all

Figure 4-6. Graph and color map for an unsupervised *K*-means clustering result of 27,137 anther and pollen expressed probe sets using the Acuity 4.0 software. Number displayed above each cluster represents the number of probe sets for each group. Stage of development indicated at the top and the line graph shows overall trend for each cluster. Similarity metric: Euclidean Squared, $k = 45$ (number of clusters). TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

probe sets that showed no significant differential expression throughout all samples were removed. The dataset of 25,813 probe sets was further analyzed to identify probe sets that showed an increase in expression by an absolute fold-change (FC) of at least 2 and only these were considered to be differentially expressed (Figure 4-7, top table). For comparison, DEPs were further filtered to include only probe sets with an expression value ≥ 7 ($\log_2$) representing an absolute signal value of 128 (Figure 4-7, bottom table).

Generally, the further apart the anther developmental stages, the higher the number of DEPs (Figure 4-7, top table), and MPG has lower numbers of up-regulated genes overall especially when compared to TCP (1,330). The values highlighted in red (Table 4-6) represent the numbers of up-regulated DEPs across consecutive stages, demonstrating essentially a gradual decrease in the number of up-regulated probe sets as the anther and pollen mature. However, after PM II a slightly higher number of DEPs is observed (TCP vs BCP) that may be attributed to the fact that UNM and BCP have very similar transcriptomes (Figure 4-4E). The above trends are maintained with the datasets composed of probe sets with an expression value cutoff of ≥ 7 (Figure 4-7, bottom table). Furthermore, approximately 75% of the DEPs are ≥ 7 and therefore in agreement with data in Table 3-3.

An analysis was also carried out to identify probe sets whose expression is down-regulated at each successive stage of development (Figure 4-8). Generally, as with up-regulated genes, the further apart the developmental stages the higher the number of down-regulated DEPs, although this now includes the MPG. The DEPs highlighted in green (Figure 4-8) represent probe sets that are repressed with respect to the preceding

	TET	UNM	BCP	TCP	MPG
TET	—	5,398	6,898	5,497	3,329
UNM		—	3,700	4,640	3,284
BCP			—	3,725	2,760
TCP				—	1,330
MPG					—

	TET	UNM	BCP	TCP	MPG
TET	—	4,140	4,538	4,537	2,841
UNM		—	2,126	3,807	2,782
BCP			—	3,103	2,353
TCP				—	982
MPG					—

Figure 4-7. Differentially expressed probe sets (DEPs) up-regulated throughout anther development and in mature pollen. Top table: probe sets up-regulated at least 2-Fold Change (2-FC) at an adjusted P-value < 0.05; bottom table: probe sets up-regulated at least 2-FC, P-value < 0.05, and expression value ≥ 7 ($\log_2$). Stages labeled on the top have at least a 2-FC increase in expression over the stages indicated on the left hand side. Numbers highlighted in red indicate up-regulated probe sets between consecutive developmental stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

stage of anther development. The observed trends also apply to the dataset consisting of probe sets down-regulated from an original expression value ≥ 7 (Figure 4-8, bottom table). A significant increase in the number of repressed probe sets occurs post PM II and this is especially true within the MPG. There is at least twice the number of down-regulated probe sets after PM II in comparison to the earlier stages. As a result, with both up- and down-regulated data, the pre-mitosis II stages (TET, UNM, BCP) show the least variation in expression or in other words, the highest number of genes whose expression does not change significantly among the three stages.

4.1.5 Transcripts enriched and specific to the developing anther and mature pollen

An in-depth analysis was performed to identify genes that were preferentially expressed (enriched) in the anther and mature pollen (Figure 4-9). To find enriched genes, anther- and pollen-expressed probe sets were compared to that of the female reproductive organs (stigma and ovary) at four stages of development (TET, UNM, BCP, TCP), as well as to an additional Wheat GeneChip[®] dataset of 16 other triticale tissues including leaf tissue (at seedling, tillering, early boot, early and late senescence), coleoptile, root, stem tissue (throughout development based on Zadoks' scale Z21, Z45, Z59, Z75), and seed tissues (embryo, epidermis, endosperm, pericarp, and crease), totaling 24 non-anther/pollen microarray datasets. Only anther or pollen transcripts showing at least a 2-fold up-regulation over all other tissues were considered enriched (see Materials and Methods).

We found a total of 5,032 probe sets whose expression was enriched in the different anther stages and/or in mature pollen, with the highest proportions being found

	TET	UNM	BCP	TCP	MPG
TET	—	4,356	4,617	10,939	14,427
UNM		—	2,588	10,880	14,777
BCP			—	11,946	16,764
TCP				—	8,196
MPG					—

	TET	UNM	BCP	TCP	MPG
TET	—	3,283	3,718	8,506	11,119
UNM		—	2,137	8,517	11,749
BCP			—	7,964	11,668
TCP				—	6,133
MPG					—

Figure 4-8. Differentially expressed probe sets (DEPs) down-regulated throughout anther development and in mature pollen. Top table: probe sets showing at least a 2-Fold Change (2-FC) down-regulated at an adjusted P-value < 0.05; bottom table: probe sets down-regulated at least 2-FC from an original expression value ≥ 7 ($\log_2$), P-value < 0.05. Stages labeled on the top have at least a 2-FC decrease in expression over the stages indicated on the left hand side. Numbers highlighted in green indicate down-regulated probe sets between consecutive developmental stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

post PM II (Figure 4-8). In particular, even though MPG shows the least number of expressed probe sets, it displays the highest number of enriched probe sets (Figure 4-9). When expressed as percentages, the TCP and MPG Enriched probe sets make up 9.6% and 17.8% of the total number of expressed probe sets for each stage, respectively, whereas all stages prior to PM II only make up a combined total of 8.6%.

As mentioned above, the Enriched dataset includes 5,032 probe sets showing enrichment in at least one stage of pollen development and therefore includes probe sets whose expression is recorded in more than one anther stage. When this redundancy is removed, the Enriched dataset actually represents a total of 2,982 individual probe sets and this number will be used in subsequent analyses. The Enriched dataset was further screened to identify Stage-Enriched probe sets (Figure 4-9), which represent anther/pollen Enriched probe sets that show a 2-fold enrichment within a particular stage when further compared to the other anther/pollen stages. For example, there are 388 anther/pollen enriched probe sets which are expressed at the TET stage, of those, 237 probe sets are enriched at the TET stage in comparison to the other stages and therefore considered Stage-Enriched (Figure 4-9). A total of 1,387 Stage-Enriched probe sets were identified. The sum of the Stage-Enriched probe sets does not equal the original number of non-redundant Enriched probe sets because it excludes those that show less than a 2-fold change among different stages or those that do show 2-fold change but occur in more than one stage. The percentages of non-redundant Enriched probe sets that are also Stage-Enriched in one anther stage or mature pollen are: TET (7.9%), UNM (5.5%), BCP (10.2%), TCP (6.4%), and MPG (16.4%). This represents 46.4% of the total number of anther- and pollen-enriched genes with MPG representing the largest proportion followed

Figure 4-9. Anther- and mature pollen-expressed, -enriched, and -specific probe sets throughout development. Enriched probe sets have at least a 2-fold higher expression over all other triticale tissues (leaf tissue at seedling, tillering, early boot, early and late senescence; coleoptile; root; stem tissue throughout development based on Zadoks' scale Z21, Z45, Z59, Z75; embryo; epidermis; endosperm; pericarp; crease; ovary and stigma at TET, UNM, BCP and TCP stages). Stage-Enriched probe sets are probe sets that are also enriched in a particular stage compared to other anther stages or mature pollen. Specific probe sets correspond to Enriched probe sets with no expression in non-anther/pollen tissues. Stage-Specific probe sets are Specific probe sets that show no expression in other anther stages or mature pollen. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

by BCP (Figure 4-8). Interestingly, this finding is not in agreement with previous results where we found the largest proportion of enriched probe sets occurring post PM II.

Lastly, to discriminate probe sets that are specifically expressed in the anther or pollen as opposed to preferentially expressed, we filtered the Enriched dataset further. Enriched probe sets were considered specific when present in the anther or pollen and absent (as opposed to a 2-fold enrichment) in all the other tissues mentioned above (see Materials and Methods). A total of 1,455 probe sets displayed specific expression in at least one stage of anther development or mature pollen, which represents 778 individual anther- and pollen-specific genes (Figure 4-9). The 3 stages following PM I have more than twice the number of specific genes when compared to the two earliest stages. In terms of proportions, the two stages post PM II (TCP and MPG), have the highest proportion of expressed probe sets that are also considered specific with 2.6% and 4.2% respectively, whereas the three stages prior to PM II have a combined total of 3%. From this dataset, Stage-Specific probe sets were identified (Figure 4-9) and they correspond to probe sets that are not only anther- or pollen-specific, but are also specific to a particular stage (ie. not expressed in the other stages). The number of Stage-Specific genes was: TET (29), UNM (31), BCP (96), TCP (28), and MPG (51). Each Stage-Specific group contains probe sets also present in clusters showing stage exclusive expression: TET (cluster 4, 17), UNM (cluster 38), BCP (clusters 5, 18), TCP (cluster 44), and MPG (cluster 20), validating these results (Figure 4-6). However, the number of Stage-Specific probe sets presented is considerably lower than the numbers from the cluster analysis and this reduction is due to the fact that the Stage-Specific probe sets result from a

comparison including 24 other tissues, whereas in the cluster analysis only anther-expressed genes were being analyzed.

4.1.6 Characterization of anther- and pollen-specific genes

A total of 778 probe sets were specific to the male reproductive tissues and these were divided into three groups for analyses: (1) anther-only: 379 probe sets specifically expressed in the developing anther but not in the mature pollen grain, (2) anther + pollen: 348 probe sets specifically expressed in the developing anther but also expressed in the mature pollen grain, and (3) pollen-only: 52 probe sets expressed in the mature pollen grain but not the anther. In theory, group (1) should be enriched for the anther wall, tapetum, and early microspore genes, group (2) should be enriched for late pollen genes, while group (3) should be enriched for genes that are activated at the very last stage of pollen maturation. For each group, we assigned annotations and gene ontology (GO) categories for biological process using the web-based program Blast2GO (see Materials and Methods). Probe sets with no significant matches were classified as ‘unknown’, whereas probe sets that showed a match to an existing protein or domain but did not have an associated annotation GO ID for biological process were considered ‘unclassified’. The distributions of the different biological processes for the 3 different anther/pollen-specific probe sets are summarized in Figures 4-10, 4-11, and 4-12 respectively. It is important to realize that proteins can fall in more than one biological process and as a result some redundancy will occur in their distribution.

Almost half (47%) of the anther-only probe sets fell into the unknown and unclassified categories (Figure 4-10). The most highly represented category for anther-only probe sets was transport, while the next largest groups included carbohydrate,

Figure 4-10. Distribution of biological process categories for anther-only probe sets (anther/pollen-specific probe sets not expressed in MPG).

protein, and general metabolic processes. Regulation of cellular processes was also a significant category. The transport category was made up of a variety of genes encoding proteins that could transport metal ions, protons, oligopeptides, amino acids, and lipids. A large proportion was involved in lipid transport and 8 lipid transfer proteins were identified specifically in the anther. The metabolic process category contained a highly varied group of enzymes which often occurred only once within the group. Exceptions included two peroxidases, two 5-phosphoribosyl-5-aminoimidazole synthetases, and 4 cytochrome P450 proteins.

The protein metabolic process group was largely represented by various types of proteases and protein kinases. The proteolytic enzymes included 2 carboxypeptidase d, a prolylcarboxypeptidase, a serine carboxypeptidase, 2 serine proteases, 2 serine carboxylases, 2 cysteine proteases, an aspartic protease, and a subtilisin-like protease. The protein (de)phosphorylation group included 2 receptor-like kinases, a TAK33 receptor-like kinase, 3 serine/threonine kinases, an EF-hand calcium-binding containing domain serine/threonine kinase, a serine-threonine/tyrosine kinase, and a cdc 2 kinase. It also included a serine/threonine phosphatase, a phosphatase 2C, a metallo-dependent phosphatase, as well as a protein phosphatase 4 core regulatory subunit R2. The carbohydrate metabolic process category was comprised of anther-specific glycosyl hydrolases, a β -amylase, a β -galactosidase, polygalacturonases, and a β -glucanase, for example. The cellular process group represented many important functions related to male reproductive development. Many genes identified were involved in pollen development including exine formation for which raftin genes appeared often. Other proteins which might be involved in exine synthesis include a cinnamoyl-CoA reductase,

a chalcone synthase-like enzyme, β -propeller containing proteins, a strictosidine synthase also known as *LAP3* (Dobritsa et al., 2009), and the previously mentioned cytochrome P450 proteins. A considerable portion of this group is represented by cell wall modifying proteins including proline-rich extensins, a pectate lyase, a pectinesterase, pectinesterase inhibitors, polygalacturonases, and cell wall invertases. Some of these enzymes also appeared in the carbohydrate metabolic process category thus illustrating the redundancy previously mentioned.

Lastly, because they often play key roles in development, the groups representing transcription regulation and protein (de)phosphorylation will be presented briefly. Six transcription factors were found to be expressed in the anther-only. They consisted of a myb, B3 domain and 4 basic-helix-loop-helix (bHLH) containing transcription factors. The protein phosphorylation group was made up of 7 different protein kinases including two receptor-like protein kinases, a calcium-dependent kinase, and three serine/threonine protein kinases. The protein dephosphorylation group included a protein phosphatase 2C, a serine/threonine protein phosphatase, and a metallo-dependent phosphatase.

The top fifty highest expressors in the anther-only specific dataset are summarized in Table 4-1. It is striking how many of these genes fall under the unknown or unclassified categories. For those that are classified, no single biological process predominates within this group of highly expressed anther-only genes, although there is a small bias towards transport. There are 3 raftin genes, two cysteine proteases, and two cell wall invertases.

	Probe Set ID	MAX	Description		Probe Set ID	MAX	Description
1	Ta 4313 1 S1_at	TCP	cysteine proteinase	26	Ta 11690 1 S1_at	BCP	knox 2
2	Ta 3812 2 A1_a_at	UNM	18S ribosomal RNA	27	TaAffx 106410 2 A1_at	BCP	prolylcarboxypeptidase
3	TaAffx 114178 1 S1_s_at	UNM	metabolic process	28	Ta 10483 1 A1_a_at	BCP	unclassified
4	TaAffx 87253 3 S1_at	BCP	unknown	29	Ta 4401 1 S1_at	BCP	unknown
5	Ta 4136 1 S1_at	BCP	unknown	30	Ta 23687 2 S1_x_at	UNM	18S ribosomal RNA
6	Ta 3781 1 A1_a_at	BCP	unclassified	31	TaAffx 93263 1 A1_at	UNM	nodulin-26 precursor
7	Ta 4176 1 S1_x_at	BCP	raftin1a	32	Ta 20651 1 S1_at	TET	peroxidase 40
8	Ta 23091 1 S1_at	TET	ATPase	33	TaAffx 87185 1 S1_at	UNM	membrane protein
9	Ta 3812 1 A1_a_at	UNM	unclassified	34	TaAffx 9640 1 S1_s_at	UNM	unknown
10	Ta 28807 1 S1_a_at	TET	unknown	35	Ta 3956 1 A1_at	BCP	drug/metabolite transporter
11	Ta 22377 1 S1_at	BCP	unknown	36	Ta 3791 1 S1_at	TCP	cysteine protease
12	Ta 30685 2 S1_x_at	BCP	10 kDa secreted protein	37	Ta 3849 1 S1_at	TCP	pectin methylesterase inhibitor
13	Ta 25940 1 A1_at	BCP	unclassified	38	Ta 30400 1 A1_at	TET	bHLH transcription factor
14	Ta 23687 1 S1_at	BCP	unclassified	39	TaAffx 32344 1 S1_at	BCP	unknown
15	Ta 28854 1 S1_at	BCP	gibberellin-regulated	40	Ta 4195 1 S1_at	UNM	raftin1a
16	Ta 20686 1 A1_at	UNM	unknown	41	TaAffx 12411 1 S1_at	BCP	unclassified
17	TaAffx 9588 1 S1_x_at	BCP	unclassified	42	Ta 4046 1 S1_a_at	UNM	unknown
18	Ta 4272 1 S1_at	BCP	unknown	43	Ta 4100 1 S1_at	BCP	unknown
19	Ta 686 1 S1_x_at	BCP	unknown	44	Ta 4322 1 A1_at	BCP	cell wall invertase
20	Ta 10359 1 S1_at	BCP	hydrolase	45	Ta 198 1 S1_at	BCP	cell wall invertase
21	Ta 4091 1 S1_x_at	UNM	raftin1a	46	Ta 4212 1 A1_at	BCP	cytochrome P450
22	TaAffx 59661 1 S1_at	UNM	galactosyltransferase	47	Ta 23250 1 S1_at	BCP	unknown
23	Ta 15373 1 S1_at	BCP	unclassified	48	TaAffx 65977 1 A1_at	TET	unknown
24	Ta 4411 1 S1_at	UNM	LTP family protein	49	Ta 4037 1 S1_at	BCP	unknown
25	Ta 11496 1 A1_at	TCP	cyamine dioxygenase	50	Ta 648 1 S1_a_at	TET	unknown

Table 4-1. The 50 most highly expressed anther-only specific probe sets. Stage with highest expression for each probe set is shown under 'MAX'. Top 50 expression values range from 14.15 – 10.88 based on log₂ scale. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

Close to 48% of the 348 anther + pollen probe sets fell into the unknown and unclassified categories (Figure 4-11). The distribution of the different biological processes for the anther + pollen probe sets was somewhat similar to that of the anther-only probe sets with some significant trends. The anther + pollen probe sets showed a considerable increase in the carbohydrate metabolic process category, and substantial decreases in the transport, as well as the lipid and protein metabolic process categories. The anther + pollen carbohydrate metabolic process category included 4 β -galactosidases, 2 glycosyl transferases, 2 glycosyl hydrolases, 2 glucan endo-1,3- β -glucosidases, 1 sucrose phosphate synthase, 1 α -amylase, 1 carbohydrate phosphorylase, as well as carbohydrate and sucrose transporters. Furthermore, this category comprised numerous cell wall modifying proteins including 2 pectin esterase, 10 pectin esterase inhibitors, 3 pectate lyases, and one polygalacturonase.

Although the metabolic process category remained similar, the composition of the anther + pollen set showed practically no overlap with the anther-only set, with the exception of peroxidases. However, this category was still made up of a wide variety of proteins, mostly enzymes, and re-occurrence was observed only with peroxidases (3), soluble inorganic pyrophosphatases (2) and Tri a 4 proteins (3) which are believed to have oxidoreductase activity. The anther + pollen and the anther-only groups showed considerable similarity in their composition of cell wall associated enzymes, although the former lacked cell wall invertases. The anther + pollen also lacked many of the proteins described to be involved in exine synthesis in agreement with this group being enriched in anther genes expressed later in development. However, the anther + pollen group showed a considerable enrichment in proteins that might play a role in vesicular transport

Figure-4-11 Distribution of biological process categories for anther + pollen probe sets (anther/pollen-specific probe sets not expressed in MPG).

such as 3 clathrin assembly proteins, a synaptobrevin-like protein, an Exo70 protein, and a dynamin-like protein. The 3 transcription factors that were specific to the anther + pollen group were different from those found in the anther-only. They consisted of an AP2 and two MYB transcription factors. The protein phosphorylation/dephosphorylation category consisted of 4 serine/threonine protein kinases, 4 calcium-dependent protein kinases, but no phosphatases.

The top 50 annotated anther + pollen specific genes are summarized in Table 4-2. They include a significant number of genes involved in carbohydrate metabolism, 4 protein kinases, and several known pollen allergens.

We identified 52 mature pollen-only probe sets and their relative distribution among different biological processes is summarized in Figure 4-12. Nearly two thirds of the probe sets (60%) were unknown or unclassified. The remaining probe sets represented a variety of genes encoding a cytochrome P450 protein, an AP2 transcription factor, a leucine-rich repeat protein, an argonaute protein, and two disease resistance proteins. Not surprisingly given the small number of probe sets most are found in the top 50 most abundant mature pollen-only genes (Table 4-3).

