DEVELOPMENTAL RESPONSES TO POLYCOMB COMPLEX MUTATIONS AND PLANTING DENSITY IN HEXAPLOID WHEAT

By

Selena Lopez

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Breeding, Genetics and Biotechnology – Crop and Soil Sciences – Master of Science

2021

ABSTRACT

DEVELOPMENTAL RESPONSES TO POLYCOMB COMPLEX MUTATIONS AND PLANTING DENSITY IN HEXAPLOID WHEAT

By

Selena Lopez

Hexaploid wheat (*Triticum aestivum*) has undergone two hybridization events resulting in severe genetic bottlenecks, particularly in the D genome contributed by the diploid progenitor species, *Aegilops tauschii*. One method to regain this genetic diversity is through interspecific hybrids with diploid progenitors. This requires embryo rescues to recover progeny due to endosperm failure of the hybrids. Previous studies in *Arabidopsis* and rice have demonstrated that mutations in imprinted Polycomb Repressive Complex 2 (PRC2) genes *medea* (*MEA*), *fertilization independent endosperm* (*FIE*), and *fertilization independent seed 2* (*FIS2*) can, or potentially can, restore endosperm function to interploidy hybrids. Homologs of *MEA*, *FIE*, and *FIS2* were identified in the wheat A, B and D genomes. Pyramids of double and triple mutants were developed using a combination of Kompetative allele specific PCR (KASP) markers and amplicon sequencing. Single, double, and triple mutants at *MEA*, *FIE*, and *FIS2* associated with the three subgenomes were hybridized with *Ae. tauschii* to identify mutant combinations that may restore endosperm function.

The wheat breeding program at Michigan State University utilizes a modified bulk breeding method to rapidly advance populations from the F2 to F4 stage under greenhouse conditions. To determine the influence of planting density on yield components and plant architecture in greenhouse settings, two experiments were carried out using two spring and winter wheat. Experiments with both spring and winter wheat demonstrated that increased planting density reduces plant height, spikelets per spike, grain length and grain width.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Modern wheat

Wheat (*Triticum aestivum*, 2n=6x=42, AABBDD) is a food source that billions of people rely on for calories and nutrition. In 2019, 35,468 hg/ha of wheat was harvested from 2.159e8 hectares worldwide (FAOSTAT 2020). Wheat accounts for 20% of protein and 28% of calories consumed in people's diets (Velu et al., 2016). As the second most produced cereal in the world, wheat is a major component in the agricultural market (Velu et al., 2016).

Modern day wheat is an allohexaploid consisting of an A, B, and D genomes (AABBDD), which contain a total of 17Gb of DNA sequences. The A genome is most related to *Triticum urartu* (2n = 2x = 14), the B genome is a close relative of *Aegilops speltoides* (2n = 14x, SS), and the D genome is *Aegilops tauschii* (2n=2x=14, DD) or goat grass (Ling et al., 2013); these genomes were merged over two separate hybridization events (Figure 1.1). The first hybridization event occurred 300,000-500,000 years ago between the A and B genome forming *Triticum turgidum* (AABB) (Huang et al., 2002). Tetraploid wheat (*T. turgidum*, AABB) is used in modern agriculture for pasta as well as breeding synthetic hexaploid wheat (Zhang et al., 2008). The second hybridization occurred between tetraploid wheat (AABB) and *Ae. tauschii* (DD) to form modern hexaploid wheat, *T. aestivum* (AABBDD) (Mcfadden and Sears 1946).

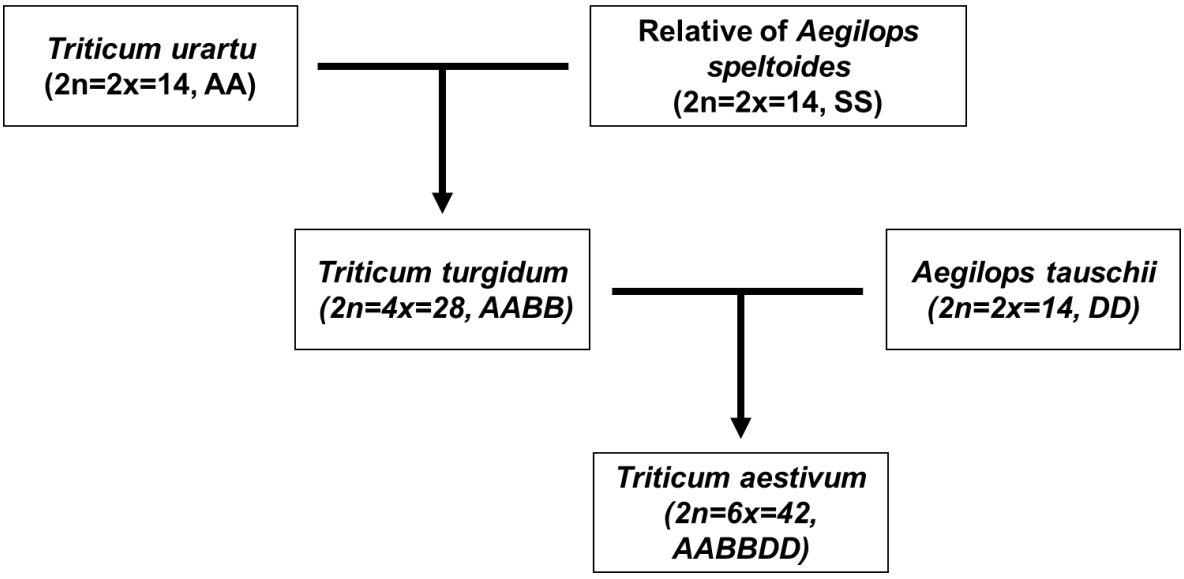


Figure 1.1: Hybridizations that formed hexaploid wheat.