	Probe Set ID	MAX	Description		Probe Set ID	MAX	Description
1	Ta 3717 1 A1_s_at	MPG	H ⁺ ATPase	26	Ta 22357 1 S1_s_at	TCP	unknown
2	TaAffx 95401 1 S1_at	MPG	polygalacturonase	27	Ta 3875 1 S1_at	TCP	unknown
3	Ta 12765 1 A1_at	MPG	protein kinase	28	Ta 30647 1 S1_at	MPG	major allergen Phl p 5
4	Ta 25610 1 S1_at	MPG	ralf precursor	29	Ta 8249 1 A1_at	MPG	Ca ⁺ dep protein kinase
5	Ta 16069 1 S1_s_at	BCP	unknown	30	Ta 10427 1 S1_x_at	TCP	actin-depolymerizing factor
6	Ta 3767 2 S1_x_at	TCP	polygalacturonase	31	Ta 25912 1 A1_at	MPG	unknown
7	Ta 28077 1 S1_x_at	TCP	unknown	32	Ta 4139 1 S1_s_at	TCP	major allergen Phl p 5
8	Ta 4344 1 A1_at	TCP	actin	33	TaAffx 122596 1 S1_at	MPG	beta-glucosidase
9	Ta 11712 1 A1_at	MPG	unknown	34	TaAffx 119713 1 S1_at	BCP	unknown
10	Ta 4081 1 S1_at	MPG	unclassified	35	Ta 28020 1 S1_s_at	TCP	pollen allergen (group 2)
11	Ta 14414 1 S1_at	TCP	expansin EXB4	36	Ta 24617 1 S1_x_at	TCP	unknown
12	Ta 10100 2 S1_s_at	TCP	inorganic pyrophosphatase	37	Ta 480 1 S1_at	TCP	pollen allergen (group 2)
13	Ta 30686 1 S1_s_at	BCP	unknown	38	Ta 10256 3 S1_at	MPG	major allergen Phl p 5
14	Ta 3988 1 S1_s_at	TCP	unknown	39	Ta 10427 1 S1_at	TCP	actin-depolymerizing factor
15	TaAffx 128418 48 A1_at	MPG	glycoside hydrolase	40	TaAffx 65316 1 A1_at	MPG	Ca ⁺ dep protein kinase
16	Ta 10401 1 S1_a_at	MPG	unknown	41	Ta 3919 1 A1_at	MPG	zinc ion binding protein
17	Ta 10100 2 S1_a_at	MPG	inorganic pyrophosphatase	42	Ta 10100 2 S1_x_at	MPG	inorganic pyrophosphatase
18	Ta 16011 1 S1_at	BCP	unclassified	43	TaAffx 87016 1 S1_at	BCP	pollen allergen Phl p 11
19	Ta 10300 1 S1_at	TCP	unknown	44	TaAffx 23544 2 S1_at	MPG	phosphodiesterase
20	TaAffx 88655 1 S1_at	TCP	unknown	45	Ta 3837 1 A1_at	TCP	peroxidase
21	Ta 3702 1 A1_at	MPG	clathrin assembly protein	46	Ta 3966 1 S1_at	MPG	b12d protein
22	Ta 4320 1 S1_at	TCP	ascorbate oxidase	47	Ta 3913 1 S1_at	MPG	unclassified
23	Ta 10268 1 S1_at	TCP	pectinesterase inhibitor	48	Ta 10261 1 S1_at	TCP	unknown
24	Ta 10469 1 A1_at	MPG	GAMyb transcription factor	49	Ta 15997 1 S1_at	BCP	ATP synthase
25	Ta 10415 1 A1_at	MPG	protein kinase	50	TaAffx 128418 131 A1_s_at	MPG	glycoside hydrolase

Table 4-2. The 50 most highly expressed anther + pollen specific probe sets. Stage with highest expression for each probe set is shown under 'MAX'. Top 50 expression values range from 14.62 – 13.22 based on log₂ scale. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen, TCP, tri-cellular pollen; MPG, mature pollen grain

Figure 4-12. Distribution of biological process categories for pollen-specific probe sets.

	Probe Set ID	MAX	Description		Probe Set ID	MAX	Description
1	Ta.28507.1.S1_at	MPG	unknown	26	Ta.25001.3.A1_at	MPG	unknown
2	TaAffx.28543.1.S1_at	MPG	argonaute protein	27	TaAffx.54300.1.S1_at	MPG	pathogenesis protein
3	Ta.25193.1.S1_at	MPG	heat shock protein	28	TaAffx.50392.1.S1_at	MPG	zinc finger/unclassified
4	TaAffx.85068.1.S1_at	MPG	unclassified	29	Ta.21899.3.A1_x_at	MPG	unclassified
5	TaAffx.97425.1.S1_at	MPG	unknown	30	TaAffx.110344.1.S1_at	MPG	unknown
6	TaAffx.28590.1.S1_at	MPG	aldehyde dehydrogenase	31	TaAffx.119697.2.S1_at	MPG	unknown
7	TaAffx.57685.1.S1_at	MPG	unknown	32	TaAffx.51217.1.S1_s_at	MPG	RNA polymerase
8	TaAffx.93442.1.S1_at	MPG	unknown	33	TaAffx.113599.1.S1_at	MPG	unknown
9	TaAffx.93442.1.S1_x_at	MPG	unknown	34	TaAffx.49970.1.A1_at	MPG	unknown
10	TaAffx.9487.1.S1_s_at	MPG	agenet domain/unclassified	35	TaAffx.8907.1.S1_at	MPG	unknown
11	TaAffx.31662.1.S1_at	MPG	unknown	36	Ta.15650.1.S1_at	MPG	resistance protein rga2
12	TaAffx.53490.1.S1_at	MPG	cytochrome P450	37	TaAffx.38933.1.S1_at	MPG	unknown
13	Ta.8657.1.A1_at	MPG	unknown	38	Ta.3168.3.S1_at	MPG	unknown
14	TaAffx.70069.2.S1_at	MPG	aspartic peptidase	39	Ta.17664.1.S1_at	MPG	formate dehydrogenase
15	Ta.744.1.S1_at	MPG	unclassified	40	Ta.15396.1.S1_at	MPG	unknown
16	TaAffx.113821.1.S1_at	MPG	unknown	41	Ta.4404.1.S1_at	MPG	unknown
17	TaAffx.85980.1.S1_at	MPG	unclassified	42	TaAffx.4718.1.S1_at	MPG	unknown
18	TaAffx.122623.3.S1_at	MPG	unclassified	43	TaAffx.113454.1.S1_at	MPG	tsn1
19	TaAffx.64713.1.S1_at	MPG	oligopeptide transporter	44	TaAffx.107998.1.S1_at	MPG	unknown
20	TaAffx.29721.1.S1_at	MPG	vtc2-like protein	45	TaAffx.109945.1.S1_at	MPG	unknown
21	TaAffx.23431.1.S1_at	MPG	zinc finger/unclassified	46	Ta.15395.1.S1_at	MPG	unknown
22	Ta.26522.1.A1_at	MPG	decarboxylase	47	TaAffx.52195.1.S1_at	MPG	unknown
23	Ta.9427.2.S1_at	MPG	hga1/unclassified	48	TaAffx.50414.1.S1_at	MPG	unknown
24	Ta.21124.2.S1_at	MPG	ethylene-responsive element	49	TaAffx.12359.1.A1_at	MPG	transposase
25	Ta.9445.2.S1_at	MPG	ATPase	50	Ta.17741.1.S1_at	MPG	60s ribosomal protein

Table 4-3. The 50 most highly expressed pollen specific probe sets. Stage with highest expression for each probe set is shown under 'MAX'. Top 50 expression values range from 9.85 – 3.19 based on log₂ scale. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

4.2 Discussion

The goal of this study was to generate the most extensive and comprehensive transcriptome dataset on reproductive development in grasses to date, using triticale as a representative. The Affymetrix Wheat GeneChip[®] was used to profile both male (anther) and female (ovary, stigma) reproductive tissues at four stages of pollen development: tetrad (TET), uninucleate microspore (UNM), bi-cellular pollen (BCP), tri-cellular pollen (TCP), as well as mature pollen grain (MPG). The expression profiles generated should act as a fundamental gene expression resource for efforts to uncover the dynamic molecular mechanisms involved in reproductive development. Although we are using the most comprehensive wheat microarray chip available, it is important to keep in mind that it does not represent the complete transcriptome, and furthermore, that it was derived from sequences available in 2004. The design of the array was largely based on ESTs taken from public databases such as GenBank and is not representative of all transcripts at all stages of wheat development.

4.2.1 Anther and pollen demonstrate distinctive gene expression patterns

Since the developmental landmarks used to identify the stages at which the different tissues were harvested for microarray analysis relied on critical phases of pollen development, the anther transcriptome data should be the easiest to correlate with the cellular events occurring at those stages. When comparing the number of genes expressed at different anther developmental stages, a substantial decline in transcript numbers occurred late in anther development following PM II. In fact, mature pollen

grains displayed the smallest number of transcripts of all tissues and stages analyzed in this study. This is consistent with previous findings in rice where the least number of expressed probe sets in developing anthers was observed at the TCP stage (Fujita et al., 2010) and even lower numbers occurred in mature pollen grain (Wei et al., 2010). Similarly, mature anthers displayed the lowest number of expressed genes of any barley tissue examined (Druka et al., 2006), while mature pollen grain was not part of this study.

In the anther, lower numbers of transcripts at the TCP stage may reflect the shutdown of various cellular activities, such as cell differentiation events, prior to dehiscence. Since the pollen grain consists only of three cells and is essentially transcriptionally inactive at maturity (McCormick, 1993), it is not all that surprising that it would display a low number of expressed genes. It is widely speculated that many of the mature pollen transcripts are stored for downstream cellular activities such as pollen germination and tube growth which tend to have a high metabolic cost (Taylor et al., 1997). High transcript levels may also be required, for example, in wheat pollen the accumulation of carbohydrates, proline, and glycine betaine acts as a mechanism to prevent damage of proteins and of cell structures such as membranes during pollen dehydration (Schwacke et al., 1999; Pearce et al., 1976).

A global analysis of overlapping gene expression within the anther and mature pollen transcriptomes throughout development revealed that the largest group of common genes was found in the section that included all stages (TET/UNM/BCP/TCP/MPG). This group likely brings together both the common “housekeeping” and “anther-specific” genes required by anther and pollen throughout development. The next largest groups TET/UNM/BCP/TCP and TET/UNM/BCP, again probably include genes common to

early anther development while reflecting the significant changes in gene expression characteristic of the TCP and MPG mentioned above. The next largest group of overlapping expression is UNM/BCP and this is part of a recurring trend that will be discussed below. Generally, consecutive stages tend to have larger gene expression overlap. A similar tendency was observed in developing maize anthers. For example, a comparison of anther tissues at pre-meiotic, UNM, and BCP stages, as well as MPG revealed that 92% of the probe sets were shared during early anther development, whereas MPG demonstrated greater overlap with UNM/BCP as opposed to probe sets expressed earlier at the pre-meiotic stages (Ma et al., 2008).

When taking into consideration the similarity in the expression levels of the overlapping probe sets, the general trend was that the closer two stages were in the developmental sequence, the higher the similarity between transcriptomes. This is to be expected and also indicates a gradual divergence in gene expression as development progresses. Consequently, the most striking example is the comparison between the TET and MPG transcriptomes, which showed a correlation coefficient of 0.1007, thus demonstrating that these transcriptomes are essentially unrelated. This trend is also consistent with previous studies on the developing wheat anther (Crismani et al., 2006). Conversely, the UNM/BCP comparison showed the highest correlation ($R^2 = 0.7906$) among all stages compared, suggesting that the two stages share similar molecular functions in anther development. This was somewhat unexpected since the two transcriptomes display significant differences. For example, BCP showed 14% more expressed probe sets than UNM (Figure 4-3). Furthermore, BCP is comprised of a generative and a larger vegetative cell, whereas UNM is a single celled body. However,

the cytoplasm of the generative cell is greatly reduced and the cell nucleus is highly condensed, and therefore transcriptionally less active in comparison to the vegetative cell (McCormick, 1993). As a result, the male germline transcriptome is most likely masked by that of the larger vegetative nucleus and this may account for the persisting similarity between the UNM and BCP transcriptomes (Honys & Twell, 2004). Two studies comparing pollen development have also reported similar findings. The UNM and BCP transcriptome profiles were found to be the most highly correlated both in rice ($R^2 = 0.82$, Wei et al., 2010) and Arabidopsis ($R^2 = 0.96$, Honys & Twell, 2004). Our study demonstrates a slightly lower correlation and this may be partly due to the fact that we used whole anthers as opposed to isolated microspore/pollen populations, therefore our results likely also reflect changes in gene expression occurring in different anther cell types such as tapetal cells. In wheat, the tapetal cells begin to degenerate at UNM until only traces persist at TCP (Koonjul et al., 2005). Surprisingly, when rice microspore/pollen and tapetum cell types were characterized separately, 51% of the genes were expressed specifically in the male gametophyte, 5% were found exclusively in the tapetum, and 44% were expressed in both tissues (Suwabe et al., 2008). Furthermore, hierarchical clustering also revealed that the rice tapetum transcriptome at the UNM stage is most similar to the BCP stage (Tang et al., 2010). This indicates that the tapetum probably contributes only a small proportion of the discrepancies in the correlation between co-expressed probe sets of UNM/BCP.

Generally, more probe sets were up-regulated during early anther development, especially in the transition from TET to UNM, whereas more probe sets were down-regulated post PM II, especially in the transition from BCP to TCP. These trends are

likely correlated with the initiation and the end of anther differentiation respectively. Up-regulation of gene expression in late pollen development likely reflects the accumulation of the transcripts that are stored in MPG to be translated once it lands and hydrates on the stigma (Taylor et al., 1997). Not surprisingly, when compared to anther TCP (which contains nearly mature pollen), the MPG showed the lowest number of both up-regulated and down-regulated genes. Although a considerable number of probe sets are up-regulated in the transition from UNM to BCP, this transition shows the lowest number of down-regulated probe sets by far. One study specifically surveyed the number of probe sets that “appeared” and “disappeared” with respect to the previous stage in rice anthers at equivalent developmental stages (Fujita et al., 2010). Although comparison with the triticales data is difficult since it also includes probe sets that were present in the previous stage that changed in expression by 2-fold, the main difference was the significantly higher number of “disappearing” probe sets in the transition from UNM to BCP. This result is difficult to explain with the available information.

The trends in overlapping and in differential gene expression throughout anther development and in mature pollen grain were evident in the cluster analysis which grouped probe sets with similar developmental expression patterns. The majority of the anther probe sets are expressed in all stages, whereas the remaining probe sets display every combination of stage-specific expression. These results are consistent with previously published cluster analyses in *Arabidopsis* (Honys & Twell, 2004) and rice (Hobo et al., 2008), for example. Such cluster analyses will be useful in identifying genes specifically associated with particular cellular events (eg. cell division) or to obtain anther-specific promoters with specific expression patterns.

A total of 778 probe sets were identified as being specific to the male reproductive tissues with the majority occurring late in development thus reflecting an increased degree of specialization of these tissues. To simplify the analysis of the male reproductive tissue-specific probe sets, they were divided in three groups: “anther-only” which consisted of probe sets expressed specifically in the developing anther but not in the mature pollen grain, “anther + pollen” which includes anther-specific probe sets that are also expressed in the mature pollen grain, and “mature pollen-only” which are expressed exclusively in the mature pollen grain. As discussed below, this division resulted in very different data sets. Generally, about half of all the male tissue-specific probe sets fell into the unknown and unclassified categories indicating that considerable work is still required before a complete molecular picture of anther- and pollen-specific expression is obtained.

4.2.2 Anther-only gene expression

A total of 379 probe sets were found to be anther-specific but absent from the mature pollen transcriptome. This anther-only dataset should be enriched for transcripts occurring specifically within anther cells such as the tapetum or found only in early microspore development. GO process annotations were assigned to 187 anther-only probe sets and the categories believed to be most relevant to anther/microspore development will be discussed separately below.

4.2.3 Anther-only transport proteins

Transport was the most highly represented category of the anther-only gene set. Most of the genes in this category are expressed during early stages of development especially during the TET or UNM stage (data not shown). Transporters identified included a potassium transporter, a nicotianamine synthase, and a ZIP1 family transporter protein. Nicotianamine is a metal chelator known to play a role in the intracellular delivery of metals during plant reproduction (Takahashi et al., 2003), whereas ZIP1 transports zinc and has been identified in rice shoots, roots, and panicles (Chen et al., 2008). Several ATPases were found and many were H⁺ ATPases. These proteins have important roles in plant growth such as translocating metabolites for pollen development. Three nodulin-like proteins were identified one expressing exclusively at TCP and the other two were similar to nodulin 26 and expressed abundantly at UNM (data not shown). In rice, nodulin 26 has been reported to express at the TET-UNM stage and encodes an aquaporin protein functioning in transporting water or nutrients to the developing pollen mother cell through the tapetal layer (Fujita et al., 2010; Gomes et al., 2009).

We identified 8 anther-specific lipid transfer proteins (LTPs) making up the largest family within the transport category. Most express before the BCP stage and none express passed this stage. Many LTPs have been reported to be enriched in rice anthers (Fujita et al., 2010) and specifically within the tapetum (Lauga et al., 2000; Zhang et al., 2009). Five rice LTPs were identified whose expression peaked in anthers at the TET stage but were absent in MPG (Huang et al., 2009). Lipid transfer proteins are a group of highly conserved proteins that are non-specific transporters of lipid molecules (Yeats & Rose, 2008). They have a high affinity for lipid molecules due to the presence of a

hydrophobic cavity and are believed to participate in the transport of lipid precursors from the tapetum to the anther locule for sporopollenin/exine formation (Zhang et al., 2009). Pollen wall components are believed to be synthesized in the tapetum during early development and then gradually deposited to the surface of the microspore. One hypothesis postulates that the exine precursors are synthesized in the tapetum by linking acyl and ether bonds to the intermediates in the ER and then transferred to the locule via secretory vesicles stabilized by lipid transfer proteins (Huang et al., 2009).

4.2.4 Anther-only protein metabolism

Protein metabolism was another highly represented biological process in the anther-only dataset. The group was largely made up of proteases and protein kinases. The protein kinases will be discussed further below. The proteases identified consisted of a broad variety of enzymes including 4 carboxypeptidases, 2 cysteine proteases, and single examples of serine, aspartic, and subtilisin proteases. Proteases have been shown to be up-regulated in soybean pollen, many encoding for ubiquitin-specific types (Haerizadeh et al., 2009). Most of the proteases detected in our dataset were specifically abundant in mature anthers and are most likely transcripts accumulated for post-translation protein modifications during pollen tube growth (Haerizadeh et al., 2009). Cysteine and serine proteases have been identified in rice as having a role in pollen development and mutant analysis of a putative rice cysteine protease was shown to cause pollen degradation shortly after the microspore stage (Lee et al., 2004). Proteases play a diversity of roles ranging from protein catabolism and maturation, and it will be interesting to determine the roles the above proteases play in anther development. Two

ubiquitin carboxy-terminal hydrolase enzymes were also identified. They belong to the family of de-ubiquitinating enzymes responsible for generating free ubiquitin, recycling ubiquitin, and reversing the effects of ubiquitination (Smalle & Vierstra, 2004). This family of proteins has been shown to be crucial to pollen development in Arabidopsis (Doelling et al., 2007).

4.2.5 Anther-only cell wall biogenesis and modification

The regulation of cellular processes was another highly represented group in the anther-only collection and many of these genes were related to pollen cell wall development. Several putative raftin anther-specific transcripts were identified and these have also been reported in wheat and rice anthers (Wang et al., 2003; Li et al., 2006). In cereal crops, raftin is a structural protein of orbicules, also known as Ubisch bodies, and they have a secretory function that is essential for pollen maturation (Druka et al., 2006). Ubisch bodies have only been reported in monocot plants such as cereals and appear to be absent in dicots. RNA induced silencing of raftin in wheat caused microspore collapse and the inability of the tapetum to fully degenerate (Wang et al., 2003). It is believed that raftin is synthesized in the tapetum and packaged into Ubisch bodies for the transport of molecules such calcium to the pollen surface, it is also thought to play a role in tapetum degradation (Zhang et al., 2009; Wang et al., 2003). All the triticales anther-only raftin transcripts were expressed from the TET to the BCP stage (results not shown) consistent with their presumed tapetum-specific expression. Additional anther-only genes believed to play a role in pollen cell wall biogenesis included a cinnamoyl-CoA reductase (Grienenberger et al., 2010), a chalcone-synthase-like enzyme (Dobritsa et al., 2010), β -

propeller containing proteins (Dobritsa et al., 2009), a strictosidine synthase also known as *LAP3* (Dobritsa et al., 2009), and cytochrome P450 proteins (Li et al., 2010a). The presence of these proteins in the anther-only dataset is consistent with expression that is limited to early microspore development.