***Aegilops tauschii* (D genome)**

Ae. tauschii (2n=2x=14), or goat grass, has a center of origin in Iran containing a mixture of subspecies and lineages of *Ae. tauschii* (Singh et al. 2019). The subspecies recognized as most related to the D genome of hexaploid wheat, *Ae. tauschii ssp. strangulata*, ranges from Transcaucasia to eastern Caspian Iran (Wang et al., 2013). Studies have indicated that other lines outside of this subspecies, including *ssp. tauschii*, are also significantly related to the D subgenome of hexaploid wheat (Dvorak et al., 1998). Two lineages of *Ae. tauschii* exist: *Ae. tauschii ssp. tauschii* belonging in lineage 1 (L1) and *Ae. tauschii ssp. strangulata* belonging in lineage 2 (L2) (Mizuno et al. 2010). These lineages most likely diverged based on geographic location. Genetic analysis determined that L2 accessions generally originate in western regions, and L1 accessions generally originate in eastern regions such as Afghanistan and Pakistan (Takumi et al., 2008).

Only a few plants of *Ae. tauschii* hybridized with wheat, resulting in a very narrow genetic background of the wheat D genome (Lagudah et al., 1991). By analyzing the geographic and genetic diversity of *Ae. tauschii*, desired traits from specific *Ae. tauschii* accessions can be introgressed into hexaploid wheat. Genes conferring disease or insect resistance to leaf rust, stem rust, yellow rust, Hessian fly, greenbug, cereal leaf beetle, root lesion nematode, karnal bunt, soil borne mosaic virus, and powdery mildew have been successfully introgressed to hexaploid wheat from *Ae. tauschii* (Miranda et al., 2006; Olson et al., 2013; Cox et al., 2017; Saluja et al., 2018). In recent years, *Pm58*, a powdery mildew resistance gene, was successfully introgressed into hexaploid wheat from *Ae. tauschii* (Wiersma et al., 2017). Future genes of interest can be exploited in *Ae. tauschii*, such as temperature adaptations in the differing altitudes of L2 (Singh et al., 2019). L1 accessions are also a valuable source of genetic diversity as L1 only contributes 2.7% introgression from *Ae. tauschii* in United States winter wheat lines (Singh et al., 2019).

Endosperm formation and *Ae. tauschii* introgression

Double fertilization is an essential process for most angiosperms. Two male gametes are necessary: one to combine with the female egg cell, and the other to combine with the two polar nuclei. The latter forms the endosperm portion of the seed. Endosperm is the primary food source for the germinating embryo. The resulting combination of two polar nuclei with the sperm is a triploid (3n) endosperm. This 2:1 ratio of maternal and paternal genetic material is required for proper formation of endosperm.

There are two methods of introgressing *Ae. tauschii* into hexaploid wheat: direct hybridization and synthetic hexaploid wheat. In direct hybridizations, the endosperm ratio is severely off balance from the 2:1 maternal to paternal dosage in the endosperm. A hexaploid

wheat and *Ae. tauschii* direct hybridization would have a 6:1 maternal to paternal dosage resulting in failed endosperm formation and eventual seed abortion. Research involving these direct hybridizations typically employ embryo rescue and subsequent tissue culture of the rescued young embryo (Shen et al., 2011). Synthetic hexaploid wheat is created by hybridizing *T. turgidum* (AABB) with *Ae. tauschii* (DD) and then undergoes spontaneous or induced chromosome doubling (Cox et al., 2017). Synthetic hexaploid wheat lines are often spring type and non-free threshing (Cox et al., 2017). Additionally, changes can occur in the genome structure and/or copy numbers of genes in synthetic hexaploids (Li et al., 2018).

Polycomb repressive complex 2

A major mechanism responsible for endosperm formation is Polycomb Repressive Complex 2 (PRC2), one of several polycomb group complexes (Holec and Berger, 2012). These complexes were first discovered through research in segmentation of *Drosophila melanogaster* (Lewis, 2004). Later research determined homologs of PRC2 are also found in mammals, angiosperms, and some unicellular organisms (Margueron and Reinberg, 2011). PRC2 is not entirely conserved as some yeast species do not have the complex (Margueron and Reinberg, 2011). PRC2 catalyzes trimethylation of histone H3 on lysine 27 (H3K27me3) allowing epigenetic activity to occur in select vegetative and reproductive development (Weinhofer et al., 2010). Homology searches using online resources were carried out between *Arabidopsis* PRC2 genes and wheat. Based on these searches, it was found that the wheat PRC2 complex consists of three genes: *fertilization independent endosperm* (*FIE*), *fertilization independent seed* (*FIS2*), and *Medea* (*MEA*); these genes are maternally expressed (Köhler et al., 2003a).

PRC2 genes

Several studies have researched the actions of PRC2 genes. One found that increasing maternal dosage in *Arabidopsis* can bypass endosperm failure (Kradolfer et al., 2013). The study analyzed *MEA* and *FIS2* mutations. Both *MEA* and *FIS2* mutants formed enlarged seeds, some of which were viable and can be germinated. Imprinting regulation involves both *FIS2* and *MEA* (Köhler and Makarevich, 2006). These imprinted genes are silenced differently in *Arabidopsis*. *MEA* is silenced by H3K27 trimethylation from PRC2 and *FIS2* is silenced through DNA methylation or, more specifically, methyltransferase 1 (MET1) (Jullien et al., 2008). The methylation *FIS2* experiences is also controlled by the retinoblastoma pathway as it represses MET1 when the female gamete is forming (Jullien et al., 2008).

Medea (MEA) interacts with *DNA methyltransferase (MET1)* to repress autonomous endosperm in *Arabidopsis* (Schmidt et al., 2013). *MET1* is not orthologous to wheat, therefore altering the function of *MEA* alone may impose autonomous endosperm formation in wheat. *MEA* is known to down-regulate maternal alleles during fertilization and silences paternal alleles during seed formation (Baroux et al., 2006). Specifically, *MEA* represses expression of the *PHERES1 (PHE1)* MADS box gene as well as other genes that are expressed during the first stages of endosperm formation (Jarillo et al., 2009). Both *MEA* and *FIE* associate with the promoter of *PHE1*, meaning *PHE1* expression is likely epigenetically regulated by PRC2 proteins (Köhler et al., 2003b). *MEA* will bind to its own promoter prior to fertilization (Baroux et al., 2006). This binding might not include any other PRC2 proteins, due to unchanged *MEA* levels in *FIS2* and *FIE* mutants.

FIE proteins are in a WD repeat protein family (Nallamilli et al., 2013). The WD repeats form a beta propeller structure, a circular structure that creates surface loops used as the basis of

protein complex formation (Yadegari et al., 2000). This family promotes protein-protein interactions. *FIE* functions in seed development, grain filling, and may also function in male gametophyte development (Nallamilli et al., 2013). In Rice, the *FIE* protein contains a retinoblastoma protein (RB) region (Nallamilli et al., 2013).

Paternally expressed genes

Paternally expressed genes (*PEG*s) commonly have disrupted expression in interspecific hybrids (Kirkbride et al., 2015). These disruptions are due to methylation, *MEA*, and ploidy differences, and result in seed abortion. In *Arabidopsis,* crosses between a diploid female and a hexaploid male resulted in upregulation of *PEG*s and *PHE1* and aborted seed (Kirkbride et al., 2015). Upregulation is also found in interspecific hybrids. The same phenotype is observed between crosses of a hypomethylated parent and a normal diploid plant (Adams et al., 2000). Altering PEG expression has been shown to produce viable seed in interploidy crosses in *Arabidopsis* (Wolff et al., 2015). Mutations in genes *PEG2* and *PEG9* resulted in decreased gene expression and subsequent endosperm cellularization.

Seed size can be altered by *PEG*s due to increased ploidy in uneven ploidy crosses (FitzGerald et al., 2008). For example, a tetraploid male crossed with a diploid female will likely form larger seed than a tetraploid female crossed with a diploid male. This is, however, only if the ploidy differences are not too great in that seed formation does not occur.

Imprinting

Parents face a dilemma when determining the best way to maximize their offspring's survival: size of offspring versus number of offspring (Smith and Fretwell, 1974). It is an

evolutionary trade off determined in a tug of war between the male and female parents. The female parents tend to distribute resources evenly among offspring to increase the numbers of offspring (Haig, 2013). The male parents try to increase resources for the embryo and increase its size. One of the reasons the male parent would want a larger embryo size is because there is a possibility of other sibling seeds being fertilized by a different male plant (Haig and Westoby, 1989).

Imprinting is evolutionarily conserved, but the mechanisms and genes may differ. Yang et al. (2018) found that genes controlling imprinting are not well conserved between *A. thaliana* and monocots. Additionally, they found that a large percentage of imprinted genes in diploid or tetraploid wheat are also imprinted in hexaploid wheat. This is not surprising, as diploid and tetraploid wheat are ancestors to hexaploid wheat. These results are useful in understanding that imprinting, although it is widespread, it is not well conserved outside of related species.

KASP marker usage

Kompetitive allele-specific PCR (KASP) has allowed breeders to determine the genotypic diversity of plant germplasm quickly and affordably. These markers can be used for trait assessment ranging from disease resistance to improving eating and cooking quality (Meade et al., 2019; Yang et al., 2019). KASP markers determine if a single nucleotide polymorphism (SNP) and indels are present via fluorescence-based genotyping technology (Semagn et al., 2014). Other single-marker methods to visualize SNPs have also been used; cleaved amplified polymorphic sequences (CAPS), allele-specific PCR (AS-PCR), and re-sequencing (Rasheed et al., 2016). KASP markers are preferred over others due to their cost efficiency, high throughput, and reduce labor inputs. These markers require a forward primer containing the desired SNP, a

reverse common primer, isolated DNA, and KASP master mix. All of these components undergo multiple rounds of PCR to allow amplification of the desired sequence region, as well as the attachment of fluorescently labelled primers in the KASP master mix. These fluorescent primers allow sequences with the targeted SNP to be visualized. Although KASP markers are great for visualizing SNPs, there is potential for off-target amplification.