Many genes encoding cell wall modifying proteins and enzymes were found in the anther-only fraction and they included proline-rich extensins, a pectate lyase, a pectinesterase, pectinesterase inhibitors, and polygalacturonases. This is one of the groups of proteins, along with carbohydrate metabolism enzymes, that were shared with the anther + pollen group and these proteins will be further discussed below. Lastly, a mention should be made of the fact that two cell wall invertases were found within the anther-only fraction that did not occur in the anther + pollen fraction. Cell wall invertases are key enzymes involved in sugar metabolism and movement, and have often been shown to be critical to pollen development (Koonjul et al., 2005).

4.2.6 Anther-only transcription factors

A total of nine transcription factors were identified in the anther-only group. A B3 domain containing transcription factor was found and it belongs to a highly conserved family among higher plants with many different members involved in various plant processes including regulation of flower or reproductive development (Romanel et al., 2009). An anther calmodulin-binding transcription factor was also identified and they are believed to be calcium dependent sensor proteins (Choi et al., 2005). Due to their dependence on calcium, they may regulate kinase cascades and may be involved in regulating signalling during pollen tube growth (Choi et al., 2005). Another transcription

factor belongs to a group of homeobox genes known as knox 2 thought to be involved in development and which have been shown to be expressed in rice flowers (Ito et al., 2002). Finally a zinc finger, MYB 3, and four basic helix-loop-helix (bHLH) transcription factors were also found in the anther-only group. They are involved in the specification of various tissue types and studies in rice have identified these transcription factors among others as anther-specific or expressed during pollen development (Fujita et al., 2010; Hobo et al., 2008; Wei et al., 2010).

Many bHLH transcription factors have been characterized including *Undeveloped Tapetum 1 (Udt1)* which is required for the differentiation and development of tapetal cells (Jung et al., 2005). Similar to our results with the triticales anther-only bHLH transcription factors (data not shown), the level of the *Udt1* transcript was more abundant at the early stages of anther development and progressively decreased up until the heading stage. In the *Udt1* mutant, the meiocytes undergo normal meiotic cell division but soon degenerate due to extremely defective tapetal cells. Furthermore, the anther middle layer failed to degenerate. Another characterized bHLH domain containing protein, *Tapetum Degeneration Retardation (TDR)*, is involved in regulating tapetal program cell death in rice (Li et al., 2006). At the vacuolated pollen stage, the tapetal cells and middle layer failed to degenerate but instead expand causing the microspore to collapse. Furthermore exine deposition was not detected. Further characterization of the triticales anther-only bHLH transcription factors will be required to ascertain if they play important regulatory roles in tapetal and pollen development.

4.2.7 Anther-only protein phosphorylation

A wide array of protein kinases were found within the anther-only dataset. They included two receptor-like protein kinases. Receptor-like protein kinases have been shown to play critical roles in anther development. For example, a rice LRR receptor-like protein kinase has been reported to participate in a signaling pathway that regulates cell proliferation and proper anther wall formation and tapetum development (Nonomura et al., 2003). Interestingly, one of the anther-only triticales receptor-like protein kinase possesses an S-locus glycoprotein domain although it is unlikely to play a role in self-incompatibility. A homolog of the triticales tak33 receptor-like protein kinase has been described in wheat, however its exact role remains unknown (Feuillet et al., 2001). Three serine/threonine and one serine-threonine/tyrosine protein kinases were also found in the anther-only dataset. One corresponded to a Nek5 protein kinases and NIMA-related kinases (Neks) are believed to be involved in plant development processes (Vigneault et al., 2007). Another protein kinase belonged to the U-box domain containing proteins and might play a role in signal transduction. The latter two triticales protein kinases were expressed early in anther development (results not shown). We also found a calcium dependent-like kinase specifically expressed at the TCP stage, and they are known to be involved in signaling pathways related to pollen tube growth (Taylor et al., 1997). Finally, a cdc2 protein kinase was found and these represent conserved regulators of the cell cycle. For example, a mutation in the Arabidopsis CDKA;1 cdc2 protein kinase led to a failure of generative cell mitosis and death of the male gametophyte (Iwakawa et al., 2006). The role of the anther-only proteins involved in protein de-phosphorylation is unknown at this point.

4.2.8 Anther + pollen expression

A total of 348 probe sets expressed specifically in triticales male reproductive tissues were found in both the developing anther and the mature pollen grain. This group should be enriched for genes involved in late pollen development. The major differences between this group and the anther-only group were a substantial increase in the carbohydrate metabolic process and decreases in the transport, as well as lipid and protein metabolism. These trends are in general agreement with a shift from early microspore development and cell wall synthesis involving considerable transport from the tapetum to mature pollen starch and transcript accumulation in support of pollen germination and tube growth.

4.2.9 Anther + pollen cell wall modification and carbohydrate metabolism

The largest category in the anther + pollen dataset was carbohydrate metabolic process and it consisted of enzymes and proteins associated with the accumulation of constituents in the mature pollen necessary to support pollen germination and tube growth, namely starch synthesis (Clément et al., 1994) and cell wall modification/expansion (Krichevsky et al., 2007). The category included genes encoding several enzymes that fall under the general classification of glycoside hydrolases, as well as glycosyltransferases, a sucrose phosphate synthase, as well as carbohydrate and sucrose transporters. Furthermore, it contained numerous cell wall modifying proteins that are often found in maturing pollen (Fujita et al., 2010; Haerizadeh et al., 2009) such as polygalacturonases, pectate lyases, pectin esterases, and pectin esterase inhibitors.

It has been hypothesized that these proteins, along with α -expansins, can also mediate papillar wall expansion and loosening in order to enable pollen tube growth (Shopfer, 2001).

4.2.10 Anther + pollen metabolism

The metabolic process group was the second largest category and it consisted of a variety of proteins and enzymes that differed from those of the anther-only data set with the exception of peroxidases. In fact, the anther + pollen dataset lacked many of the anther-only genes involved in exine synthesis, as well as lipid and protein metabolism. Interestingly, we found that the anther + pollen dataset included several genes encoding proteins involved in membrane trafficking. Three clathrin assembly proteins were identified and they have a major role in forming vesicles. The synaptobrevin-like, Exo70, and dynamin GTPase proteins identified are possibly directly involved in pinching off the clathrin coated vesicles from the parent membrane in exocytosis (Thoms & Erdmann, 2005). In Arabidopsis, 6 *EXO70* genes showed specific expression in the microspore and pollen, and a role in vesicle-trafficking during tip growth was implied (Li et al., 2010b). It is therefore likely that these triticales proteins are accumulating in pollen in readiness for the high requirements for vesicular trafficking associated with pollen tube growth, although additional experimentation will be required to confirm this hypothesis.

4.2.11 Anther + pollen regulatory factors

Unique regulatory factors of the anther + pollen dataset included an AP2 and two MYB transcription factors. Both these types of transcription factors have been identified in gene expression studies of soybean pollen (Haerizadeh et al., 2009), developing rice

anthers (Fujita et al., 2010) and pollen (Hobo et al., 2008; Wei et al., 2010). Plant MYB proteins belong to a superfamily of transcription factors known to regulate various processes including floral development (Li et al., 1999). The AP2 transcription factor is involved in a variety of different plant processes, including floral organ development (Jofuku et al., 1994). The only triticales mature pollen-only transcription factor found was also an AP2-related factor (see below).

Four serine/threonine protein kinases and 4 calcium-dependent protein kinases were also part of the anther + pollen dataset and they were all expressed at significantly high levels in TCP and MPG (data not shown). It is difficult to ascertain the role of the serine/threonine protein kinases at this point and additional experimentation will be required. However, calcium-dependent protein kinases are known to be critical to pollen tube growth (Cheung & Wu, 2008) and since those found in triticales express in MPG, it is likely that they serve a similar role. The anther + pollen dataset also possessed 2 kinase partner proteins that were expressed late in triticales pollen development with maximum expression in the mature pollen grain (data not shown), a tomato kinase partner protein has been shown to interact with pollen-specific receptor kinases and furthermore, pollen tubes over-expressing this protein had defective pollen tube tips (Kaothien et al., 2005).

4.2.12 Mature pollen-only gene expression

Only 51 probe sets were found to be expressed in the triticales mature pollen grain and not in any other tissue including developing anthers. Approximately two thirds were either unknown or unclassified. Since the developing anther contains the developing pollen grain, a pollen-specific gene will also be present in the anther, therefore the 51

probe sets only represent a small fraction of the genes that are expressed solely in the pollen and they correspond to genes whose transcripts appear in the pollen grain following dehiscence. Since the pollen is believed to be transcriptionally silent following dehydration, the occurrence of such probe sets is somewhat surprising.

4.2.13 Mature pollen-only metabolism

An eclectic collection of enzymes and proteins were found in the mature pollen-only metabolic process category. They included a cytochrome P450 protein, an aldehyde dehydrogenase, a formate dehydrogenase, an S-adenosylmethionine decarboxylase, an aspartic protease, as well as two glycosyl transferases. As many of these proteins are involved in a broad range of processes, the exact role of each of these proteins in the mature pollen grain is difficult to surmise and elucidation will require additional experiments. Nonetheless, S-adenosylmethionine decarboxylase which is involved in polyamine synthesis has been shown to be important to pollen germination especially under heat stress (Song et al., 2002). Aldehyde dehydrogenases have been found to express highly in the male gametophyte and their activity has been demonstrated to be important for pollen tube growth (Op den Camp & Kuhlemeier, 1997).

4.2.14 Mature pollen-only defense and stress response

As mature pollen grains are released from the anther they must rapidly achieve pollination before they succumb to environmental stress and eventually die. Defense and response to stimulus transcripts were highly represented groups in the rice mature pollen grain (Wei et al., 2010). Since pollination generally occurs in summer, one of the major

abiotic stresses is excess heat. In addition to the S-adenosylmethionine decarboxylase mentioned above which may have a possible role in conferring heat tolerance, a DnaJ protein was also found in the mature pollen-only dataset. There are 33 proteins with homology to DnaJ proteins expressed highly in the Arabidopsis pollen, in fact the *THERMOSENSITIVE MALE STERILE 1* mutant resulted from a mutation in a protein containing a DnaJ motif (Yang et al., 2009). In view of the climate change predictions, it may be interesting to evaluate the role of the triticales S-adenosylmethionine decarboxylase and DnaJ proteins in conferring heat tolerance to pollen.

The released pollen grain not only faces abiotic stresses, but might also have to overcome biotic stresses. Two defense-related genes were found in the mature pollen-only dataset, a pathogenesis-related (PR) 1 protein and a resistance gene analog (rga) 2 resistance protein. A role for such proteins in specifically protecting the pollen from pathogens still needs to be demonstrated.

4.3 Conclusion

Transcriptome analysis throughout anther and pollen development provides a snapshot of the molecular mechanisms involved in male reproductive development. The developing anther and pollen transcriptomes were found to reflect closely the biological events underlying anther and pollen development. The triticales male tissue transcriptome dataset will serve as an important resource to the research community. Although nearly half of the anther/pollen-specific genes were unknown or unclassified, numerous novel Triticeae genes were revealed including key regulatory genes, thus setting the stage for much future work.

5. Transcriptome Analysis of Triticale Ovary Development

5.1 Results

5.1.1 Microscopic analysis of ovary development

A microscopic analysis of ovary development was performed at developmental stages coinciding with the previously described anther TET, UNM, BCP, and TCP stages (Figure 5-1A, B). Phenotypic characterization of the ovary throughout development was performed using SEM (Figure 5-1C). The external morphological changes are subtle yet visible. At the earliest stage, the ovary is semi-transparent and is 250 μm in width across the mid-section (Figure 5-1B, TET) with epidermal cells that are easily distinguished (Figure 5-1C, TET). Major changes occur beginning at the UNM stage including the initiation of hair growth at the crown which persists until maturity, the enlargement of the ovary, and the formation of longitudinal indentations on both sides of the ovary that progressively alter its shape from round to heart shaped (Figure 5-1B and C, UNM-TCP).

5.1.2 Gene expression profiling throughout ovary development

The flow chart in Figure 5-2 summarizes the process of microarray data analysis for each subsequent section below. A total number of 26,962 probe sets were expressed in at least one stage of ovary development covering 44.1% of the array. As shown in Figure 5-3, the first two stages of development (TET, UNM) displayed similar levels of expressed probe sets and the same observation can be made for the last two stages (BCP, TCP). A sharp increase in the number of expressed probe sets was also evident following the UNM stage.

Figure 5-1. Microscopy of triticale reproductive tissues at four stages of pollen development. (A) Aniline blue and DAPI stained microspore/pollen; (B) light microscopy of developing floret; and (C) SEM of ovary during development. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen. Scale bar = 10µm (top, panel A), 1mm (middle, panel B).

Figure 5-2. Schematic diagram of microarray data analysis for ovary development. A) Workflow for finding differentially expressed probe sets (DEPs) and (B) tissue-enriched and -specific probe sets. Analysis performed for the 3P + 2P (at least 2 Present calls) and 3P datasets, only intersecting probe sets are relevant as they represent 3P data, but also consider 2P values. Number in parentheses indicate the number of different sample arrays used. OV, ovary at 4 different stages of development; RMA, robust multichip average; ANOVA, analysis of variance; DEPs, differentially expressed probe sets.

Figure 5-3. Number of probe sets expressed (present calls in all replicates, 3P) at each stage of ovary development. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

To further expand on the relationship among the four ovary transcriptomes, a dashboard model was constructed to examine the amount of co-expression occurring across different developmental stages. The dashboard diagram consists of the 26,962 probe sets expressed in the ovary in at least one stage of development (Figure 5-4). There are 16,589 probe sets whose expression occurs in all four stages, representing 61.5% of the total number of ovary expressed probe sets. This was by far the largest section of overlapping probe sets indicating that nearly two thirds of the genes expressed in the ovary are found throughout its development. The next largest sections of overlapping probe sets include: UNM/BCP/TCP (1,990; 7.4%), BCP/TCP (1,500; 5.6%), and TET/BCP/TCP (874; 3.2%), which interestingly, always consisted of the BCP and TCP stages. In contrast, groups that included the UNM stage show lower numbers of overlapping genes including the three smallest groups, all of which were below 1%: TET/UNM (39; 0.1%), TET/UNM/TCP (55; 0.2%), and UNM/TCP (133; 0.5%). A significant proportion of the ovary probe sets (15%) were expressed during only one stage of development (expression in other tissues is not being considered): TET (398; 1.5%), UNM (202; 0.8%), BCP (1,547; 5.7%), and TCP (1,987; 7.4%). Most of these probe sets occur later in development (BCP and TCP) which may be a reflection of the greater number of probe sets expressed at these stages (Figure 5-3).

The level of correlation between the co-expressed transcriptomes of all four stages of ovary development was described using scatter plots (Figure 5-5). Remarkably, all the pairwise comparisons displayed relatively high correlation scores. As mentioned above, similar numbers of expressed probe sets were observed between the early (TET/UNM) and late (BCP/TCP) transcriptomes (Figure 5-3), however, the most highly

Figure 5-4. Dartboard model showing the relationship between the 26,962 genes expressed throughout the four stages of ovary development. Numbers shown represent the number of expressed genes in individual or overlapping samples. Total number of expressed probe sets at each stage: Te: 19,081, U: 19,875, B: 23,788, Tr: 23,488. Te, tetrad; U, uninucleate microspore; B, bi-cellular pollen; Tr, tri-cellular pollen.

correlated transcriptomes were UNM and BCP ($R^2 = 0.9495$, Figure 5-5D). Generally, the further apart in the developmental sequence, the lower the correlation of the transcriptome pair, with the lowest correlation obtained between the earliest and latest stages, TET and TCP ($R^2 = 0.7067$, Figure 5-5C).

5.1.3 Co-regulated gene expression in the developing ovary

To identify various expression patterns that exist throughout ovary development, an unsupervised *K*-means clustering was performed. All 26,962 ovary-expressed genes were clustered using the Acuity 4.0 software (Figure 5-6). Based on gap statistics, the optimal number of gene clusters was $k = 36$. Clusters are represented as a color map with the overall expression trend displayed as a line graph. Probe sets preferentially expressed in one stage of development could be observed for TET (cluster 24), UNM (cluster 21), BCP (clusters 15, 31), and TCP (clusters 20, 36). Clusters showing preferential expression in only two stages include TET/BCP (clusters 3, 7), TET/TCP (clusters 4, 8), UNM/BCP (clusters 13, 29), UNM/TCP (cluster 12), and BCP/TCP (clusters 11, 19, 35), but not TET/UNM which was also the smallest group of the dartboard analysis (Figure 5-5). Only one cluster showed expression limited to the first three stages of development TET/UNM/BCP (cluster 23), whereas several clusters displayed expression in the last three stages UNM/BCP/TCP (clusters 9, 17, 25, 33). Furthermore, one quarter of the clusters showed expression in all stages of ovary development, they consistently represented the largest groups of probe sets, and several clusters maintained a relatively high expression throughout all four stages (eg. clusters 2, 6, 18, 34).

Figure 5-5. Scatter plots comparing gene co-expression between different stages of ovary development with the linear correlation (R^2) indicated. Pairwise transcriptome comparison between (A) TET versus UNM, (B) TET versus BCP, (C) TET versus TCP, (D) UNM versus BCP, (E) UNM versus TCP, (F) BCP versus TCP. The linear regression is represented as a purple line. The middle green line represents probe sets expressed at the same level in both datasets whereas the two outer lines represent probe sets with at least a two-fold change in intensity value in one of the datasets. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

Figure 5-6. Graph and color map for an unsupervised *K*-means clustering result of 26,962 ovary expressed genes using the Acuity 4.0 software. Number displayed above each cluster represents the number of probe sets for each group. Stage of development indicated at the top and the line graph shows overall trend for each cluster. Similarity metric: Euclidean Squared, $k = 36$ (number of clusters). TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

5.1.4 Identification of differentially expressed genes in the developing ovary

To find ovary genes that are differentially expressed among the four stages of development, an analysis of variance (ANOVA) statistical test was performed across all 26,962 ovary-expressed genes. A total of 22,040 probe sets met an adjusted P value < 0.05 and therefore showed statistically significant variation in expression. Of these, probe sets that also demonstrate an absolute fold-change of at least 2 (2-FC) were considered differentially expressed probe sets (DEPs). The number of up-regulated DEPs is shown in Figure 5-7, with the bottom table further filtered to include only DEPs having a minimum expression value ≥ 7 . The number of DEPs decreased significantly when filtered to ≥ 7 , although the relative trends were maintained. On average, 56.3% of the up-regulated DEPs have an intensity value ≥ 7 signifying that a little below half of the DEPs within the ovary express at low levels (below 128).

Up-regulation of gene expression between pairs of stages were compared and the further apart the stages, the larger the number of DEPs observed (Figure 5-7, top table). For example, when comparing the TCP stage with the earlier stages, the trend is TET (6,484), UNM (3,672), and BCP (3,057), and this is consistent with the scatter plot data (Figure 5-5) where the correlation between transcriptomes increases as the stages get closer together: TET-UNM ($R^2 = 0.7067$), UNM-TCP ($R^2 = 0.8325$), BCP-TCP ($R^2 = 0.8675$). The values highlighted in red in Figure 5-7 represent probe sets that are up-regulated between successive stages of ovary development. The highest number of up-regulated probe sets was observed during the transition from TET to UNM followed by that from BCP to TCP. Approximately half the number of probe sets was up-regulated in

	TET	UNM	BCP	TCP
TET	—	4,398	5,711	6,484
UNM		—	1,743	3,672
BCP			—	3,057
TCP				—

	TET	UNM	BCP	TCP
TET	—	3,298	3,486	4,129
UNM		—	505	1,871
BCP			—	1,772
TCP				—

Figure 5-7. Differentially expressed probe sets (DEPs) up-regulated throughout ovary development. Top table: probe sets up-regulated at least 2-Fold Change (2-FC) at an adjusted P-value < 0.05; bottom table: probe sets up-regulated at least 2-FC, P-value < 0.05, and expression value ≥ 7 ($\log_2$). Stages labeled on the top have at least a 2-FC increase in expression over the stages indicated on the left hand side. Numbers highlighted in red indicate up-regulated probe sets between consecutive developmental stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

the transition from UNM to BCP and this is in close agreement with the very high correlation between these two transcriptomes (Figure 5-4D).

Down-regulated DEPs were also investigated throughout ovary development (Figure 5-8). Generally, the more distant the developmental stage, the higher the number of repressed genes observed. The values highlighted in green represent down-regulated DEPs at each consecutive stage of ovary development and the trend observed is very similar to that seen with up-regulated probe sets (Figure 5-7). Only a small group of probe sets were down-regulated going from UNM to BCP (442), whereas more than five times as many probe sets decreased in expression between TET and UNM (2,399) or between BCP and TCP (2,541).