Amplicon sequencing

Amplicon sequencing allows researchers to sequence a small, known section of DNA approximately 200-500bp long. This type of sequencing has been implemented in polyploid species due to a lack of homoallele specificity when using markers (Balázs and Cowling, 2010). Homoallele specificity is crucial when isolating a specific mutation, such as a SNP, insertion, or deletion. Without this specificity, regions that do not include the desired mutation will be amplified. These off-target regions cloud marker results by indicating the mutation is present when it is not. Additionally, it is difficult to determine if a sample is homozygous or heterozygous for the desired mutation with these off-target regions. Amplicon sequencing removes this uncertainty by allowing researchers to directly compare the sequences to the reference genome and determine if the anticipated mutation is present. Amplicon sequencing, along with RNA-seq, has made locating mutations in complex polyploid genomes more accurate (Li et al., 2019).

Planting density

The wheat breeding program at Michigan State University has implemented a modified bulk method, Minibulk, that rapidly advances crosses from the F2 to F4 stage (Olson et al., 2021). The Minibulk process uses relatively high planting densities of 300 seeds at F2 and then 500 seeds at F3 in 8x8 inch square pots (Olson et al., 2021). Minimal research has been conducted for grain planting density in greenhouse settings. Cisar et al. (1982) examined planting densities of oats in greenhouse settings. The goal of the study was to determine the optimal planting density that would result in the greatest number of spikes with few or single seeds for the use of single seed descent (Cisar et al., 1982). Additionally, the oats were grown in stressed conditions without fertilizer to ensure few seeds will develop per spike (Cisar et al., 1982). This differs from the goals of the wheat breeding program at Michigan State University.

BIBLIOGRAPHY

BIBLIOGRAPHY

Adams S, Vinkenoog R, Spielman M, Dickinson HG, Scott RJ (2000) Parent-of-origin effects on seed development in Arabidopsis thaliana require DNA methylation. Development **127**: 2493–2502

Balázs E, Cowling WA (2010) Exploiting genome-wide association in oilseed Brassica species. Genome **53**: 853–855

Baroux C, Gagliardini V, Page DR, Grossniklaus U (2006) Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in Arabidopsis. Genes Dev **20**: 1081–1086

Cisar G, Howey AE, Brown CM (1982) Optimal Population Density and Random vs. Selective Elimination of Genotypes Under a Modified Single-Seed Descent Method with Spring Oats1. Crop Sci **22**: cropsci1982.0011183X002200030034x

Cox TS, Wu J, Wang S, Cai J, Zhong Q, Fu B (2017) Comparing two approaches for introgression of germplasm from Aegilops tauschii into common wheat. Crop J **5**: 355–362

Dvorak J, Luo M-C, Yang Z-L, Zhang H-B (1998) The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat. Theor Appl Genet **97**: 657–670

FAOSTAT (2020) Food and agriculture organization of the United Nations. Food Agric data. http://www.fao.org/faostat/en/#data

FitzGerald J, Luo M, Chaudhury A, Berger F (2008) DNA Methylation Causes Predominant Maternal Controls of Plant Embryo Growth. PLOS ONE **3**: e2298

Haig D (2013) Kin Conflict in Seed Development: An Interdependent but Fractious Collective. Annu Rev Cell Dev Biol **29**: 189–211

Haig D, Westoby M (1989) Parent-Specific Gene Expression and the Triploid Endosperm. Am Nat **134**: 147–155

Holec S, Berger F (2012) Polycomb Group Complexes Mediate Developmental Transitions in Plants. Plant Physiol **158**: 35

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum Aegilops complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci **99**: 8133

Jarillo JA, Piñeiro M, Cubas P, Martínez-Zapater JM, Jarillo JA, Piñeiro M, Cubas P, Martínez-Zapater JM (2009) Chromatin remodeling in plant development. Int J Dev Biol **53**: 1581–1596

Jullien PE, Mosquna A, Ingouff M, Sakata T, Ohad N, Berger F (2008) Retinoblastoma and Its Binding Partner MSI1 Control Imprinting in Arabidopsis. PLOS Biol **6**: e194

Kirkbride RC, Yu HH, Nah G, Zhang C, Shi X, Chen ZJ (2015) An Epigenetic Role for Disrupted Paternal Gene Expression in Postzygotic Seed Abortion in Arabidopsis Interspecific Hybrids. Mol Plant **8**: 1766–1775

Köhler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W (2003a) Arabidopsis MSI1 is a component of the *MEA*/*FIE* Polycomb group complex and required for seed development. EMBO J **22**: 4804–4814

Köhler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003b) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. Genes Dev **17**: 1540–1553

Köhler C, Makarevich G (2006) Epigenetic mechanisms governing seed development in plants. EMBO Rep **7**: 1223–1227

Kradolfer D, Hennig L, Köhler C (2013) Increased Maternal Genome Dosage Bypasses the Requirement of the FIS Polycomb Repressive Complex 2 in Arabidopsis Seed Development. PLOS Genet **9**: e1003163

Lagudah ES, Appels R, Brown AHD, McNeil D (1991) The molecular–genetic analysis of Triticum tauschii, the D-genome donor to hexaploid wheat. Genome **34**: 375–386

Lewis EB (2004) A Gene Complex Controlling Segmentation in Drosophila. *In* HD Lipshitz, ed, Genes Dev. Cancer Life Work Edw. B Lewis. Springer US, Boston, MA, pp 205–217

Li A, Liu D, Yang W, Kishii M, Mao L (2018) Synthetic Hexaploid Wheat: Yesterday, Today, and Tomorrow. Engineering **4**: 552–558

Li W, Riday H, Riehle C, Edwards A, Dinkins R (2019) Identification of Single Nucleotide Polymorphism in Red Clover (Trifolium pratense L.) Using Targeted Genomic Amplicon Sequencing and RNA-seq. Front Plant Sci **10**: 1257

Ling H-Q, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, et al (2013) Draft genome of the wheat A-genome progenitor Triticum urartu. Nature **496**: 87

Margueron R, Reinberg D (2011) The Polycomb complex PRC2 and its mark in life. Nature **469**: 343–349

Mcfadden ES, Sears ER (1946) The origin of triticum spelta and its free-threshing hexaploid relatives. J Hered **37**: 81–89

Meade F, Byrne S, Griffin D, Kennedy C, Mesiti F, Milbourne D (2019) Rapid Development of KASP Markers for Disease Resistance Genes Using Pooled Whole-Genome Resequencing. Potato Res. doi: 10.1007/s11540-019-09428-x

Miranda LM, Murphy JP, Marshall D, Leath S (2006) Pm34: a new powdery mildew resistance gene transferred from Aegilops tauschii Coss. to common wheat (Triticum aestivum L.). Theor Appl Genet **113**: 1497–1504

Mizuno N, Yamasaki M, Matsuoka Y, Kawahara T, Takumi S (2010) Population structure of wild wheat D-genome progenitor Aegilops tauschii Coss.: implications for intraspecific lineage diversification and evolution of common wheat. Mol Ecol **19**: 999–1013

Nallamilli BRR, Zhang J, Mujahid H, Malone BM, Bridges SM, Peng Z (2013) Polycomb group gene Os*FIE*2 regulates rice (Oryza sativa) seed development and grain filling via a mechanism distinct from Arabidopsis. PLoS Genet **9**: e1003322–e1003322

Olson EL, Noble A, Forsberg L, Brisco-McCann E (2021) The 'Minibulk' system for advancing winter cereal breeding programs. Manuscript in preparation.

Olson EL, Rouse MN, Pumphrey MO, Bowden RL, Gill BS, Poland JA (2013) Simultaneous transfer, introgression, and genomic localization of genes for resistance to stem rust race TTKSK (Ug99) from Aegilops tauschii to wheat. Theor Appl Genet **126**: 1179–1188

Rasheed A, Wen W, Gao F, Zhai S, Jin H, Liu J, Guo Q, Zhang Y, Dreisigacker S, Xia X, et al (2016) Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. Theor Appl Genet **129**: 1843–1860

Saluja M, Kaur S, Bansal U, Subhash Chander Bhardwaj, Chhuneja P (2018) Molecular mapping of linked leaf rust resistance and non-glaucousness gene introgressed from Aegilops tauschii Coss. in hexaploid wheat Triticum aestivum L. Plant Genet Resour **16**: 82–88

Schmidt A, Wöhrmann HJP, Raissig MT, Arand J, Gheyselinck J, Gagliardini V, Heichinger C, Walter J, Grossniklaus U (2013) The Polycomb group protein MEDEA and the DNA methyltransferase MET1 interact to repress autonomous endosperm development in Arabidopsis. Plant J **73**: 776–787

Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed **33**: 1–14

Shen X, Gmitter FG, Grosser JW (2011) Immature Embryo Rescue and Culture. *In* TA Thorpe, EC Yeung, eds, Plant Embryo Cult. Methods Protoc. Humana Press, Totowa, NJ, pp 75–92

Singh N, Wu S, Tiwari V, Sehgal S, Raupp J, Wilson D, Abbasov M, Gill B, Poland J (2019) Genomic Analysis Confirms Population Structure and Identifies Inter-Lineage Hybrids in Aegilops tauschii. Front Plant Sci **10**: 9

Smith CC, Fretwell SD (1974) The Optimal Balance between Size and Number of Offspring. Am Nat **108**: 499–506

Takumi S, Mizuno N, Okumura Y, Kawahara T, Matsuoka Y (2008) Two major lineages of Aegilops tauschii Coss. revealed by nuclear DNA variation analysis.