5.1.5 Transcripts enriched and specific to the developing ovary

To provide further insight into the transcripts underlying ovary development, we identified probe sets that were preferentially expressed in the ovary (Figure 5-9). An ANOVA statistical test was performed across all 29 tissues to find probe sets that were considered differentially expressed at a confidence level of 95%. Probe sets were considered ovary-enriched if they also displayed a 2-FC up-regulation when compared to the other 25 reproductive, vegetative, and seed tissues, whereas they were considered ovary-specific if they were present in the ovary, but absent in all other datasets.

Of the 26,909 detected probe sets within the ovary, a total of 322 probe sets were found to be enriched in the ovary at either one or more stages of development (Figure 5-9). This Enriched dataset consists of 246 non-redundant probe sets. The fact that these numbers are relatively low is probably due to the close relationship between the ovary

	TET	UNM	BCP	TCP
TET	—	2,399	2,226	3,614
UNM		—	442	2,261
BCP			—	2,541
TCP				—

	TET	UNM	BCP	TCP
TET	—	1,775	1,815	3,045
UNM		—	310	1,725
BCP			—	996
TCP				—

Figure 5-8. Differentially expressed probe sets (DEPs) down-regulated throughout ovary development. Top table: probe sets showing at least a 2-Fold Change (2-FC) down-regulation at an adjusted P-value < 0.05; bottom table: probe sets down-regulated at least 2-FC, P-value < 0.05, and expression value ≥ 7 ($\log_2$). Stages labeled on the top have at least a 2-FC decrease in expression over the stages indicated on the left hand side. Numbers highlighted in green indicate down-regulated probe sets between consecutive developmental stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

and the stigma, and possibly seed tissues being compared. Approximately 63% of the Enriched probe sets occur within the later stages of ovary development, with 27.6% and 35.4% occurring at the BCP and TCP stages, respectively. This dataset was further filtered to find Stage-Enriched probe sets. A total of 144 probe sets displayed stage-enrichment representing 44.7% of the total number of Enriched probe sets, with most being contributed by the latest stage (Figure 5-9).

The Stage-Enriched dataset was filtered further to find probe sets that were specifically expressed in the ovary. There were 112 ovary-specific probe sets that showed no expression in any other tissue, and these were then filtered to identify those that were Stage-Specific (Figure 5-9). Stage-Specific probe sets correspond to probe sets that were not only specific to the ovary, but exclusive to one stage of development. Similar to Stage-Enriched probe sets, the proportion of Stage-Specific probe sets was highest late in development TCP (30.4%) and BCP (22.3%), relative to the earlier stages, TET (10.7%) and UNM (2.7%). Interestingly, UNM had the lowest number of specific probe sets (11).

5.1.6 Characterization of the ovary-specific genes

The distribution of the 112 ovary-specific probe sets based on GO biological process is summarized in Figure 5-10. Over half of the ovary-specific probe sets fall into the unknown or unclassified categories. Of the genes that could be identified, the larger categories included macromolecule biosynthetic process, protein (de)phosphorylation, and protein metabolic process.

Figure 5-9. Ovary-expressed, -enriched, and -specific probe sets throughout development. Enriched probe sets have at least a 2-fold higher expression over all other triticales tissues (leaf tissue at seedling, tillering, early boot, early and late senescence; coleoptile; root; stem tissue throughout development based on Zadoks' scale Z21, Z45, Z59, Z75; embryo; epidermis; endosperm; pericarp; crease; anther and stigma at TET, UNM, BCP and TCP stages, mature pollen grain). Stage-Enriched probe sets are probe sets that are also enriched in a particular stage compared to other ovary stages. Specific probe sets correspond to Enriched probe sets with no expression in non-ovary tissues. Stage-specific probe sets are Specific probe sets that show no expression in other ovary stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

Probe sets related to the macromolecule biosynthetic process group included genes that are involved in RNA/DNA replication and translation. For example, such genes included two DNA and one RNA-directed polymerase, a translation initiation factor and constituents of the ribosome, all of which were specifically expressed in the ovary. Interestingly, all of the genes in the protein metabolism category were made up of different types of proteolytic enzymes such as aspartic, cysteine, or subtilisin-like proteases. The protein (de)phosphorylation group consisted of three protein kinases, a receptor-like kinase, a leucine-rich repeat kinase, and two phosphatases. A single ovary-specific transcription factor was identified and it belonged to the B3 domain transcription factor family. It is noteworthy that ovary-specific carbohydrate and lipid metabolic processes were very poorly represented.

The top 50 highest expressing ovary-specific probe sets are summarized in Table 5-1 and a little more than half have unknown functions or are unclassified. Of the top 20 highest expressors, 18 were most abundant during BCP or TCP. The genes encoding the four proteases are also found within the top 20 ovary-specific probe sets. Interestingly, the second most abundant ovary-specific gene encodes the B3 transcription factor.

Figure 5-10. Distribution of biological process categories for ovary-specific probe sets.

Probe Set ID	MAX	Description	Probe Set ID	MAX	Description
1 Ta 4898 1 A1_s_at	TCP	aspartic protease	26 Ta 10219 2 S1_at	TET	unknown
2 TaAffx 65336 1 A1_at	TCP	B3 transcription factor	27 Ta 20863 1 S1_x_at	TET	actin
3 Ta 12792 1 S1_x_at	TCP	unknown	28 Ta 26231 1 A1_at	BCP	phosphoesterase
4 TaAffx 80103 1 S1_at	TCP	leucine-rich repeat protein	29 TaAffx 92373 1 S1_at	TCP	unknown
5 TaAffx 79990 1 S1_at	TCP	subtilisin-like protease	30 TaAffx 87067 1 S1_at	UNM	unknown
7 Ta 4557 1 S1_at	BCP	unknown	32 Ta 5998 1 S1_x_at	TET	unknown
8 TaAffx 30463 1 S1_at	BCP	stress-related protein	33 TaAffx 97850 1 S1_at	UNM	unknown
9 Ta 9952 1 S1_at	TCP	unknown	34 TaAffx 38788 1 S1_at	TCP	unclassified
10 Ta 25262 1 S1_at	BCP	ferredoxin reductase	35 Ta 11882 1 S1_at	UNM	oxoglutarate-Fe(II) oxygenase
11 TaAffx 91428 1 S1_at	TET	unknown	36 TaAffx 77835 1 S1_at	TCP	Ribosome constituent
12 Ta 13844 1 S1_at	TCP	vacuolar protein sorting	37 Ta 20413 3 S1_at	TCP	receptor protein kinase
13 Ta 6512 2 S1_a_at	TCP	DUF1668 protein/unclassified	38 Ta 9365 2 S1_at	TET	DUF827 protein/unclassified
14 Ta 27768 1 S1_at	BCP	unknown	39 TaAffx 85625 1 S1_at	TCP	unknown
15 Ta 4715 1 S1_at	TCP	unknown	40 Ta 13835 1 S1_at	TCP	unknown
16 Ta 27995 1 S1_x_at	TCP	unknown	41 Ta 12662 3 S1_at	UNM	DUF2828 protein/unclassified
17 TaAffx 119549 1 S1_at	TCP	aspartic protease	42 TaAffx 11972 1 A1_at	BCP	unknown
18 TaAffx 59731 1 S1_at	TCP	unknown	43 TaAffx 31845 1 S1_at	BCP	npr1-like protein
19 Ta 3691 1 S1_at	BCP	cysteine-rich protein	44 Ta 12774 1 A1_at	UNM	unknown
20 TaAffx 70061 4 S1_at	UNM	DUF639 protein/unclassified	45 TaAffx 52640 1 S1_at	TCP	BURP protein/unclassified
21 TaAffx 87029 1 S1_at	TCP	zinc finger protein/unclassified	46 TaAffx 79006 1 S1_at	TCP	set domain protein/unclassified
22 TaAffx 75697 1 A1_s_at	TET	auxin influx carrier component	47 Ta 6359 2 S1_x_at	TET	DNA polymerase epsilon
23 TaAffx 48377 1 S1_at	BCP	unknown	48 Ta 26027 1 A1_at	TCP	unknown
24 Ta 23301 2 A1_at	TET	serine/threonine kinase	49 TaAffx 12223 1 S1_at	TET	chromatin assembly factor-1
25 Ta 18416 1 S1_at	BCP	unknown	50 Ta 28127 3 S1_at	TCP	cop9 protein

Table 5-1. The 50 most highly expressed ovary-specific probe sets. Stage with highest expression for each probe set is shown under 'MAX'. Top 50 expression values range from 9.23-5.38 based on log₂ scale. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

5.2 Discussion

The research presented here provides significant insight into the molecular biology of the developing ovary in a grass species. In fact, this is the first study to generate transcriptome profiles for triticales ovaries at four stages of development using the Affymetrix Wheat GeneChip[®]. Previous gene expression analyses have been predominantly generated only for mature ovaries in *Arabidopsis* (Swanson et al., 2005; Tung et al., 2005) and in rice (Li et al., 2007; Fujita et al., 2010). The study by Fujita et al. (2010) surveyed a wide range of transcriptome profiles in rice reproductive tissues including developing anthers, as well as mature ovary and stigma before and after pollination. The remaining studies profiled the mature ovary mainly for comparison purposes within studies of the stigma. Within the Triticeae, the most relevant gene expression analyses were performed on mature pistils of barley (Druka et al., 2006) and wheat (Schreiber et al., 2009) which should also contain mature ovary transcripts.

5.2.1 Developmental changes in the ovary transcriptome

We have based our staging for the female tissues on established phases of pollen development: TET, UNM, BCP, TCP. This will help compare the gene expression profiles of different organs at similar developmental time points. The duration of female and male meiosis has been shown to be quite synchronous in wheat and barley, but none so much in rye (Bennett et al., 1973). The timing of the mitotic phases of the embryo sac and pollen were not synchronized and divisions leading to the differentiation of the seven celled female gametophyte progressed much more rapidly than the male gametophyte (Timar et al., 1997; Bennett et al., 1973). The megaspore in the ovule goes through two

mitotic divisions producing an 8 nucleate embryo sac concurrent with PM I (Bennett et al., 1973). The embryo sac becomes fully mature once the pollen grains reach the second pollen mitosis and enter the TCP stage. Nonetheless, the male and female gametophytes are regulated to mature concurrently to facilitate self-pollination (Xi & Cui, 1983).

The two earliest stages of tritcale ovary development showed a very similar number of expressed probe sets and the same was observed with the last two stages, although they showed a considerable increase in expressed probe sets. This indicates that the ovary probably enters a more complex developmental program beginning at the BCP stage, in fact the female gametophyte is octonucleate beginning at BCP and this through to maturity. Probe sets occurring in all stages make up approximately 62% of the total number of overlapping ovary-expressed probe sets. This is a considerably larger proportion than was observed in the anther and probably points to a lesser level of differentiation within ovary tissues throughout development. The next four largest overlapping sections include BCP and/or TCP, and interestingly, the numbers of probe sets found only in BCP or TCP are greatly superior to those found only in TET or UNM. These results are somewhat consistent with the higher number of probe sets found at these later stages and again point to increased levels of specialization in the maturing ovary probably in preparation for pollination and fertilization events.

Interestingly, in contrast to the wide range in the levels of correlation among the anther transcriptomes, comparisons between the different co-expressing developing ovary transcriptomes revealed not only a relatively high degree of correlation, but also a much reduced level of variation throughout development with correlation values ranging from 0.7067 (TET/TCP) to 0.9495 (UNM/BCP). Based on the number of expressed probe sets

and the number of overlapping probe sets mentioned above, one might have expected the BCP and TCP transcriptomes to be the most highly correlated and although they did show the second highest correlation value, like in the anther, the UNM and BCP transcriptomes were the most highly correlated. This indicates that despite the differences in numbers, the UNM and BCP transcriptomes likely underscore largely similar functions.

Expression profiling of ovary development revealed a large proportion of genes expressed in all stages, in agreement with the overlapping expression data. What is remarkable is the fact that many clusters revealed gene expression exclusively in one or two stages, demonstrating distinctive stage-specific expression even though the developmental staging was based on pollen development. There are many critical events associated with female gametophyte formation and ovule patterning, and the cluster analysis would indicate that these processes require gene expression within restricted developmental windows.

Most of the ovary-expressed probe sets (82%) displayed differential expression throughout development. Given that 62% of the probe sets are shared among all the stages, this would point to varying levels of transcription regulation of the same genes throughout development. As with the anther, the further apart the stages, the higher the number of up- and down-regulated probe sets generally. Accordingly, the highest numbers of DEPs in both cases resulted from comparisons made between the TET and TCP stages. This suggests that cellular activities during early development most likely are different or vary in emphasis later in the mature ovary. The transition from TET to UNM and from BCP to TCP involved the largest changes observed between consecutive

developmental stages. In agreement with the correlation data, the smallest changes in differentially expressed genes were found in the transition from the UNM to the BCP stage, especially when comparing highly expressed probe sets. These results would seem to suggest that ovary development is divided into three phases consisting of an early (TET to UNM) and a late (BCP to TCP) phase in which gene activity is diverse, separated by a middle phase (UNM-BCP) of more conserved activity which in turn may reflect more stable biological processes. These phases may in turn reflect activities associated with the initiation of ovary development, the preparations for pollination and fertilization, and a period of growth and expansion, respectively

A total of 112 ovary-specific probe sets were identified and the majority was expressed at low levels. Compared to the male reproductive tissues, significantly less female-specific probe sets were found. In fact, most of the expressed probe sets in the female tissues were common between the ovary and stigma (Chapter 7.1) and comparing their transcriptomes at the TCP stage revealed a high correlation ($R^2 = 0.8024$, data not shown). This would mean that pistil-specific transcripts found in both the ovary and stigma would not appear in either ovary-specific or stigma-specific groups. More than half of the ovary-specific probe sets were classified as unknown or unclassified. The probe sets that could be classified are discussed below.

5.2.2 Ovary-specific phosphoregulation

Five protein kinases and two protein phosphatases were ovary-specific. The triticale ovary-specific checkpoint protein kinase (chk1) expressed early in development (data not shown). The chk1 protein kinases have an important role in cell cycle arrest

and apoptosis in the event of DNA damage such as UV (Tang et al., 2006). Since they are required for mitotic progression, it may explain why our gene was found to be TET-specific. A calcium-dependent kinase was also found to be ovary-specific. Calcium-dependent protein kinases are important in plant development as they regulate downstream calcium mediated signaling pathways by sensing the flux of Ca^{+2} ions (Ray et al., 2007). Calcium is known to be important to pollen tube guidance within the ovary (Malhó, 1998) and high concentrations of Ca^{+2} have been observed in the synergid cells of the ovule (Chaubal & Reger, 1992). Two ovary-specific leucine-rich-repeat (LRR) receptor-like kinases were identified and one was similar to Clavata 1. Clavata 1 functions in cell proliferation and differentiation of flower meristems (Stone et al., 1998) and the *clavata 1* mutation is known to affect pistil development in Arabidopsis (Okada et al., 1989). The remaining kinase is a serine/threonine protein kinase whose role is unknown. Similarly, little is known of the role of the two ovary-specific phosphatases, one is a phosphatase 2C and the other is a calcineurin-like phosphoesterase, both may be involved in signaling. Given that these triticale ovary-specific kinases and phosphatases are being characterized for the first time, much work remains to confirm their role in ovary development.

5.2.3 Ovary-specific proteolysis

Most of the genes in the protein metabolic process category were involved in protein degradation and included aspartic, cysteine, and subtilisin-like proteases. They expressed late in ovary development (data not shown). This finding was also reported for the mature ovary transcriptome of Arabidopsis (Swanson et al., 2005). Barley and wheat

proteases have been identified in the nucellus tissue layer of the ovary (Bi et al., 2005; Linnestad et al., 1998; Dominguez & Cejudo, 1998). The nucellus makes up a large component of the ovule and is very important in female gametophyte development. Following fertilization, it undergoes programmed cell death, a process that requires proteases. Accordingly, many proteases tend to express exclusively in the nucellus during late development and after fertilization (Bi et al., 2005; Chen & Foolad, 1997). Proteases expressed late in development may also be required for the degeneration of pistil tissues in order to facilitate pollen tube penetration (Wu & Cheung, 2000). Others have proposed that proteases have a role in processing cell wall proteins of the developing embryo or cereal grain (Linnestad et al., 1998). All these putative roles are associated with pollination, fertilization, and post-fertilization events, and as the pollen stores RNA for rapid translation of proteins involved in germination, the ovary may be storing mRNAs encoding proteases for their rapid production during these processes.

5.2.4 Ovary-specific transcription factor

Transcription factors were not abundant in our ovary-specific dataset, even though gene expression studies of particular female gametophyte cell types such as synergid, central and egg cells have shown an over-representation of transcription factors (Wuest et al., 2010). This would seem to indicate that microarray analysis of the entire ovary is not sufficiently sensitive to detect the expression of transcription factors expressed in a limited number of specialized cells, especially if they are expressing at low levels. A single ovary-specific B3 domain containing transcription factor was identified and it was the second most abundant ovary-specific transcript. Male tissue-specific B3

type transcription factors were identified, and this may indicate that they serve a fundamental role in reproductive development.

5.2.5 Ovary-specific genes

The remaining ovary-specific probe sets represented a very broad array of metabolic and biosynthetic processes. There was very little obvious connection among the probe set-encoded functions or with ovary development in particular. Some interesting genes included those involved in the defense response (rgh1a, γ -thionin, lr10, npr1-like), genes involved in chromosome organization (chromatin assembly factor 1, sister chromatid cohesion protein DCC1), genes involved in nucleic acid metabolism (DNA polymerase ϵ , RNA polymerase Rpb4, a reverse transcriptase, a DNA/RNA helicase), genes involved in translation (ribosomal constituents, translation initiation factor 3g), genes for transporters (auxin, ammonium, phosphate), and some individual genes of interest included cop9, M cyclin, and an endosperm transfer cell-specific pr9 precursor. Since a number of these genes relate to nucleic acid biosynthesis, chromosome organization, and cell division, it is tempting to speculate that they may be involved in the processes associated with the cell divisions in embryo sac development. Additional studies will be necessary to validate such speculations.

5.3 Conclusion

Here we present the first gene expression study on the developing Triticeae ovary. Although we based our staging on pollen development, significant trends were observed in the transcriptomes of the developing ovary and many stage-specific genes were identified. The ovary-specific genes were associated with a variety of biological processes, with specialized functions in macromolecule biosynthesis, protein phosphorylation, and protein metabolism. Future functional analyses of these genes will provide further insight in ovary development and their roles before, during, and after fertilization.

6. Transcriptome Analysis of Triticale Stigma Development

6.1 Results

6.1.1 Microscopic analysis of stigma development

As in many other members of the Poaceae family, the triticale stigma is a double branched plumose structure and is considered a “dry” type stigma. Developing stigmas were examined using light and scanning electron microscopy at stages corresponding with the pollen TET, UNM, BCP and TCP stages as described previously (Figure 4-1, Chapter 4). Throughout development, the tip of the stigma was observed to develop sooner than the proximal part. Furthermore, the older the stigmas the easier it was to obtain a high quality image as young stigmas tended to collapse, and this likely coincides with the gradual deposition of cuticular wax as the stigma matures.

Early in development, nodular-like protrusions begin appearing on both primary stigmatic branches (Figure 6-1B, C, TET). At the UNM stage, these protrusions have elongated into multicellular and multiseriate stigmatic secondary branches (Figure 6-1C, UNM). Later in development, the tips of individual secondary branches begin to curl outwards (Figure 6-1C, BCP) becoming fully extended and covered with papillae at maturity (Figure 6-1C, TCP). The secondary branches only proliferate on the ventral side of each main stigma branch, and eventually the two primary branches completely grow apart between BCP and TCP (Figure 6-1B, BCP, TCP). At this point, the stigma is fully extended and covers a large surface area ready for pollen capture (Figure 6-1C, TCP).

Figure 6-1. Microscopy of triticale reproductive tissues at four stages of pollen development. (A) Aniline blue and DAPI stained microspore/pollen; (B) light microscopy of developing stamen and pistil; and (C) SEM of the tip of the stigma during development, section shown in SEM close-up indicated by arrow in panel B. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen. Scale bar =10 μ m (panel A), 1 mm (panel B).

6.1.2 Gene expression profiling throughout stigma development

A schematic diagram outlining the method of analysis for each result section below is illustrated in Figure 6-2. A total of 26,147 probe sets were detected in the stigma in at least one stage of development, covering 42.8% of the array. Figure 6-3 summarizes the number of probe sets expressed at each stage. The lowest number of expressed probe sets occurs at the earliest stage (TET) followed by the latest stage (TCP), while BCP showed the highest number of probe sets.