Velu G, Guzman C, Mondal S, Autrique JE, Huerta J, Singh RP (2016) Effect of drought and elevated temperature on grain zinc and iron concentrations in CIMMYT spring wheat. J Cereal Sci **69**: 182–186

Wang J, Luo M-C, Chen Z, You FM, Wei Y, Zheng Y, Dvorak J (2013) Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. New Phytol **198**: 925–937

Weinhofer I, Hehenberger E, Roszak P, Hennig L, Köhler C (2010) H3K27me3 Profiling of the Endosperm Implies Exclusion of Polycomb Group Protein Targeting by DNA Methylation. PLoS Genet **6**: e1001152

Wiersma AT, Pulman JA, Brown LK, Cowger C, Olson EL (2017) Identification of Pm58 from Aegilops tauschii. Theor Appl Genet **130**: 1123–1133

Wolff P, Jiang H, Wang G, Santos-González J, Köhler C (2015) Paternally expressed imprinted genes establish postzygotic hybridization barriers in Arabidopsis thaliana. eLife **4**: e10074

Yadegari R, Kinoshita T, Lotan O, Cohen G, Katz A, Choi Y, Katz A, Nakashima K, Harada JJ, Goldberg RB, et al (2000) Mutations in the *FIE* and *MEA* Genes That Encode Interacting Polycomb Proteins Cause Parent-of-Origin Effects on Seed Development by Distinct Mechanisms. Plant Cell **12**: 2367

Yang G, Chen S, Chen L, Gao W, Huang Y, Huang C, Zhou D, Wang J, Liu Y, Huang M, et al (2019) Development and utilization of functional KASP markers to improve rice eating and cooking quality through MAS breeding. Euphytica **215**: 66

Yang G, Liu Z, Gao L, Yu K, Feng M, Yao Y, Peng H, Hu Z, Sun Q, Ni Z, et al (2018) Genomic Imprinting Was Evolutionarily Conserved during Wheat Polyploidization. Plant Cell **30**: 37–47

Zhang W, Chao S, Manthey F, Chicaiza O, Brevis JC, Echenique V, Dubcovsky J (2008) QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. Theor Appl Genet **117**: 1361–1377

CHAPTER II

DEVELOPMENTAL CHANGES IN POLYCOMB REPRESSIVE COMPLEX 2 (PRC2) MUTANTS

Abstract

Hexaploid wheat (*Triticum aestivum*) has undergone two hybridization events resulting in severe genetic bottlenecks, particularly in the D genome contributed by the diploid progenitor, *Aegilops tauschii*. One method to regain this genetic diversity is through interspecific hybridizations with diploid progenitors, which requires embryo rescue to recover progeny due to endosperm failure in seed development. Previous studies in *Arabidopsis* and rice have demonstrated that mutations in imprinted Polycomb Repressive Complex 2 (PRC2) genes *medea* (*MEA*), *fertilization independent endosperm* (*FIE*), and *fertilization independent seed 2* (*FIS2*) can, or potentially can, restore endosperm function to interploidy hybrids. Homologs of *MEA*, *FIE,* and *FIS2* were identified in the wheat A, B and D genomes. Pyramids of double and triple mutants of these genes were developed using a combination of KASP markers and amplicon sequencing. Single, double, and triple mutants at *MEA*, *FIE*, and *FIS2* were hybridized with *Ae. tauschii* to identify mutant combinations that may restore endosperm function.

Introduction

Wheat (*Triticum aestivum*, 2n=6=42, AABBDD) is a food source that billions of people rely on for calories and nutrition. Wheat accounts for 20% of protein and 28% of calories consumed in people's diets (Velu et al., 2016). In 2019, 35,468 hg/ha of wheat were harvested from 2.159e8 hectares worldwide (FAOSTAT 2020). As the second most produced cereal in the world, wheat is a major component in the agricultural market (Velu et al., 2016).

Wheat is an allohexaploid consisting of A, B, and D genomes. The A genome is most closely related to *Triticum urartu* (2n = 2x = 14), the B genome is a relative of *Aegilops*

speltoides (2n = 14, SS), and the D genome is *Aegilops tauschii* (2n=2x=14, DD); these genomes were merged over two separate hybridization events (Ling et al., 2013). The first hybridization event occurred 300,000-500,000 years ago between the A and B genome progenitors, which resulted in *Triticum turgidum* (AABB) (Huang et al., 2002). The second hybridization occurred between *Triticum turgidum* (AABB) and *Ae. tauschii* (DD) to form modern hexaploid wheat, *T. aestivum* (AABBDD) (Mcfadden and Sears, 1946). Only a few *Ae. tauschii* plants hybridized with *Triticum turgidum*, resulting in a narrow genetic background of the wheat D genome in hexaploid wheat (Lagudah et al., 1991). In an effort to increase genetic diversity, *Ae. tauschii* has been used as a novel source of genetic diversity in hexaploid wheat for insect, disease, and abiotic stress resistance (Lagudah et al., 1991). Genes conferring disease or insect resistance to leaf rust, stem rust, yellow rust, Hessian fly, greenbug, cereal leaf beetle, root lesion nematode, karnal blunt, soil borne mosaic virus, powdery mildew, and others have been introgressed into hexaploid wheat from *Ae. tauschii* (Miranda et al., 2006; Olson et al., 2013; Cox et al., 2017; Wiersma et al., 2017; Saluja et al., 2018). Finding novel genes that are well adapted to undesirable conditions, drought, disease, heat, etc., is an important aspect of wheat breeding.

Two primary methods have been used to introgress *Ae. tauschii* into hexaploid wheat: synthetic hexaploid wheat production and direct hybridization. Synthetic hexaploid wheat is created by hybridizing *Triticum turgidum* with *Ae. tauschii* and then undergoing spontaneous or induced chromosome doubling (Cox et al., 2017). Synthetic hexaploid wheat lines are often spring type and non-free threshing (Cox et al., 2017). Additionally, changes can occur in the genome structure and/or copy numbers of genes in synthetic hexaploids (Li et al., 2018).

Direct hybridization is possible between hexaploid wheat as the maternal parent and *Ae. tauschii* as the paternal parent. However, only the embryo develops in these hybrids, not endosperm. Thus, the embryo has to be rescued and subsequent tissue culture is necessary to generate viable offspring (Shen et al., 2011). The necessity of embryo rescue is potentially due to the imbalance of maternal to paternal dosages in double fertilization. Two male gametes are necessary in double fertilization: one to combine with the female egg cell, and the other to combine with two polar nuclei (Faure J.E., 2011). The latter forms the endosperm portion of the seed. The resulting combination of two polar nuclei with the one sperm cell is a triploid (3n) endosperm. This 2:1 ratio of maternal and paternal genetic material is required for proper formation of endosperm. In the direct hybridization procedure, the endosperm ratio is 6:1 maternal to paternal dosage gene dosage, which may cause the failure in endosperm development.

A primary component of endosperm development is the Polycomb Repressive Complex 2 (PRC2), one of several polycomb group complexes (Holec and Berger, 2012). These complexes were first discovered through research in segmentation of Drosophila (Lewis, 2004). PRC2 catalyzes trimethylation of histone H3 on lysine 27 (H3K27me3) allowing epigenetic activity to occur in select vegetative and reproductive development (Weinhofer et al., 2010). PRC2 consists of several genes, three of which are *fertilization independent endosperm* (*FIE*), *fertilization independent seed* (*FIS2*), and *medea* (*MEA*) (Köhler et al., 2003). These genes are maternally expressed and are present among angiosperms including rice and wheat (Nallamilli et al., 2013).

Several studies have researched the actions of PRC2 genes. Most notably, Kradolfer et al. (2013) found that increasing maternal dosage of mutant *MEA* or *FIS2* in *Arabidopsis* can bypass endosperm failure in maternal tetraploid and paternal diploid crosses. These crosses produced

viable triploid seeds (Kradolfer et al., 2013). The maternal excess of mutant *MEA* or *FIS2* formed endosperm at a slower pace of 10 days after pollination instead of the usual 6 days after pollination (Kradolfer et al., 2013).

Other research has focused on the specific mechanisms of *MEA*, *FIE*, and *FIS2*. *MEA* is silenced by H3K27 trimethylation from PRC2 and *FIS2* is silenced through DNA methylation or, more specifically, methyltransferase 1 (MET1) in *Arabidopsis* (Jullien et al., 2008). *MEA* interacts with DNA methyltransferase (MET1) to repress autonomous endosperm in *Arabidopsis* (Schmidt et al., 2013). Since MET1 is not orthologous to wheat, altering the function of *MEA* alone may impose autonomous endosperm formation in wheat. *FIE* proteins are in a WD repeat protein family (Nallamilli et al., 2013). The WD repeats form a beta propeller structure, a circular structure that creates surface loops used as the basis of protein complex formation (Yadegari et al., 2000). This family promotes protein-protein interactions. *FIE* functions in seed development, grain filling, and may also function in male gametophyte development (Nallamilli et al., 2013). In rice, the *FIE* protein contains a retinoblastoma protein region (Nallamilli et al., 2013).

Based on these previous studies, it is expected that having one or several mutant copies of *MEA*, *FIS2*, or *FIE* genes in maternal hexaploid wheat may allow proper endosperm formation, and therefore viable seed, to occur in interploidy hybridizations of hexaploid wheat and *Ae. tauschii*. Since wheat is an allohexaploid comprised of three homoeologous genomes (AABBDD), three hypotheses can be tested.

1. One mutant homeoallele of *MEA*, *FIS2*, or *FIE* will restore viable endosperm formation in $6n_{mat}X\ 2n_{pat}$ hybridizations. One mutant homeoallele of *MEA*, *FIS2*, or *FIE* will lead the central cell of the maternal hexaploid wheat to function as 4n, instead of 6n. The 4:1

maternal to paternal endosperm ratio may be close enough to the desired 2:1 maternal to paternal endosperm ratio to allow endosperm to form.

2. Two mutant homeoalleles of *MEA*, *FIS2*, or *FIE* will restore viable endosperm formation in $6n_{mat}X\ 2n_{pat}$ hybridizations. Two mutant homeoalleles of *MEA*, *FIS2*, or *FIE* will lead the central cell of the maternal hexaploid wheat to function as 2n, instead of 6n restoring the appropriate 2:1 maternal to paternal endosperm ratio.
3. Three mutant homeoalleles of *MEA*, *FIS2*, or *FIE* will restore viable endosperm formation in $6n_{mat}X\ 2n_{pat}$ hybridizations. Three mutant versions of *MEA*, *FIS2*, or *FIE* will result in autonomous endosperm formation.