A further analysis of the relationships among different stigma stages was performed to identify the degree of overlap in gene expression using a dartboard model (Figure 6-4). As observed for the anther and ovary, the highest proportion of overlap consisted of probe sets expressing in all four stages of stigma development (15,147, 58%). The next largest datasets included those showing overlap between the last three stages UNM/BCP/TCP (2,053), the first three stages TET/UNM/BCP (1,691), or between the UNM and BCP (1,123) stages. Accordingly, co-expressing datasets that did not include both UNM and BCP displayed the lowest values of overlap: TET/UNM/TCP (98), UNM/TCP (149), TET/BCP (178), or TET/TCP (196). Datasets with exclusive expression in one stigma stage (expression in other triticales tissues is not being considered in this analysis) showed an increase in the number of probe sets as stigma development progressed, with the TCP stage showing one of the largest datasets overall (1,635).

Figure 6-2. Schematic diagram of microarray data analysis for stigma development. A) Workflow for finding differentially expressed probe sets (DEPs) and (B) tissue-enriched and -specific probe sets. Analysis performed for the 3P + 2P (at least 2 Present calls) and 3P datasets, only intersecting probe sets are relevant as they represent 3P data, but also consider 2P values. Number in parentheses indicate the number of different sample arrays used. ST, stigma at 4 different stages of development; RMA, robust multichip average; ANOVA, analysis of variance; DEPs, differentially expressed probe sets.

Figure 6-3. Number of probe sets expressed (present calls in all replicates, 3P) at each stage of stigma development. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

Figure 6-4. Dartboard model showing the relationship between the 26,147 genes expressed throughout the four stages of stigma development. Numbers shown represent the number of expressed genes in individual or overlapping samples. Total number of expressed probe sets at each stage: Te: 18,531, U: 21,396, B: 22,374, Tr: 20,577. Te, tetrad; U, uninucleate microspore; B, bi-cellular pollen; Tr, tri-cellular pollen.

The correlation between transcriptomes was examined using scatter plots of co-expressed genes (Figure 6-5). The variation among transcriptomes increased with the developmental distance between subsequent stages, accordingly, the smallest correlation was found when comparing the earliest (TET) with the latest stage (TCP), revealing a relatively weak score ($R^2 = 0.4686$, Figure 6-5C). The most similar transcriptomes were UNM and BCP demonstrating the highest correlation value ($R^2 = 0.9178$, Figure 6-5D). This result is in concordance with the dartboard model demonstrating the groups with the highest overlapping probe sets always included the UNM and BCP stages (Figure 6-4).

6.1.3 Co-regulated gene expression in the developing stigma

To identify different co-regulated gene expression patterns throughout stigma development, an unsupervised *K*-means clustering analysis was performed on the 26,147 stigma-expressed probe sets. Probe sets were grouped into 30 clusters ($k = 30$) which is the optimal number of clusters generated by gap statistical analysis (Figure 6-6). This analysis revealed that probe sets showing preferential expression at a single time point in development could be found for each stage: TET (cluster 28), UNM (cluster 6), BCP (cluster 25), and TCP (clusters 5, 19). Profiles demonstrating higher expression throughout early development included those that are expressed in the first two (cluster 14) or the first three (clusters 2, 10, 24) stages. The reverse patterns of expression could also be observed and included probe sets showing expression in the last two (clusters 1, 15) or the last three (clusters 7, 21, 29) stages. Other clusters highlighted probe sets with trends that showed preferential expression in only two (clusters 20, 27) or three (clusters 9, 13, 23) stages, but with at least one intermediate stage with no expression. However,

Figure 6-5. Scatter plots comparing gene co-expression between different stages of stigma development with the linear correlation (R^2) indicated. Pairwise transcriptome comparison between (A) TET versus UNM, (B) TET versus BCP, (C) TET versus TCP, (D) UNM versus BCP, (E) UNM versus TCP, (F) BCP versus TCP. The linear regression is represented as a purple line. The middle green line represents probe sets expressed at the same level in both datasets whereas the two outer lines represent probe sets with at least a two-fold change in intensity value in one of the datasets. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

Figure 6-6. Graph and color map for an unsupervised *K*-means clustering result of 26,147 stigma-expressed genes using the Acuity 4.0 software. Number displayed above each cluster represents the number of probe sets for each group. Stage of development indicated at the top and the line graph shows overall trend for each cluster. Similarity metric: Euclidean Squared, $k = 30$ (number of clusters). TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

consistent with the dartboard data (Figure 6-4), most probe sets were found in clusters showing expression at relatively high levels in all stages with subtle pattern differences (clusters 4, 8, 12, 18, 22, 26, 30).

6.1.4 Identification of differentially expressed genes in the developing stigma

An analysis of variance (ANOVA) statistical test was performed on the 26,147 probe sets detected within the stigma. A total of 21,422 probe sets met a cutoff P-value < 0.05 and were therefore regarded as showing a statistically significant variation in gene expression. Furthermore, probe sets that also showed a 2-fold up- or down-regulation above the other stages were considered differentially expressed probe sets (DEPs). The number of up-regulated DEPs for each stage of stigma development is summarized in Figure 6-7, and includes those that meet a cutoff expression value ≥ 7 , with the latter group showing a substantial drop in the number of probe sets. An increase in the number of DEPs was observed throughout the course of development as the stages diverge, a trend predominant throughout the entire experiment. Accordingly, up-regulated DEPs were most abundant when comparing the earliest and latest TET and TCP stages (7,281), whereas the lowest number of DEPs occurred between UNM and BCP (1,686). The above trends were consistent with previous results pointing to the least and most highly correlated transcriptome pairs (Figure 6-5). Upon investigation of the probe sets which are up-regulated in consecutive stages (highlighted in red in Figure 6-7), the highest number of up-regulated probe sets occurred between TET and UNM, whereas the lowest number was found between UNM and BCP, the latter result being consistent with previous analyses indicating the close relationship between these two transcriptomes. An

	TET	UNM	BCP	TCP
TET	—	5,125	6,316	7,281
UNM		—	1,686	4,994
BCP			—	3,420
TCP				—

	TET	UNM	BCP	TCP
TET	—	1,736	2,068	2,354
UNM		—	439	1,725
BCP			—	1,317
TCP				—

Figure 6-7. Differentially expressed probe sets (DEPs) up-regulated throughout stigma development. Top table: probe sets up-regulated by at least a 2-Fold Change (2-FC) at an adjusted P-value < 0.05; bottom table: probe sets up-regulated at least 2-FC, P-value < 0.05, and expression value ≥ 7 ($\log_2$). Stages labeled on the top have at least a 2-FC increase in expression over the stages indicated on the left hand side. Numbers highlighted in red indicate up-regulated probe sets between consecutive developmental stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

increase in the number of up-regulated probe sets was then observed in the BCP-TCP transition.

The total numbers of stigma probe sets displaying a 2-fold decrease in expression throughout development are summarized in Figure 6-8 (top). The DEPs showing a 2-fold decrease in expression with respect to probe sets with original expression values ≥ 7 are also presented (Figure 6-8, bottom). The latter table excludes probe sets whose original low expression values might be questionable. The average proportion of down-regulated DEPs from a signal intensity ≥ 7 equals 73% of the original number of down-regulated DEPs. This is in concordance with Table 3-3 (Chapter 3), which indicated that 73 % of the 3P data have a signal value above 128. Generally, the number of repressed probe sets increased with the developmental distance between stages. Not surprisingly, the lowest number of repressed probe sets was observed during the transition from the UNM to BCP stage, while the highest numbers of down-regulated genes occurred at the TCP stage (Figure 6-8). Interestingly, but not unexpectedly, the trend in the numbers of down-regulated DEPs between consecutive stages (highlighted in green in Figure 6-8) was similar to that observed for the up-regulated DEPs in that the UNM and BCP transition showed the least differential expression, but differed in that the number of down-regulated DEPs are lower in TET to UNM than BCP to TCP, whereas the up-regulated DEPs showed the opposite relationship (Figure 6-7).

	TET	UNM	BCP	TCP
TET	—	2,116	2,968	5,561
UNM		—	1,329	5,721
BCP			—	4,355
TCP				—

	TET	UNM	BCP	TCP
TET	—	1,614	2,460	4,454
UNM		—	865	4,028
BCP			—	2,726
TCP				—

Figure 6-8. Differentially expressed probe sets (DEPs) down-regulated throughout stigma development. Top table: probe sets showing at least a 2-Fold Change (2-FC) down-regulation at an adjusted P-value < 0.05; bottom table: probe sets down-regulated at least 2-FC from an original expression value ≥ 7 ($\log_2$), P-value < 0.05. Stages labeled on the top have at least a 2-FC decrease in expression over the stages indicated on the left hand side. Numbers highlighted in green indicate down-regulated probe sets between consecutive developmental stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

6.1.5 Transcripts enriched and specific to the developing stigma

Stigma-enriched genes were identified by comparing relative expression levels across 25 other triticales reproductive, leaf, and seed tissues (see Materials and Methods). An ANOVA statistical test was performed across the 26,147 stigma expressed probe sets. A total of 26,108 probe sets met a cutoff P-value < 0.05 and therefore showed significant differential gene expression. Stigma probe sets that also show a 2-FC difference in expression above all other tissues were considered enriched. A total of 857 probe sets were found to be enriched in the stigma at one or more stages of development (Figure 6-9), and when redundant probe sets were removed, 624 individual stigma-enriched probe sets were identified. The number of enriched probe sets increased as stigma development progressed: TET (48), UNM (118), BCP (243), and TCP (448). The TCP stage possessed the largest number of enriched probe sets even though it had the second lowest number of total expressed genes among the four stages (Figure 6-3).

The Enriched dataset was further filtered to identify Stage-Enriched probe sets which correspond to probe sets that are enriched at one particular stage when also compared to the other stigma developmental stages. The highest proportions of Stage-Enriched probe set were found in the first and last stages with the large majority found at the TCP stage (Figure 6-9).

We found 117 stigma-specific genes that were expressed in one or more stages of stigma development. Specific probe sets were further sub-divided into Stage-Specific probe sets which are expressed exclusively at one stage of stigma development. Almost half of the stigma-specific probe sets were Stage-Specific, with more probe sets expressed during the TCP (39) stage, than all other stages combined (Figure 6-9).

Figure 6-9. Stigma-expressed, -enriched, and -specific probe sets throughout development. Enriched probe sets have at least a 2-fold higher expression level over all other triticale tissues (leaf tissue at seedling, tillering, early boot, early and late senescence; coleoptile; root; stem tissue throughout development based on Zadoks' scale Z21, Z45, Z59, Z75; embryo; epidermis; endosperm; pericarp; crease; anther and ovary at TET, UNM, BCP and TCP stages, mature pollen grain). Stage-Enriched probe sets are stigma-enriched probe sets that are also enriched in a particular stage compared to other stigma stages. Specific probe sets correspond to Enriched probe sets with no expression in non-stigma tissues. Stage-specific probe sets are Specific probe sets that show no expression in other stigma stages. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

6.1.6 Characterization of the stigma-specific genes

The distribution of the biological processes in which the stigma-specific probe sets were found is shown in Figure 6-10. Combined, the ‘unknown’ and ‘unclassified’ groups make up the largest portion of stigma-specific probe sets. The next largest groups included lipid metabolic process and regulation of transcription, followed by metabolic process. Many enzymes are represented in the lipid metabolic process category including lipases, a hydrolase, a stearyl-ACP desaturase, a cyclopropane fatty acyl-phospholipid synthase, three very-long-chain 3-ketoacyl-CoA synthases, as well as an acyl-transferase. Surprisingly, regulation of transcription was a highly represented group and included stigma-specific transcription factors with the following identified domains: roc1, GATA zinc-finger, MADS, MYB, and YABBY. The metabolic process group represents a broad array of genes including metabolic functions relating to transferase activity and oxidation/reduction. The protein (de)phosphorylation category was comprised of two protein kinases and two receptor-like protein kinases, as well as a protein phosphatase 2C.

The 50 highest expressing stigma-specific probe sets are presented in Table 6-1 and remarkably, almost half of these stigma probe sets were classified as unknown or unclassified. The most abundantly expressed stigma-specific probe sets represented a varied collection of genes involved in numerous processes of which lipid biosynthesis, transcription regulation, and protein phosphorylation and dephosphorylation are well represented.

Figure 6-10. Distribution of biological process categories for stigma-specific probe sets.

Probe Set ID	MAX	Description	Probe Set ID	MAX	Description
1 Ta.28079.1.S1_x_at	TCP	unknown	26 TaAffx.26550.1.S1_at	TCP	roc1-transcription factor
2 Ta.11473.1.S1_at	TCP	unknown	27 TaAffx.4884.1.S1_at	BCP	cyclopropane fatty acid synthase
3 Ta.4972.1.A1_at	TCP	unknown	28 TaAffx.10328.1.S1_at	TCP	unknown
4 TaAffx.52744.1.S1_at	TCP	retrotransposon/ unclassified	29 TaAffx.100076.1.S1_at	TCP	protein phosphatase
5 TaAffx.129312.1.S1_at	TCP	stearoyl-acp desaturase	30 Ta.13460.1.S1_at	TCP	Tamyb10-A -transcription factor
6 Ta.14474.1.S1_x_at	TCP	unknown	31 TaAffx.29157.1.S1_at	UNM	Ribosomal protein
7 Ta.10977.1.S1_at	TCP	BTB/POZ / NPH3 domain	32 TaAffx.119672.1.A1_at	BCP	unknown
8 TaAffx.25228.1.S1_at	BCP	unknown	33 TaAffx.79475.1.S1_x_at	BCP	unknown
9 TaAffx.92417.1.S1_at	TCP	receptor-like protein kinase	34 TaAffx.86998.1.S1_at	TCP	O-acyltransferase, WSD1, N-terminal
10 TaAffx.29977.1.S1_at	TCP	lipase hydrolase family protein	35 TaAffx.48364.1.S1_at	TCP	VLCFA condensing enzyme cut1
11 Ta.29372.1.S1_at	TCP	polygalacturonase-inhibiting protein	36 Ta.16366.1.S1_at	TCP	unknown
12 TaAffx.50917.1.S1_at	TCP	DNA heat shock protein	37 TaAffx.121723.2.S1_s_at	UNM	RNA polymerase subunit
13 TaAffx.113464.1.S1_s_at	TCP	disease resistance protein	38 TaAffx.55466.1.S1_at	TCP	O-fucosyltransferase
14 TaAffx.119097.1.S1_at	TCP	unknown	39 Ta.7227.2.S1_at	TET	yabby protein-transcription factor
15 Ta.6378.1.S1_at	BCP	unknown	40 Ta.4814.2.A1_at	TCP	unknown
16 TaAffx.78780.1.S1_at	TCP	amidase	41 Ta.5778.3.S1_at	TCP	flavonol glucosyltransferase
17 Ta.18809.1.S1_at	TCP	unknown	42 TaAffx.38665.1.S1_at	UNM	unknown
18 TaAffx.38572.1.A1_at	TET	transcription factor	43 Ta.3132.1.S1_x_at	TCP	unknown
19 TaAffx.59709.1.S1_at	TCP	lipase class 3 family protein	44 TaAffx.9527.1.S1_at	TCP	lipase class 3 family protein
20 TaAffx.78610.1.S1_at	TET	glycosyl transferase family 14	45 Ta.8198.1.S1_at	UNM	unknown
21 TaAffx.109043.1.S1_at	TCP	ibr domain / Unclassified	46 Ta.19512.1.S1_at	TCP	tpr domain protein / Unclassified
22 TaAffx.28071.1.S1_at	TCP	protein kinase	47 TaAffx.15773.1.A1_at	TET	unknown
23 TaAffx.111169.2.S1_s_at	TCP	Unclassified	48 TaAffx.106722.1.S1_x_at	TCP	Unclassified
24 TaAffx.100076.1.S1_x_at	TCP	protein phosphatase	49 Ta.5626.1.S1_at	BCP	Unclassified
25 TaAffx.114045.1.S1_at	TCP	auxin response-transcription factor	50 TaAffx.36982.1.S1_at	BCP	lipid transfer protein family

Table 6-1. The 50 most highly expressed stigma-specific probe sets. Stage with highest expression for each probe set is shown under 'MAX'. Top 50 expression values range from 13.76 – 6.20 based on log₂ scale. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen.

6.2 Discussion

We have identified genes expressed in the triticales stigma using the Affymetrix Wheat GeneChip®. To our knowledge, this is the first study profiling stigma gene expression patterns throughout development in any plant. Furthermore, the stigma dataset was compared to that of 25 other triticales tissues at different stages of development to identify stigma-enriched genes. Relevant transcriptomic studies have been performed only on mature stigmas in *Arabidopsis* (Swanson et al., 2005; Tung et al., 2005) and rice (Li et al., 2007; Fujita et al., 2010), as well as on pollinated rice stigmas (Fujita et al., 2010). In the Triticeae, transcript profiling data was generated for whole mature pistils in wheat (Schreiber et al., 2009) and barley (Druka et al., 2006), but was not described in detail.

6.2.1 Developmental changes in the stigma transcriptome

Currently, key landmark stages in cereal stigma development have not been reported, we therefore profiled the developing stigma transcriptome according to stages of pollen development: TET, UNM, BCP, and TCP. The structural features of the developing triticales stigma were characterized using microscopy. To our knowledge, this is the first phenotypic analysis of the developing stigma in cereals. At the onset of stigma development, small protuberations become evident on the two main branches that eventually elongate and become secondary branches covered with papillae. At TCP, the two triticales stigmatic branches are fully extended and well separated, resembling those of other grass species (Heslop-Harrison and Heslop-Harrison, 1980; Ciampolini et al., 2001). The secondary branches were only observed on the ventral side of the stigma

suggesting an evolutionary adaptation as the dorsal portion of the stigma typically does not come into contact with pollen.

Unlike the anther, the number of stigma-expressed probe sets did not differ greatly among the stages, although a steady increase was observed which peaks at the BCP stage. The lowest number of expressed probe sets occurred at the TET stage when stigma development is just beginning. A small decrease in expressed probe sets was also seen during the transition from BCP to TCP, this may parallel a reduction in stigma growth and differentiation at the mature stage.

The TET stage had the least number of overlapping probe sets of any stigma stage which suggests some distinction in its transcriptome as well as some similarity among the later stages as the stigma expands and differentiates. Nevertheless, 58% of all stigma expressed probe sets were shared among all stages. As expected, the next largest sections included overlapping expression between the first three (TET/UNM/BCP) or last three (UNM/BCP/TCP) consecutive stages of development. Interestingly, when looking at probe sets expressed in only one stage, there seemed to be a gradual increase in “stage-specific” probe sets as the stigma develops and there was a considerable difference between the TCP stage and the previous stages.

Considering the expression levels of co-expressed probe sets, the further apart the development stages, the least correlated their transcriptomes. Although the correlation values tend to be higher than those observed for anther stages, this trend was predominant throughout the entire analysis and again suggests that the transcriptomes diverge as the stigma matures. The stigma UNM and BCP stages appear functionally very alike, not

only do they have a similar number of expressed probe sets, but they also show a remarkably high correlation (0.9178).

As with the anther and the ovary, the majority of the stigma-expressed probe sets (82%) showed differential expression throughout development. Consistent with previous observations, the further apart the developmental stages the higher the number of up- or down-regulated probe sets. Not surprisingly, the UNM to BCP transition displayed the smallest number of differentially expressed genes. More than twice as many probe sets were up-regulated than down-regulated in the transition from TET to UNM, which may suggest an increase in gene activity associated with the beginning of stigma branch differentiation. Interestingly, the transition from BCP to TCP demonstrated relatively high numbers of both up- and down-regulated probe sets, and this again reflects a certain degree of molecular disparity as the stigma matures. Presumably, once the stigma is mature, many genes involved in its development are no longer required and the stigma patiently awaits pollen capture, possibly explaining the dramatic increase in down-regulated genes. Only when the stigma is pollinated does an increase in gene expression resume (Fujita et al., 2010).

Based on our stringent approach, we have identified 634 stigma-enriched and 117 stigma-specific probe sets. The TET stage shows both the lowest number of stigma-specific probe sets and the highest proportion of stage-specific probe sets (most with unknown functions, results not shown). Stigma morphology at the TET stage shows very limited development with the secondary stigmatic secondary branches appearing as only small nodules and one could speculate that this initiation process requires the expression of specific genes. On the other hand, the TCP stage accounts for 82% and 44% of the

stigma-enriched and -specific probe sets, respectively, in agreement with previous results pointing to some degree of specialization in the maturing stigma as it prepares for pollination. This also means that most of the large numbers of differentially expressed probe sets found during the BCP to TCP transition must therefore also be expressed elsewhere in the plant.

A total of 59 stigma-specific probe sets matched known genes and could be associated with a particular biological process. Even though their expression pattern was found to be very limited within the triticales plant, these genes covered a broad array of metabolic, cellular, and regulatory processes. The following discussion will focus on the major groups considered to have the largest impact on stigma development and its role in plant reproduction.

6.2.2 Stigma-specific transcription factors

A total of 9 transcription factors were found to be specifically expressed in the stigma representing one of the largest groups of related genes. Transcription-related proteins were also the most highly represented in the rice stigma (Li et al., 2007). The triticales stigma-specific transcription factors included MADS or B3 domain containing factors, as well as proteins related to previously characterized transcription factors such as YABBY, *roc1*, *ocl5*, *rolled expressed*, and *gata9*. There are two stigma-specific B3 domain containing transcription factors and this large family of proteins is involved in many plant processes including flowering (Romanel et al., 2009). Interestingly, *ocl5*, *roc1*, and *rolled expressed* are related homeodomain containing transcription factors containing START domains. The START domain is a lipid/sterol-binding domain shown

to be predominantly associated with homeodomain transcription factors in plants (Schrack et al., 2004). Homeodomain proteins are involved in organ development and differentiation, for example, *roc 1* was shown to express in a position dependent manner and believed to play an important role in epidermis differentiation (Ito et al., 2002). Similarly, the homeodomain MADS box transcription factors represent a large family of transcription factors important for floral organ identity (Heuer et al., 2001). In female tissues, ovule-specific MADS box genes have been identified and shown to have a regulatory role for proper ovule development (Angenent et al., 1995). The *YABBY* gene family is a small group of transcription factors known to specify cell fate in lateral organs (Bowman, 2000). Many characterized rice *YABBY* genes have been shown to regulate cell fate determination events such as differentiation or maintenance during early flower development (Toriba et al., 2007). It is very tempting to speculate that these stigma-specific transcription factors play an important role in triticales stigma development and differentiation; additional studies will be required to confirm this hypothesis.

6.2.3 Stigma-specific lipid metabolism

The next largest group of stigma-specific genes were involved with lipid metabolism and included three lipases (two class 3 lipases and one *gds1*-like), a lipid transport protein, a cyclopropane-fatty-acyl-phospholipid synthase, a diacylglycerol O-acyltransferase, a fatty acid desaturase, and three very-long-chain 3-ketoacyl-CoA synthases, one of which was identified as *CUT1*. Many of these proteins demonstrated high levels of expression throughout the entire stigma development and peaked at TCP (results not shown). Previous expression studies in *Arabidopsis* have also found these

types of lipid metabolism proteins in the stigma and they were shown to be induced by 30-40% when compared to other tissues (Swanson et al., 2005). Unfortunately, no studies characterizing the lipid constituents of developing cereal stigmas is available to explain the specific needs for these proteins in the stigma. However, the stigma papillae are known to be covered by a protective waxy lipid-rich cuticle layer which is also a barrier for fertilization (Tung et al., 2005). We identified a putative condensing enzyme shown to express specifically at TCP that is similar to CUT1 from Arabidopsis (Fiebig et al., 2000). CUT1 is identical to CER6 and is required for very long chain synthesis in pollen and in stems (Fiebig et al., 2000). This is the first study to report that a putative CER6/CUT1 gene may be specifically required for stigma cuticle formation. It will be very interesting to confirm the role of the putative CUT1 protein, as well as that of the other two very-long-chain 3-ketoacyl-CoA synthases, in stigma cuticle production.

6.2.4 Stigma-specific signaling proteins

Given its role in pollen-stigma interactions, genes involved in signaling might be expected to be abundant in the stigma. Arabidopsis stigma transcriptome studies have shown genes involved in cellular communication and signal transduction to be highly represented (Swanson et al., 2005; Tung et al., 2005). In our analysis, a limited number of stigma-specific genes could be ascribed a role in signaling. Nevertheless, two receptor-like protein kinases were identified and it will be interesting to ascertain whether they play a role in pollen-stigma interactions. In addition, two genes implicated the auxin signaling pathway, it is well established that this phytohormone plays a critical role in flower development (Cheng and Zhao, 2007). One of the genes encodes a member of the

B3 domain transcription factor family mentioned above, whereas the other encodes an auxin efflux carrier. In view of the stigma-specific expression of these genes, this would imply that auxin transport and regulation are important to triticales stigma development.

6.2.5 Stigma-specific defense response

Genes involved in defense responses made up another significant group of stigma-specific genes. These included a polygalacturonase inhibiting protein, a NBS-LRR disease resistance gene, a mlo family protein, and a NPR-1-like protein. These are all well characterized components of the plant defense arsenal, however it is unclear at this point why stigma-specific expression would be required. One possibility could be that their expression in the stigma is not pathogen-induced but is developmentally regulated as an added precautionary measure to protect an organ critical to plant reproduction.

6.3 Conclusion

We have successfully generated expression profiles of the developing stigma and are the first to identify Triticeae stigma stage-specific genes. A large proportion of these genes could not be identified and further characterization will undoubtedly reveal additional novel functions important to stigma development. A significant number of stigma-specific genes were similar to those identified in rice and Arabidopsis stigmas indicating conserved functions within the stigmas of monocots and dicots. The triticales stigma-specific genes represented a broad array of metabolic, regulatory, and cellular processes with the major groups related to regulation of transcription, lipid biosynthesis/

metabolism, cell-signaling, and defense response. The discovery of stigma-specific homeodomain transcription factors that may be crucial to stigma development and differentiation is of particular interest. Similarly, the identification of genes encoding stigma-specific enzymes possibly involved in cuticle development is also noteworthy. In our efforts to uncover the molecular networks involved in plant reproduction, we have generated the most comprehensive molecular study of stigma development of any species and hopefully contributed to future work attempting to unlock the mysteries of cereal female reproductive development.

7. Global Transcriptome Analysis of Triticale Reproductive Development

7.1 Results

7.1.1 Gene expression across reproductive tissues

Throughout triticale development, individual reproductive tissues (anther, ovary, stigma) and pollen are formed to carry out specialized functions to successfully complete the cycle of sexual reproduction. The expression of 31,550 probe sets was detected in at least one reproductive tissue at different developmental stages, covering 51.6% of the total Wheat GeneChip[®]. When comparing the 3 different reproductive organs (anther, stigma, ovary), the level of transcripts expressed at the TET, UNM, and BCP stages was quite similar (Figure 7-1). The gradual increase in the number of expressed probe sets as development progresses to reach a peak at the BCP stage and then decrease at the TCP stage was a consistent trend across all tissues. Although the stigma displayed a larger decrease in expressed probe sets at the TCP stage than the ovary, the anther showed the largest drop displaying the lowest amount of probe sets in any organ or stage. The mature pollen grain which is made up of only three cells showed 9,549 expressed probe sets. This is by far the lowest number of probe sets observed across all triticale tissues and stages.

7.1.2 Co-regulation of gene expression within male and female tissues

Figure 7-2 shows the overlapping expression of co-expressed genes in the different triticale reproductive organs. In all datasets, the largest section of co-expressing

Figure 7-1. Comparison of the number of expressed probe sets among different reproductive tissues at different developmental stages and including mature pollen. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

probe sets corresponded to the section including all stages of development. Typically, the closer in development the two stages being compared the larger the numbers of co-expressed genes, although exceptions occurred. When comparing all male datasets (Figure 7-2; AN/MP), only 6,921 probe sets were common throughout all stages of anther development and mature pollen. This is almost half of the number of probe sets common to all four stages of the anther (AN) and represented an even smaller fraction of the shared probe sets of all the female tissue stages (Figure 7-2; ST, OV). This lower number is likely due to the fact that the mature pollen transcriptome is considerably smaller than the other transcriptomes (Figure 7-1). The smaller TCP transcriptome may also explain why the anther section including all four stages remained smaller than that of the female tissues even in the absence of the MPG dataset (Figure 7-2; AN). The largest difference between the anther data in the presence of the MPG dataset was the reduction from 1,137 to 635 probe sets within the TCP stage, and from 11,506 to 6,921 for the TET/UNM/BCP/TCP (TeUBTr) section which likely resulted from a considerable degree of co-expression between TCP and MPG.

The stigma and ovary co-expressed transcriptomes across the different developmental stages displayed a rather similar overall pattern (Figure 7-2; ST, OV). However, the stigma tended to show significantly larger numbers of co-expressed genes mid-development (most sections including UNM and BCP), whereas the ovary generally displayed larger numbers than the stigma in the sections including later stages, especially TCP.

Although the dashboard model gives a good indication of shared transcriptomes across developmental stages, it does not take into account the level of expression.

Figure 7-2. Dartboard model showing the relationship between expressed probe sets during anther, anther and pollen, ovary, and stigma development. Numbers shown represent the number of expressed genes in individual or overlapping samples. AN, anther; MP, mature pollen; ST, stigma; OV, ovary; Te, tetrad; U, uninucleate microspore; B, bi-cellular pollen; Tr, tri-cellular pollen; P, mature pollen grain.

For each organ, the correlation among transcriptomes at different developmental stages was compared using scatter plots, and the linear regression for each developmental time sequence is shown in Figure 7-3. All three organs (anther, ovary, stigma) demonstrated a similar pattern of correlation between the different stages throughout development (Figure 7-3A). Consistently for each organ, the most highly correlated transcriptomes were UNM/BCP and the least were TET/TCP. The male transcriptome correlation coefficients were always smaller than those for each of the female tissues.

Comparisons between the first three anther stages (TET, UNM, BCP) and MPG revealed extremely low correlations $R^2 < 0.2$ (Figure 7-3B), in fact, the anther displayed higher correlation coefficients when compared with the female tissues (Figure 7-3A). However, the comparison between TCP and MPG showed a much higher correlation, almost as high as that of the UNM/BCP anther stages.

To gain insight in the degree of overlap in gene expression occurring across different organs, shared transcripts among different reproductive tissues were examined (Figure 7-4A). This included all the probe sets that were expressed in the anther, pollen, ovary, and stigma tissues at any stage of development. In total, 8,278 detected probe sets were shared by all four reproductive samples making up just above a quarter (26%) of the total number of expressed transcripts. The largest proportion (44%) of shared transcripts was found between the anther, ovary, and stigma tissues. Not surprisingly, when comparing female tissues to mature pollen probe sets that are not expressed in the anther, very few co-expressed transcripts are observed (Figure 7-4A). With regards to the

Figure 7-3. Pearson correlation coefficient of co-expressed probe sets at different developmental stages of A) the anther, ovary, and stigma, as well as between B) the anther and mature pollen. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

number of probe sets expressed in one tissue only, the anther displayed the most transcripts (1,958), followed by the ovary (1,356), stigma (1,046), and finally mature pollen (203). Interestingly, comparing the female tissues only, revealed that 85% of the expressed probe sets were shared (Figure 7-4B), whereas on the male side, only 34% were shared, while the majority of probe sets occurred in the anther tissue only, and just 1% were found only in the mature pollen grain (Figure 7-4C).

A more in-depth analysis was performed comparing the anther-, ovary-, and stigma-expressed probe sets at the four developmental stages. The relationship among the transcriptomes of the anther, ovary, and stigma at each stage is shown in Figure 7-5. Again, the highest number of co-expressed probe sets were those found in all three organs with an increase in shared probe sets at the BCP stage and a sizeable reduction at TCP. As expected, the number of co-expressed probe sets between the anther and either female organ was consistently lower than observed between the two female organs. This is especially evident at the TCP stage. Between the anther and ovary, the least number of shared probe sets was found at UNM (2.4%), and the highest at BCP (5.1%). In contrast, co-expressed probe sets between the anther and stigma were highest during UNM (3.8%), and lowest at the TCP stage (1.6%). In fact, the anther always shared more probe sets with the ovary than with the stigma except at the UNM stage. Comparison of overlapping probe sets between the female tissues at different stages, revealed interesting patterns of co-expression (Figure 7-5). During the TET stage, the stigma and ovary had the least number of overlapping probe sets (1,166). As development progressed, they maintained similar numbers of co-expressed probe sets during UNM (2,212) and BCP

(A)

(B)

OV ST
2,584 24,378 1,769

(C)

AN MP
17,588 9,233 316

Figure 7-4. Relationship between expressed transcripts in triticales reproductive tissues. (A) Venn diagram of expressed probe sets in the anther, mature pollen, ovary, and stigma tissue. Probe sets common and expressed in the (B) female (ovary, stigma) and (C) male (anther, mature pollen) tissues. AN, anther; MP, mature pollen; OV, ovary; ST, stigma.

(2,016) stages, but at the most mature stage TCP, the number of overlapping probe sets expanded to 6,129. This translates into the stigma and ovary transcriptomes being most divergent at the beginning of development (TET) and most similar at maturity (TCP). Narrowing down to probe sets expressed in one organ but not the other two, the anther showed the highest consistency in numbers of probe sets, with the most and least found at the BCP and TCP stages respectively (Figure 7-5C, D). Within the stigma only, the number of expressed probe sets varied in each stage with no discernable trend. The least number of probe sets was found at TET (649), and the most at UNM (1,800). A dramatic increase in the number of ovary transcripts occurred during the last two stages of development more than doubling the number at the previous stage. In fact, the number of ovary probe sets at TCP was the highest of any organ at any stage.

In order to simultaneously demonstrate the relationship among all transcriptomes from different organs and developmental stages while taking into consideration the expression levels of co-expressed probe sets, a principal component analysis (PCA) was performed (Figure 7-6). The major tendencies observed from the PCA plot were that the different organ transcriptomes are most similar at the early TET stage and most divergent at the late TCP stage; the stigma and ovary transcriptomes were relatively similar especially at the UNM stage; the developing anther transcriptomes formed a distinct group quite removed from all other transcriptomes; and finally, the mature pollen grain transcriptome diverged the most from any other transcriptome.

Figure 7-5. Relationship among reproductive tissues at four stages of pollen development. Venn diagram of the anther, ovary, and stigma datasets at the (A) TET, tetrad, (B) UNM, uninucleate microspore, (C) BCP, bi-cellular pollen, and (D) TCP, tri-cellular pollen stages.

Figure 7-6. Principal component analysis (PCA) plot of anther, ovary, and stigma transcriptomes across development, including mature pollen. A, anther; O, ovary; S, stigma; TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

7.1.3 Differentially expressed probe sets show similar tendencies in reproductive tissues

Probe sets considered differentially expressed must meet two criteria: 1) a cutoff P-value < 0.05 based on an ANOVA statistical test performed across all reproductive tissues, and 2) at least a 2-fold change (2-FC) in gene expression. A summary of the up-regulated differentially expressed probe sets (DEPs) for the anther, ovary, and stigma are summarized in Figure 7-7. Each table represents probe sets that showed at least a 2-FC up-regulation in pairwise comparison between stages within one particular organ. Furthermore, up-regulated DEPs showing an intensity value of ≥ 7 ($\log_2$) and therefore representing more 'reliable' signal values are also summarized for each tissue (Figure 7-7B). Generally, when comparing the four stages shared by the three reproductive organs, the further the developmental stage the more up-regulated DEPs were observed, the exception being the DEPs found when comparing anther TET-BCP (6,898) and TET-TCP (5,497). A pattern might also be invoked for all three organs when examining the DEPs throughout sequential stages (Figure 7-7, in red). In this case, the most DEPs occurred in the transition from TET to UNM, while going from UNM to BCP showed the least number of DEPs. From BCP to TCP, the number of DEPs increased again but not as high as initially observed with the TET to UNM transition. The numbers of DEPs in comparisons within the anther that included MPG as one of the datasets were quite distinct and tend to show low numbers (Figure 7-7, AN last column). The number of DEPs decreased with the proximity of the compared stage. DEPs that meet a cutoff value of ≥ 7 maintained the same trends as described above (Figure 7-7B). In the anther, an

	AN						OV					ST				
	TET	UNM	BCP	TCP	MPG		TET	UNM	BCP	TCP		TET	UNM	BCP	TCP	
(A)	TET	—	5,398	6,898	5,497	3,329	TET	—	4,398	5,711	6,484	TET	—	5,125	6,316	7,281
	UNM		—	3,700	4,640	3,284	UNM		—	1,743	3,672	UNM		—	1,686	4,994
	BCP			—	3,725	2,760	BCP			—	3,057	BCP			—	3,420
	TCP				—	1,330	TCP				—	TCP				—
	MPG					—										
(B)	TET	—	4,140	4,538	4,537	2,841	TET	—	3,298	3,486	4,129	TET	—	1,736	2,068	2,354
	UNM		—	2,126	3,807	2,782	UNM		—	505	1,871	UNM		—	439	1,725
	BCP			—	3,103	2,353	BCP			—	1,772	BCP			—	1,317
	TCP				—	982	TCP				—	TCP				—
	MPG					—										

Figure 7-7. Differentially expressed probe sets (DEPs) up-regulated throughout anther (including mature pollen), ovary, and stigma development. (A) Probe sets up-regulated at least 2-Fold Change (2-FC) at an adjusted P-value < 0.05. (B) Probe sets up-regulated at least 2-FC, P-value < 0.05, and expression value ≥ 7 ($\log_2$). Stages labeled on the top have at least a 2-FC increase in expression over the stages indicated on the left hand side. Numbers highlighted in red indicate up-regulated probe sets between consecutive developmental stages. AN, anther; OV, ovary; ST, stigma; TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

overall average of 77.7% of the DEPs were expressed at a $\log_2$ value ≥ 7 and it was 67.7% in the ovary. Within the stigma, only 56.3% of the DEPs had values ≥ 7 indicating that almost half of the differentially expressed probe sets express at low levels. Comparison of the trends with down-regulated DEPs was also performed across tissues. Down-regulated DEPs consist of probe sets that meet a P-value < 0.05 and have at least a 2-FC down-regulation in expression. A summary of the down-regulated DEPs including those that are filtered to meet an original intensity value ≥ 7 is shown in Figure 7-8. Much larger numbers of DEPs were observed in the anther relative to female tissues, and this is especially true when MPG was included in the comparison. Similar to what was observed with the up-regulated DEPs, generally the further apart the stages being compared the larger the number of down-regulated DEPs, with the exceptions being TCP/BCP in anther and ovary which showed relatively high numbers of DEPs even though they are consecutive stages. Again, the smallest numbers of DEPs occurred between UNM and BCP when comparing anther, ovary, and stigma tissues (Figure 7-8A). When focusing on the number of repressed probe sets at sequential stages (Figure 7-8A, in green), the largest number of down-regulated probe sets occurred at the transition to the most mature stage (BCP to TCP), whereas the smallest number occurred at the transition from the UNM to the BCP as mentioned above. When the probe sets were filtered to include only those whose expression values was ≥ 7 (Figure 7-8B), the pattern again reflected that observed in Figure 7-8A. The ovary maintained the lowest number of down-regulated DEPs overall, however, all tissues maintained a high percentage (76.8%, 71%, and 73% respectively) of down-regulated probe sets after

(A)

AN

OV

ST

	TET	UNM	BCP	TCP	MPG
TET	—	4,356	4,617	10,939	14,427
UNM		—	2,588	10,880	14,777
BCP			—	11,946	16,764
TCP				—	8,196
MPG					—

	TET	UNM	BCP	TCP
TET	—	2,399	2,226	3,614
UNM		—	442	2,261
BCP			—	2,541
TCP				—

	TET	UNM	BCP	TCP
TET	—	2,116	2,968	5,561
UNM		—	1,329	5,721
BCP			—	4,355
TCP				—

(B)

	TET	UNM	BCP	TCP	MPG
TET	—	3,283	3,718	8,506	11,119
UNM		—	2,137	8,517	11,749
BCP			—	7,964	11,668
TCP				—	6,133
MPG					—

	TET	UNM	BCP	TCP
TET	—	1,775	1,815	3,045
UNM		—	310	1,725
BCP			—	996
TCP				—

	TET	UNM	BCP	TCP
TET	—	1,614	2,460	4,454
UNM		—	865	4,028
BCP			—	2,726
TCP				—

Figure 7-8. Differentially expressed probe sets (DEPs) down-regulated throughout anther (including mature pollen), ovary, and stigma development. (A) Probe sets showing at least a 2-Fold Change (2-FC) down-regulation at an adjusted P-value < 0.05. (B) Probe sets down-regulated at least 2-FC from an original expression value ≥ 7 ($\log_2$), P-value < 0.05. Stages labeled on the top have at least a 2-FC decrease in expression over the stages indicated on the left hand side. Numbers highlighted in green indicate down-regulated probe sets between consecutive developmental stages. AN, anther; OV, ovary; ST, stigma; TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen, MPG, mature pollen grain.

filtering to those ≥ 7 . This indicates that the majority of down-regulated probe sets having a reliable expression value.

7.1.4 Distribution of enriched and specific transcripts throughout reproductive development

A detailed breakdown of the reproductive enriched and specific probe sets across development is shown in Figure 7-9. Noticeably, there were significantly more enriched probe sets found in the male as opposed to the female tissues. Nonetheless, the anther, ovary, and stigma all demonstrated a gradual increase in the number of Enriched probe sets right through development and peaked at the final TCP stage. The MPG had the most number of enriched probe sets (1,699) while the stigma at the TET stage showed the least (48). Similarly, probe sets that were Stage-Enriched were most abundant in the MPG, followed by the anther at the BCP stage, and within the female tissues, Stage-Enriched transcripts were highest at the TCP stage. As with the Enriched probe sets, the number of Specific probe sets increased with maturity and this is true for all reproductive tissues analyzed. Similarly, the number of specific probe sets was also significantly higher on the male side especially later in development. The highest number of Stage-Specific probe sets occurred in anther BCP (96), followed by the MPG (51).

7.1.5 Reproductive tissue-specific genes

The anther-only, anther + pollen-, ovary-, and stigma-specific probe sets were divided into eighteen different biological process categories based on gene ontology annotations and the results are summarized in Table 7-1. Due to the fact that very few mature pollen-only probe sets were identified, this data was not included in the

Figure 7-9. Summary of the number of anther-, pollen-, ovary-, and stigma-Enriched, Stage-Enriched, Specific, and Stage-Specific probe sets. Enriched probe sets have at least a 2-fold higher expression over all other different triticales tissues (leaf tissue at seedling, tillering, early boot, early and late senescence; coleoptile; root; stem tissue throughout development based on Zadoks' scale Z21, Z45, Z59, Z75; embryo; epidermis; endosperm; pericarp; crease; pollen; anther/ ovary/ stigma at TET, UNM, BCP, TCP stages (if applicable). Stage-Enriched probe sets are probe sets that are also enriched in a particular stage compared to other stages of the same organ. Specific probe sets correspond to Enriched probe sets with expression in only one organ. Stage-Specific probe sets are Specific probe sets that show no expression in other stages of the same organ. TET, tetrad; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain.

comparison. However, it should be kept in mind that the anther + pollen gene set represents many transcripts that are likely to be pollen-specific. In all cases, a considerable proportion of the probe sets were unknown or unclassified. Considerable differences can be observed when comparing the allocations of the different tissues to different biological processes. In the male tissues, transcripts in the anther-only dataset were mostly represented in the transport, metabolic process, regulation of cellular process, and protein metabolism, whereas the top categories in the anther + pollen dataset were carbohydrate metabolic process, metabolic process, regulation of cellular process, and transport. In the female tissues, the most highly represented categories in the ovary were macromolecule biosynthetic process, protein (de)phosphorylation, signaling, and protein metabolic process, while in the stigma lipid metabolic process, regulation of transcription, metabolic process, and response to stimulus was highly represented.

Biological Process	Anther-only	Anther + Pollen	Ovary	Stigma
Unknown	29.1	33.2	46.8	30.4
Unclassified	17.8	12.5	4.8	11.9
Transport	10.5	5.4	2.4	4.4
Signaling	1.2	1.7	4.8	3.7
Response to stimulus	2.9	4.5	2.4	5.2
Regulation of transcription	2.7	2.0	2.4	7.4
Regulation of cellular process	6.1	7.1	3.2	3.0
Protein metabolic process	5.9	2.3	4.8	3.0
Protein (de)phosphorylation	2.7	3.1	5.6	3.7
Nucleotide metabolic process	0.7	0.0	2.4	2.2
Metabolic process	6.8	8.2	4.0	6.7
Macromolecule biosynthetic process	2.9	4.5	7.3	3.0
Lipid metabolic process	3.4	1.1	0.8	7.4
Defense response	1.0	1.1	2.4	3.0
Cytoskeleton organization	0.2	1.7	0.8	2.2
Chromosome organization	0.7	0.9	2.4	1.5
Cell cycle	0.2	0.3	1.6	0.0
Carbohydrate metabolic process	4.9	10.2	0.8	1.5

Table 7-1. Summary of the biological process distribution of tissue-specific probe sets found in the anther-only, anther + pollen, ovary, and stigma. Numbers represent percentage (%) values within each tissue.

7.2 Discussion

Global gene expression was surveyed across all triticales reproductive organs during development. Due to the wealth of information generated, only selected topics will be addressed.

7.2.1 Male and female reproductive tissues display convergent and divergent expression patterns throughout development

Generally, the number of expressed probe sets in the anther, ovary, and stigma were quite similar to each other throughout development and had a tendency to most divergent at the mature stage (TCP). At maturity, the main function of the mature anther is to shed pollen for pollination and die, which can therefore account for the sizeable reduction in expressed probe sets. In contrast, the female tissues generally maintain a similar number of expressed probe sets, which is consistent with the fact that they still need to be very active during and following pollination. The mature pollen has a significantly lower number of expressed probe sets than the other three tissues and this is consistent with previous studies in rice (Wei, et al., 2010) and Arabidopsis (Honys & Twell, 2004) and reflects its diminutive nature.

Unexpectedly, the three reproductive organs generally showed similar trends in the numbers of overlapping probe sets, as well as in the levels of correlation of transcriptomes between stages (Figure 7-3). In all three organs, the highest correlation values were obtained for the UNM and BCP stages. These stages also displayed the largest degree of overlapping probe sets except in the ovary where the largest number occurred between BCP and TCP. It is possible that most of the ovary genes shared between the UNM and BCP stage also occur in another stage. Since the staging of the

female tissues was based on landmark stages of pollen development, we did not expect such coherence in developmental trends across the organs. In the Triticeae, wheat and barley have demonstrated a degree of synchrony between the duration time of male and female meiosis (Bennett et al., 1973). Later, the female gametophyte develops more rapidly and therefore their mitotic divisions are not synchronous, nonetheless, they are both regulated to mature at the same time (Yuan & Cai, 1983). It is likely that the trends observed reflect overall developmental phases, a first phase of differentiation required by the initiation of organ development, a middle phase of growth, and a final phase of specialization in preparation for pollination and fertilization.

Upon examination of the genes shared among the different organs, the anther, ovary, and stigma share the larger portion of their genes which likely consist of genes required for basic cellular function. The degree of sharing is more than double among the female tissues than the male tissues. Interestingly, the anther shares half as many probe sets with the stigma than the ovary and this may be due to their role in developing gametophytes. Most of the mature pollen probe sets are shared with the anther, ovary, and stigma. As expected, given that the anther contains the developing pollen, the next largest pollen overlapping section is with the anther. A number of different analyses have pointed to a divergence of transcriptomes later in development and accordingly when the anther, ovary, and stigma transcriptomes are compared throughout development, the least degree of overlap occurs at the TCP stage. This was well illustrated by the principal component analysis, which showed that the mature pollen grain and the anther, ovary, and stigma at the TCP stage were at very distant points from one another.

The majority of the genes expressed in a particular organ showed significant changes in expression during the course of its development. The highest number of up-regulated probe sets was consistently observed at the transition between the TET and UNM stage, and likely signals a major change in the programming of gene expression necessitated by early organ differentiation. The next highest numbers of up-regulated DEPs occurred in the transition from BCP to TCP in all organs and perhaps relate to the activation of genes required for maturation and pollination-related processes. The highest numbers of down-regulated DEPs also occurred simultaneously at the TCP stage in all organs, although more so in the anther which is essentially shutting down. When comparing the anther at the TCP stage which contains nearly mature pollen, with the released mature pollen grain, a relatively small number of probe sets were up-regulated which possibly represented mRNAs stored in readiness for pollen germination (Honys & Twell, 2004), but nearly half the probe sets were down-regulated and these are likely a combination of anther and pollen gene expression shut down.

A significantly larger number of enriched probe sets were found in the male tissues especially in mature pollen indicating unique roles within the male reproductive tissues. Generally, the highest number of enriched and specific probe sets occurred during the TCP stage in all organs, as well as in the mature pollen grain. These results are in complete agreement with previous results pointing to a higher degree of specialization at maturity. It also indicates that even though the transition from TET to UNM involves a large number of genes, these display a considerable lesser level of enrichment or specificity when compared to other triticales tissues.

7.2.2 Reproductive tissue-specific genes

Characterization of anther-, pollen-, ovary-, and stigma-specific genes based on biological process provided a glimpse of the different molecular networks necessary for pollination and fertilization events. Many of the genes represented were unique to a particular tissue suggesting that each reproductive organ, as well as the mature pollen grain, have particular molecular activities that reflect their specific roles. Nonetheless, a large proportion of the tissue-specific genes had unknown functions and this confirms the fact that there is much research to be done in order to characterize the molecular biology underlying Triticeae reproductive development. It is interesting to notice that when one looks at the top category of the different tissues that it often reflects what would be one of the major activity within this tissue, for example, anther-only/Transport: this data set is enriched for genes involved in early microspore development which relies heavily on “imports” from the tapetum especially during pollen cell wall biogenesis; anther + pollen/Carbohydrate metabolic process: this data set is enriched with late pollen transcripts and reflects the preparations made in anticipation of pollen germination and tube growth which involve starch reserve accumulation in the mature pollen, as well as cell wall modifying proteins; ovary/Macromolecule biosynthesis process: this may be reflecting activities in embryo sac development; stigma/Lipid metabolic process and Regulation of transcription: this probably reflects the fact that the stigma is covered with a multitude of papillae making up a large surface area for pollen capture that is coated by a cuticular protective layer. The fact that the stigma has such a large proportion of transcription factors is interesting and since most seemed to be involved with

organogenesis, this suggests that the stigma is a complex highly differentiated plant structure.

7.3 Conclusion

We have surveyed the transcriptomes of triticales reproductive tissues at four stages of development. Extensive gene expression profiling was carried out and anther-, mature pollen-, ovary-, and stigma-specific genes were identified. Many characteristic genes particular to each tissue were found, of those which matched known genes much work is required to ascertain their role in reproduction, of those which did not match known genes, a greater challenge lies ahead. Nevertheless, additional research in Triticeae reproductive development will undoubtedly uncover novel and important genes and expand our knowledge of the complex processes involved in plant reproduction.

References

1. Allen, A.M., Lexer, C., Hiscock, S.J. (2010) Comparative analysis of pistil transcriptomes reveals conserved and novel genes expressed in dry, wet, and semidry stigmas. *Plant Physiol.* 154:1347-1360.
2. Alves-Ferreira, M., Wellmer, F., Banhara, A., Kumar, V., Riechmann, J.L., Meyerowitz, E.M. (2007) Global expression profiling applied to the analysis of Arabidopsis stamen development. *Plant Physiol.* 145:747-762.
3. Alves, S.C., Worland, B., Thole, V., Snape, J.W., Bevan, M.W., Vain, P. (2009) A protocol for *Agrobacterium*-mediated transformation of *Brachypodium distachyon* community standard line Bd21. *Nat Prot.* 4:638-649.
4. Angenent, G.C., Franken, J., Busscher, M., Van Dijken, A., Van Went, J.L., Dons, H.J.M., Van Tunen, A.J. (1995) A novel class of MADS box genes is involved in ovule development of petunia. *Plant Cell.* 7:1569-1582.
5. Barskaya, E., Balina, N.V. (1971) The role of callose in plant anthers. *Soviet Plant Physiol.* 18:605-610.
6. Berger, F., Hamamura, Y., Ingouff, M., Higashiyama, T. (2008) Double fertilization—caught in the act. *Trends Plant Sci.* 13:437-443.
7. Bennett, M.D., Chapman, V., Riley, R. (1971) The duration of meiosis in pollen mother cells of wheat, rye and Triticale. *Phil Trans Proc R Soc.* 178:259-275.
8. Bennett, M.D., Finch, R.A., Smith, J.B., Rao, M.K. (1973) The time and duration of female meiosis in wheat, rye and barley. *Proc R Soc Lond B.* 183:301-319.
9. Bennett, M. D., Smith, J.B., Barclay, I. (1975) Early seed development in the Triticeae. *Phil Trans R Soc Lond.* 272:199-227.
10. Bhandari, N.N., Khosla, R. (1982) Development and histochemistry of anther in triticale cv Tri-1 I. Some new aspects in early ontogeny. *Phytomorph.* 32:18-27.
11. 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 Cell Physiol.* 46:87-98.
12. Bolstad, B.M., Irizarry R.A., Astrand, M., Speed, T.P. (2003), A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. *Bioinform.* 19:185-193.

13. Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijo, J.A., Becker, J.D. (2008) Comparative transcriptomics of *Arabidopsis thaliana* sperm cells. *Plant Physiol.* 148:1168-1181.
14. Bowman, J.L. (2000) The YABBY gene family and abaxial cell fate. *Curr Opin Plant Biol.* 3:17-22.
15. Chaubal, R., Reger, B.J. (1990) Relatively high calcium is localized in synergid cells of wheat ovaries. *Sex Plant Reprod.* 3:98-102.
16. Chen, F. and 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 Mol. Biol.* 35: 821–831.
17. Chen, W.R., Feng, Y., Chao, Y.E. (2008) Genomic analysis and expression pattern of *OsZIP1*, *OsZIP3*, and *OsZIP4* in two rice (*Oryza sativa* L.) genotypes with different zinc efficiency. *Russian J Plant Physiol.* 55:400-409.
18. Cheng, Y., Zhao, Y. (2007) A role for auxin in flower development. *J Integr Plant Biol.* 49:99-104.
19. Cheung, A.Y., Wu, H. (2008) Structural and signalling networks for the polar cell growth machinery in pollen tubes. *Annu Rev Plant Biol.* 59:547-572.
20. Choi, M.S., Kim, M.C., Yoo, Jae, H., Moon, B.C., Koo, S.C., Park, B.O. (2005) Isolation of a calmodulin-binding transcription factor from rice (*Oryza sativa* L.). *J Biol Chem.* 49:40820-40831.
21. Christensen, J.E., Horner, H.T. (1974) Pollen pore development and its spatial orientation during microsporogenesis in the grass *Sorghum bicolor*. *Amer J Bot.* 61:604-623.
22. Ciampolini, F., Shivanna, K.R., Cresti, M. (2001) Organization of stigma and transmitting tissue of rice, *Oryza sativa* (L.). *Plant Biol.* 3:149-155.
23. Clément, C., Chavant, L., Burrus, M., Audran, J.C. (1994) Anther starch variations in *Lilium* during pollen development. *Sex Plant Reprod.* 7: 347-356.
24. Coffey, M. T., Gerrits, W.J. (1988) Digestibility and feeding value of B858 triticale for swine. *J Anim Sci.* 66:2728-2735.
25. Cosgrove, D.J., Bedinger, P., Durachko, D.M. (1997) Group I allergens of grass pollen as cell wall-loosening agents. *Proc Nat Acad Sci USA.* 94:6559-6564.

26. Crismani, W., Baumann, U., Sutton, T., Shirley, N., Webster, T., Spangenberg, G., Langridge, P., Able, J.A. (2006) Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics*. 7: 267.
27. D'Agostino, N., Pizzichini, D., Chiusano, M.L., Giuliano, G. (2007) An EST database from saffron stigmas. *BMC Plant Biol.* 7:53.
28. Dobritsa, A.A., Nishikawa, S., Preuss, D., Urbanczyk-Wochniak, E., Sumner, L.W., Hammond, A., Carlson, A.L., Swanson, R.J. (2009) LAP3, a novel plant protein required for pollen development, is essential for proper exine formation. *Sex Plant Reprod.* 22:167-177.
29. Dobritsa, A.A., Lei, Z., Nishikawa, S., Urbanczyk-Wochniak, E., Huhman, D.V., Preuss, D., Sumer, L.W. (2010) *LAP5* and *LAP6* encode anther-specific proteins with similarity to chalcone synthase essential for pollen exine development in *Arabidopsis*. *Plant Physiol.* 153:937-955.
30. Doelling, J.H., Phillips, A.R., Soyler-Ogretim, G., Wise, J., Chandler, J., Callis, J., Otegui, M.S., Vierstra, R.D. (2007) The ubiquitin-specific protease subfamily UBP3/UBP4 is essential for pollen development and transmission in *Arabidopsis*. *Plant Physiol.* 145:801-813.
31. Domínguez, F., Cejudo, F.J. (1998) Germination –related genes encoding proteolytic enzymes are expressed in the nucellus of developing wheat grains. *Plant J.* 15:569-574.
32. Dresselhaus, T. (2006) Cell–cell communication during double fertilization. *Curr Opin Plant Biol.* 9:41-47.
33. Druka, A., Muehlbauer, G., Druka, I., Caldo, R., Baumann, U., Rostoks, N., Schreiber, A., Wise, R., Close, T., Kleinhofs, A., Graner, A., Schulman, A., Langridge, P., Sato, K., Hayes, P., McNicol, J., Marshall, D., Waugh, R. (2006) An atlas of gene expression from seed to seed through barley development. *Funct Integr Genomics*. 6:202-211.
34. Edlund, A.F., Swanson, R., Preuss, D. (2004) Pollen and stigma structure and function: the role of diversity in pollination. *Plant Cell*. 16:S84-S97.
35. Esau, K. (1965) Plant anatomy. New York, NY. John Wiley and Sons.
36. Feuillet, C., Penger, A., Gellner, K., Mast, A., Keller, B. (2001) Molecular evolution of receptor-like kinase genes in hexaploid wheat. Independent evolution of orthologs after polyploidization and mechanisms of local rearrangements at paralogous loci. *Plant Physiol.* 125:1304-1313.

37. Fiebig, A., Mayfield, J.A., Miley, N.L., Chau, S., Fischer, R.L., Preuss, D. (2000) Alterations in *CER6*, a gene identical to *CUT1*, differentially affect long-chain lipid content on the surface of pollen and stems. *Plant Cell*. 12:2001-2008.
38. Food and Agriculture Organization of the United Nations. (2009) FaoStat. Available from <http://faostat.fao.org/>.
39. Fujita, M., Horiuchi, Y., Ueda, Y., Mizuta, Y., Kubo, T., Yano, K., Yamaki, S., Tsuda, K., Nagata, T., Niihama, N., Kato, H., Kikuchi, S., Hamada, K., Mochizuki, T., Ishimizu, T., Iwai, H., Tsutsumi, N., Kurata, N. (2010) Rice expression atlas in reproductive development. *Plant Cell Physiol*. 51:2060-2081.
40. Goldberg, R.B., Beals, T.P., Sanders, P.W. (1993) Anther development: basic principles and practical applications. *Plant Cell*. 5:1217-1229.
41. Gomes, D., Agasse, A., Thiébaud, P., Delrot, S., Gerós, H., Chaumont, F. (2009) Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. *Biochim Biophys Acta*. 1788:1213-1228.
42. Greer, M.S., Kovalchuk, I., Eudes, F. (2009) Ammonium nitrate improves direct somatic embryogenesis and biolistic transformation of *Triticum aestivum*. *New Biotech*. 26:44-52.
43. Grienemberger, E., Kim, S.S., Lallemand, B., Geoffroy, P., Heintz, D., De Azevedo Souza, C., Heitz, T., Douglas, C.J., Legrand, M. (2010) Analysis of *TETRAKETIDE α -PYRONE REDUCTASE* function in *Arabidopsis thaliana* reveals a previously unknown, but conserved, biochemical pathway in sporopollenin monomer biosynthesis. *Plant Cell*. 22:4067-4083.
44. Haerizadeh, F., Wong, C.E., Bhalla, P.L., Gresshoff, P.M., Singh, M.B. (2009) Genomic expression profiling of mature soybean (*Glycine max*) pollen. *BMC Plant Biol*. 9:25.
45. Heslop-Harrison, J. (1978) Recognition and response in the pollen-stigma interaction. *Soc Exp Biol Sympos*. 32:121-138.
46. Heslop-Harrison, J. (1979a) An interpretation of the hydrodynamics of pollen. *Amer J Bot*. 66:737-743.
47. Heslop-Harrison, J. (1979b) Aspects of the structure, cytochemistry and germination of the pollen of rye (*Secale cereale* L.). *Ann Bot*. 44:1-47.
48. Heslop-Harrison, J., Heslop-Harrison, Y. (1975) Enzymatic removal of the proteinaceous pellicle of the stigma papilla prevents pollen tube entry in the Caryophyllaceae. *Ann Bot*. 39:163-165.

49. Heslop-Harrison, J., Heslop-Harrison, Y. (1980) The pollen-stigma interaction in the grasses. 1. Fine-structure and cytochemistry of the stigmas of hordeum and secale. *Acta Bot Neerl.* 29:261-276.
50. Heslop-Harrison, J., Heslop-Harrison, Y. (1981) The pollen-stigma interaction in the grasses. 2. Pollen-tube penetration and the stigma response in secale. *Acta Bot Neerl.* 30:289-307.
51. Heuer, S., Hansen, S., Bantin, J., Brettschneider, R., Kranz, H., Lörz, H., Dresselhaus, T. (2001) The maize MADS box gene *ZmMADS3* affects node number and spikelet development and is co-expressed with *ZmMADS1* during flower development, in egg cell, and in early embryogenesis. *Plant Physiol.* 127:33-45.
52. Hobo, T., Suwabe, K., Aya, K., Suzuki, G., Yao, K., Ishimizu, T., Fujita, M., Kikuchi, S., Hamada, K., Miyano, M., Fugioka, T., Kaneko, F., Kazama, T., Mizuta, Y., Takahashi, H., Shiono, K., Nakazono, M., Tsutsumi, N., Nagamura, Y., Kurata, N., Watanabe, M., Matsuoka, M. (2008) Various spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol.* 49:1417-1428.
53. Honys, D., Twell, D. (2004) Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol.* 5: R85.
54. Huang, M. Wei, F., Wu, C., Hsing, Y., Huang, A.H.C. (2009) Analyses of advanced rice anther transcriptomes reveal global tapetum secretory functions and potential proteins for lipid exine formation. *Plant Physiol.* 149: 694-707.
55. Ito, Y., Hirochika, H., Kurat, N. (2002) Organ-specific alternative transcripts of KNOX family class 2 homeobox genes in rice. *Gene.* 288:41-47.
56. Ito, M., Sentoku, N., Nishimura, A., Hong, S.-K., Sato, Y., Matsuoka, M. (2002) Position dependent expression of *GL2*-type homeobox gene, *Roc1*: significance for protoderm differentiation and radial pattern formation in early rice embryogenesis. *Plant J.* 29:497-507.
57. Iwakawa, H., Shinmyo, A., Sekine, M. Arabidopsis CDKA;1, a cdc2 homologue, controls proliferation of generative cells in male gametogenesis. *Plant J.* 45:919-831.
58. Jofuku, K.D., Boer, B.G.W.D., Montagu, M.V., Okamuro, J.K. (1994) Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell.* 6:1211-1225.
59. Johnson, A.M. (1931) Taxonomy of flowering plants. Appleton Century Crofts Inc.

60. Jones-Rhoades, M.W., Borevitz, J.O., Preuss, D. (2007) Genome-wide expression profiling of the Arabidopsis female gametophyte identifies families of small, secreted proteins. *PLOS Genet.* 3:e171.
61. Jung, K.H., Han, M.J., Lee, Y.S., Kim, Y.W., Lee, Y.S., Kim, Y.W., Hwang, I., Kim, M.J., Kim, Y.K., Nahm, B.H., An, G. (2005) Rice *Undeveloped tapetum1* is a major regulator of early tapetum development. *Plant Cell.* 17:2705-2722.
62. Kaothien, P., Ok, S.H., Shuai, B., Wengier, D., Cotter, R., Kelley, D., Kiriakopoulos, S., Muschietti, J., McCormick, S. 2005. Kinase partner protein interacts with the LePRK1 and LePRK2 receptor kinases and plays a role in polarized pollen tube growth. *Plant J.* 42:492-503.
63. Kashkush, K., Feldman, M., Levy, A.A. (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet.* 33:102-106.
64. Koonjul, P.K., Minhas, J.S., Nunes, C., Sheoran, I.S., Saini, H.S. (2005) Selective transcriptional down-regulation of anther invertases precedes the failure of pollen development in water-stressed wheat. *J Exp Bot.* 56:179-190.
65. Krichevsky, A., Kozlovsky, S.V., Tian, G., Chen, M., Zaltsman, A., Citovsky, V. (2007) How pollen tubes grow. *Dev Biol.* 303:405-420.
66. Kuleung, C., Baenziger, P.S., Dweikat, I. (2004) Transferability of SSR markers among wheat, rye, and triticale. *Theor Appl Genet.* 108:1147-1150.
67. Lalonde, S., Beebe, D.U., Saini, H.S. (1997) Early signs of disruption of wheat anther development associated with the induction of male sterility by meiotic-stage water deficit. *Sex Plant Reprod.* 10:40-48.
68. Langridge, P., Bauman, U. (2008) Self-incompatibility in the grasses. In: *Self-Incompatibility in Flowering Plants – Evolution, Diversity and Mechanisms.* Berlin, Springer, pp. 275-287.
69. Lauga, B., Charbonnel-Campaa, L., Combes, D. (2000) Characterization of MZm3-3, a *Zea mays* tapetum-specific transcript. *Plant Sci.* 157:65-75.
70. Lee, S., Jung, K.H., An, G., Chung, Y.Y (2004) Isolation and characterization of a rice cysteine protease gene, OsCP1, using T-DNA gene-trap system. *Plant Mol Biol.* 54:755-765.
71. Lersten, N.R. (2004) *Flowering Plant Embryology.* Ames, Iowa: Blackwell Publishing.

72. Li, S.F., Higginson, T., Parish, R.W. (1999) A novel MYB-related gene from *Arabidopsis thaliana* expressed in developing anthers. *Plant Cell Physiol.* 40:343-347.
73. Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., Franke, R., Zhang, P., Chen, L., Gao, Y., Liang, W., Zhang, D. (2010a) Cytochrome P450 member CYP704B2 catalyzes the ω -hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *Plant Cell.* 22:173-190.
74. Li, M., Xu, W., Yang, W., Kong, Z., Xue, Y. (2007) Genome-wide gene expression profiling reveals conserved and novel molecular functions of the stigma in rice. *Plant Physiol.* 144:1797-1812.
75. Li, S., van Os, G.M.A., Ren, S., Yu, D., Ketelaar, T., Emons, A.M.C., Liu, C.M. (2010b) Expression and functional analyses of *EXO70* genes in *Arabidopsis* implicate their roles in regulating cell type-specific exocytosis. *Plant Physiol.* 154:1819-1830.
76. Li, N., Zhang, D.S., Liu, H.S., Yin, C.S., Li, X.X., Liang, W.Q., Yuan, Z., Xu, B., Chu, H.W., Wang, J., Wen, T.Q., Huang, H., Luo, D., Ma, H., Zhang, D.B. (2006) The rice *tapetum degeneration retardation* gene is required for tapetum degeneration and anther development. *Plant Cell.* 18:2999-3014.
77. Linnestad, C., Doan, D.N.P., Brown, R.C., Lemmon, B.E., Méyer, D.J., Jung, R., Olsen, O.A. (1998) Nucellain, a barley homolog of the dicot vacuolar-processing protease, is localized in nucellar cell walls. *Plant Physiol.* 118: 1169–1180.
78. Ma, X.F., Fang, P., Gustafson, J.P. (2004) Polyploidization-induced genome variation in triticales. *Genome.* 47: 839-848.
79. Ma, J., Duncan, D., Morrow, D.J., Fernandes, J., Walbot, V. (2007) Transcriptome profiling of maize anthers using genetic ablation to analyze pre-meiotic and tapetal cell types. *Plant J.* 50:637:648.
80. Ma, J., Skibbe, D.S., Fernandes, J., Walbot, V. (2008) Male reproductive development: gene expression profiling of maize anther and pollen ontogeny. *Genome Biol.* 9:R181.
81. Malhó, R. (1998) Pollen tube guidance – the long and winding road. *Sex Plant Reprod.* 11:242-244.
82. McCormick, S. (1993) Male gametophyte development. *Plant Cell.* 5:1265-1275.
83. McCormick, S. (2004) Control of male gametophyte development. *Plant Cell.* 16:S142-153.

84. Nonomura, K., Miyoshi, K., Eiguchi, M., Suzuki, T., Miyao, A., Hirochika, H., Kurata, N. (2003) The MSP1 gene is necessary to restrict the number of cells entering into male and female sporogenesis and to initiate anther wall formation in rice. *Plant Cell*. 15:1728-1739.
85. Oettler, G. (2005) The fortune of a botanical curiosity- Triticale: past, present and future. *J Agr Sci*. 143:329:346.
86. Okada, T., Bhalla, P.L., Singh, M.B. (2006) Expressed sequence tag analysis of *Lilium longiflorum* generative cells. *Plant Cell Physiol*. 47:698-705
87. Okada, K., Komaki, M.K., Shimura, Y. (1989) Mutational analysis of pistil structure and development of *Arabidopsis thaliana*. *Cell Differ Dev*. 28:27-38.
88. op den Camp, R.L., Kuhlemeier, C. (1997) Aldehyde dehydrogenase in tobacco pollen. *Plant Mol Biol*. 35: 355-365.
89. Palanivelu, R., Brass, L., Edlund, A.F., Preuss, D. (2003) Pollen tube growth and guidance is regulated by POP2, an Arabidopsis gene that controls GABA levels. *Cell*. 114:47-59.
90. Pearce, R.B., Strange, R.N., and Smith, H. (1976) Glycine betaine and choline in wheat; distribution and relation to infection by *Fusarium graminearum*. *Phytochem*. 15:953–954.
91. Penniket, C., Selinger, B., Laroche, A. (2009) Localization and characterization of tissue-specific gene expression patterns in triticale. Canadian Plant Genomics Workshop. Saskatoon, SK. August 24-27, 2009.
92. Pina, C., Pinto, F., Feijo, J.A., Becker, J. (2005) Gene family analysis of Arabidopsis pollen transcriptome reveals biological implication for cell growth, division control and gene expression regulation. *Plant Physiol*. 138:744-756.
93. Pressey, R., Reger, B.J. (1989) Polygalacturonase in pollen from corn and other grasses. *Plant Sci*. 59:57-62.
94. Preuss, D., Lemieux, D., Yen, G., Davis, R.W. (1993) A conditional sterile mutation eliminates surface components from *Arabidopsis* pollen and disrupts cell signalling during fertilization. *Genes Devel*. 7:974-985.
95. Quiapim, A.C., Brito, M.S., Bernardes, L.A., Dasilva, I., Malavazi, I., DePaoli, H.C., Molfetta-Machado, J.B., Giuliatti, S., Goldman, G.H., Goldman, M.H. (2009) Analysis of the *Nicotiana tabacum* stigma/style transcriptome reveals gene expression differences between wet and dry stigma species. *Plant Physiol*. 149:1211-1230.

96. Ray, S., Agarwal, P., Arora, R., Kapoor, S., Tyagi, A.K. (2007) Expression analysis of calcium-dependent protein kinase gene family during reproductive development and abiotic stress conditions in rice (*Oryza sativa* L. ssp. *indica*). *Mol Genet Genomics*. 278:493-505.
97. Robinson-Beers, K., Pruitt, R.E., Gasser, C.S. (1992) Ovule development in wild-type *Arabidopsis* and two female-sterile mutants. *Plant Cell*. 4:1237-1249.
98. Romanel, E.A.C., Schrago, C.G., Counago, R.M., Russo, C.A.M., Alves Ferreira, M. (2009) Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. *PLoS One*. 4:e5791.
99. Scott, R.J., Spielman, M., Dickinson, H.G. (2004) Stamen structure and function. *Plant Cell*. 16:S46-S60.
100. Schneitz, K., Hulskamp, M., Pruitt, R.E. (1995) Wild-type ovule development in *Arabidopsis thaliana*: a light microscope study of cleared whole-mount tissue. *Plant J*. 7:731-749.
101. Schopfer, P. (2001) Hydroxyl radical-induced cell-wall loosening *in vitro* and *in vivo*: implications for the control of elongation growth. *Plant J*. 28: 679-688.
102. Schreiber, A.W., Sutton, T., Caldo, R.A., Kalashyan, E., Lovell, B., Mayo, G., Muehlbauer, G.J., Druka, A., Waugh, R., Wise, R.P., Langridge, P., Baumann, U. (2009) Comparative transcriptomics in the Triticeae. *BMC Genomics*. 10:285.
103. Schwacke, R., Grallath, S., Breitzkreuz, K.E., Stransky, E., Stransky, H., Frommer, W.B., Rentsch, D. (1999) LeProT1, a transporter of proline, glycine betaine, and (gamma)-amino butyric acid in tomato pollen. *Plant Cell*. 11:377-392.
104. Skibbe, D.S., Wang, X., Borsuk, L.A., Ashlock, D.A., Nettleton, D., Schnable, P.S. (2008) Floret-specific differences in gene expression and support for the hypothesis that tapetal degeneration of *Zea mays* L. occurs *via* programmed cell death. *J Genet Genomics*. 35:603-616.
105. Skinner, D.J., Gasser, C.S. (2009) Expression-based discovery of candidate ovule development regulators through transcriptional profiling of ovule mutants. *BMC Plant Biol*. 9:29.
106. Smalle, J., Vierstra, R.D. (2004) The ubiquitin 26S proteasome proteolytic pathway. *Annu Rev Plant Biol*. 55:555-590.

107. Song, J., Nada, K., Tachibana, S. (2002) Suppression of S-adenosylmethionine decarboxylase activity is a major cause for high-temperature inhibition of pollen germination and tube growth in tomato (*Lycopersicon esculentum* Mill.). *Plant Cell Physiol.* 43:619-627.
108. Sprunck, S., Baumann, U., Edwards, K., Langridge, P., Dresselhaus, T. (2005) The transcript composition of egg cells changes significantly following fertilization in wheat (*Triticum aestivum*). *Plant J.* 41:660-672.
109. Steffen, J.G., Kang, I.-H., Macfarlane, J., Drews, G.N. (2007) Identification of genes expressed in the Arabidopsis female gametophyte. *Plant J.* 51:281-292.
110. Stone, J.M., Trotochaud, A.E., Walker, J.C., Clark, S.E. (1998) Control of meristem development by CLAVATA1 receptor kinase and kinase-associated protein phosphatase interactions. *Plant Physiol.* 117:1217-1225.
111. Suwabe, K., Suzuki, G., Takahashi, H., Shiono, K., Endo, M., Yano, K., Fujita, M., Masuko, H., Saito, H., Fujioka, T., Kaneko, F., Kazama, T., Mizuta, Y., Kawagishi-Kobayashi, M., Tsutsumi, N., Kurata, N., Nakazono, M., Watanabe, M. (2008) Separated transcriptomes of male gametophyte and tapetum in rice: validity of a laser microdissection (LM) microarray. *Plant Cell Physiol.* 49:1406-1416.
112. Swanson, R., Clark, T., Preuss, D. (2005) Expression profiling of *Arabidopsis* stigma tissue identifies stigma-specific genes. *Sex Plant Reprod.* 18:163-171.
113. Swanson, R., Edlund, A. F., Preuss, D. (2004) Species specificity in pollen-pistil interactions. *Annu Rev Genet.* 38:793-818.
114. Takahashi, M., Terada, Y., Nakai, I., Nakanishi, H., Yoshimura, E., Mori, S., Nishizawa, N.K. (2003) Role of nicotianamine in the intracellular delivery of metals and plant reproductive development. *Plant Cell.* 15:1263-1280.
115. Tang, J., Erikson, R.L., Liu, X. (2006) Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like kinase 1 (Plk1). *PNAS.* 103:11964-11969.
116. Tang, X., Zhang, Z.Y., Zhang, W.J., Zhao, X.M., Li, X., Zhang, D., Liu, Q.Q., Tang, W.H. (2010) Global gene profiling of laser-captured pollen mother cells indicates molecular pathways and gene subfamilies involved in rice meiosis. *Plant Physiol.* 154:1855-1870.
117. Taylor, L.P., Hepler, P.K. (1997) Pollen germination and tube growth. *Annu Rev Plant Physiol Plant Mol Biol.* 48:461-491.

118. Teng, N., Huang, Z., Mu, X., Jin, B., Hu, Y., Lin, J. (2005) Microsporogenesis and pollen development in *Leymus chinensis* with emphasis on dynamic changes in callose deposition. *Flora*. 200:256-263.
119. Thoms, S., Erdmann, R. (2005) Dynamin-related proteins and Pex11 proteins in peroxisome division and proliferation. *FEBS J*. 272:5169-5181.
120. Tibshirani, R., Walther, G., Hastie, T. (2001) Estimating the number of clusters in a data set via the gap statistic. *J R Statist Soc B*. 63:441-423.
121. Toriba, T., Harada, K., Takamura, A., Nakamura, H., Ichikawa, H., Suzuki, T., Hirano, H. (2007) Molecular characterization of the *YABBY* gene family in *Oryza sativa* and expression analysis of *OsYABBY1*. *Mol Gen Genomics*. 277:457-468.
122. Tung, C.W., Dwyer, K.G., Nasrallah, M.E., Nasrallah, J.B. (2005) Genome-wide identification of genes expressed in Arabidopsis pistils specifically along the path of pollen tube growth. *Plant Physiol*. 138:977-989.
123. Vigneault, F., Lachance, D., Cloutier, M., Pelletier, G., Levasseur, C., Séguin, A. (2007) Members of the plant NIMA-related kinases are involved in organ development and vascularisation in poplar, Arabidopsis and rice. *Plant J*. 51:575-588.
124. Vishnyakova, M.A., Willemse, M.T.M. (1994) Pollen-pistil interaction in wheat. *Acta Bot Neerl*. 43:51-64.
125. Vogel, J., Hill, T. (2008) High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. *Plant Cell Rep*. 27:471-478.
126. Wang, A., Xia, Q., Xie, W., Datla, R., Selvaraj, G. (2003) The classical Ubish bodies carry a sporophytically produced structural protein (RAFTIN) that is essential for pollen development. *PNAS*. 24: 14487-14492.
127. Wang, J-R., Wang, L., Gulden, S., Rocheleau, H., Balcerzak, M., Hattori, J., Cao, W., Han, F., Zheng, Y.-L., Fedak, G., Ouellet, T. (2010) RNA profiling of fusarium head blight-resistant wheat addition lines containing the *Thinopyrum elongatum* chromosome 7E. *Can J Plant Pathol*. 32:188-214.
128. Wei, L.Q, Xu, W.Y., Deng, Z.Y., Xue, Y., Wang, T. (2010) Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in *Oryza sativa*. *BMC Genomics*. 11:338.

129. Wijeratne, A.J., Zhang, W., Sun., Y., Liu, W., Albert, R., Zheng, Z., Oppenheimer, D.G., Zhao, D., Ma, H. (2007) Differential gene expression in *Arabidopsis* wild-type and mutant anthers: insights into anther cell differentiation and regulatory networks. *Plant J.* 52:14-29.
130. Wu, H.M., Cheung, A.Y. (2000) Programmed cell death in plant reproduction. *Plant Mol Biol.* 44:267-281.
131. Wuest, S.E., Vijverberg, K., Schmidt, A., Weiss, M., Gheyselinck, J., Lohr, M., Wellmer, F., Rahnenfuhrer, J., Mering, C., Grossniklaus, U. (2010) *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. *Curr Biol.* 20:506-512.
132. Xi, X.Y., Cui, D.C., (1983) Relationship between pollen and embryo sac development in wheat, *Triticum aestivum* L. *Bot Gaz.* 144:191-200.
133. Xu, Y., Badea, C., Tran, F., Frick, M., Schneiderman, D., Robert, L., Harris, L., Thomas, D., Tinker, N., Gaudet, D., Laroche, A. (2011) Next-Gen sequencing of the transcriptome of triticale. *Plant Genet Res.* DOI: 10.1017/S1479262111000281.
134. Yang, K., Xia, C., Liu, X., Dou, X., Wang, W., Chen, L., Zhang, X., Xie, L., He, L., Ma, X., Ye, D. (2009) A mutation in *THERMOSENSITIVE MALE STERILE 1*, encoding a heat shock protein with DnaJ and PDI domains, leads to thermosensitive gametophytic male sterility in *Arabidopsis*. *Plant J.* 57:870-882.
135. Yeats TH, Rose J.K. (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Sci.* 17:191–198.
136. Yu, H-J., Hogan, P., Sundaresan, V. (2005) Analysis of the female gametophyte transcriptome of *Arabidopsis* by comparative expression profiling. *Plant Physiol.* 139:1853-1869.
137. Zadoks, J.C., Chang, T.T., Konzak, C.F. (1974) A decimal code for the growth stages of cereals. *Weed Res.* 14:415-421.
138. Zhang, D., Liang, W., Yin, C., Zong, Y.J., Gu, F., Zhang, D. (2009) *OsC6*, encoding a lipid transfer protein, is required for postmeiotic anther development in rice. *Plant Physiol.* 154:149-162.
139. Zhao, D.Z., Wang, G.F., Speal, B., Ma, H. (2002). The *excess microsporocytes1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. *Genes Dev.* 16:2021–2031.

140. Zinkl, G.M., Zweibel, B.I., Grier, D.G., Preuss, D. (1999) Pollen-stigma adhesion in *Arabidopsis*: a species-specific interaction mediated by hydrophobic molecules in the pollen exine. *Devel.* 126:5431-5440.