Components of each hypothesis are still ongoing; however, a pipeline has been established to readily identify one, two, and three mutant homeoalleles of *MEA*, *FIS2*, or *FIE*. In other words, single, double, and triple homozygous and heterozygous mutants. The results of this research will facilitate more efficient direct hybridizations between any hexaploid wheat and *Ae. tauschii*.

Materials and methods

Identification of Arabidopsis PRC2 orthologs in wheat and rice

Arabidopsis PRC2 proteins *MEA*, *FIE*, and *FIS2* were identified via arabidopsis.org. *Arabidopsis* PRC2 proteins identified are AT3G20740.1 for *FIE*, AT2G35670.1 for *FIS2*, and AT1G02580.1 for *MEA*. These proteins were then used in TBLASTn searches in the wheat genome (Gertz et al., 2006). Orthologous PRC2 genes identified in wheat were then used to search the EMS database John Innes Center. TBLASTn searches were also used to identify

orthologous PRC2 genes in rice. Rice PRC2 accessions identified and used for phylogenetic trees are AY456262.1 for *FIE,* AY321106.1 for *FIS2*, and AJ421722.3 for *MEA.*

Plant materials

Germplasm was obtained from the John Innes Center http://www.wheat-tilling.com/. These Cadenza-derived lines were treated with ethyl methanesulfonate (EMS) to create random mutations in the genome. Several Cadenza lines were identified that contained a SNP in *MEA*, *FIS2*, or *FIE* homeoalleles. A total of 9 lines were used in crosses to combine the mutations of different homeoalleles (Table 2.1). These lines contain a transcript altering mutation of either a non-sense SNP or missense SNP in the targeted homeoallele (Table 2.1). Missense mutations are also paired with a Sorting Intolerant From Tolerant (SIFT) score, or a prediction of the amino acid change impacting protein function (Ng and Henikoff, 2003). All Cadenza lines with missense mutations have a sift score less than 0.05, which is categorized as a deleterious substitution (Ng and Henikoff, 2003). Transcription information was gathered for all homeoalleles of *MEA*, *FIS2*, and *FIE* (Table 2.2). All homeoalleles except fie-B has expression in the endosperm according to expression atlas.

Plant growth conditions

Cadenza lines were germinated at 4.4C for five days on 200 M weight White Verigood Blotter paper (Anchor Paper Company, St. Paul, Minnesota) saturated with 1% Thiram (T24201, Sigma Aldrich, St. Louis, MO) and DI water solution. Seeds were transplanted to Gro Pro Premium Black Square pots (6 in. x 6 in. x 8 in.) with saturated Sure Mix potting media

(Michigan Grower Products Inc., Galesburg, MI). A subdue application was added after transplanting. *Ae. tauschii* seeds were germinated in clear plastic boxes with paper saturated with 1% Thiram solution for 6-10 weeks. 1% Thiram was periodically applied throughout vernalization. After vernalization, *Ae. tauschii* seedlings were transplanted the similarly to Cadenza lines. All germplasm was grown in greenhouse conditions with fertilizer (20-20-20 NPK) used weekly (approximately 188g / 5-gal H2O at 1:50 dilution). Fertilization may have occurred more often as needed.

Homeoallele	Zygosity	SNP type	sift scores	Line Name	Previous ensemble gene name	Previous ensembl cDNA position	Current ensembl gene name	Current cDNA position
fis2-A	Heterozygous	AA change	0.02	Cadenza 1069	Traes_5AL_0C4598F51.1	1387	TraesCS5A02G179600	1643
fis2-B	Heterozygous	Stop codon	NA	Cadenza 1722	Traes_5BL_96AAD5F61.1	286	TraesCS5B02G177400	2265
fis2-D	Heterozygous	Stop codon	NA	Cadenza 1699	Traes_5DL_600B6FCDC.1	974	TraesCS5D02G184200	849
mea-A	Homozygous	Stop codon	NA	Cadenza 0092	Traes_7AS_D9C6AC2A0.2	785	TraesCS7A02G128600	1087
mea-B	Heterozygous	AA change	0	Cadenza 1722	Traes_7BS_C79EBD23E.1	175	TraesCS7B02G028500	1963
mea-D	Heterozygous	AA change	0.02	Cadenza 1182	Traes_7DS_23A2E8EBB.1	225	TraesCS7D02G127400	2291
fie-A	Heterozygous	Stop codon	NA	Cadenza 1611	Traes_7AL_7CC969251.1	811	TraesCS7A02G308300	1562
fie-B	Homozygous	Stop codon	NA	Cadenza 1379	Traes_7BL_B7F78475F.1	487	TraesCS7B02G377900LC	650
fie-D	Homozygous	Stop codon	NA	Cadenza 1456	Traes_7DL_DF97791D5.1	484	TraesCS7D02G305100	938

Table 2.1: SNP, zygosity, SNP type, sift score, line names, gene names, and cDNA location of SNP for mutant donor parent.

Homeoallele	Protein	Exons	Coding exons	Domains and features	Transcript length (base pairs)	Translation length (amino acids)	Transcripts	Paralogs
fis2-A	TraesCS5A02G179600.1	22	21	7	2471	638	1	3
fis2-B	TraesCS5B02G177400.1	21	20	7	2620	619	4	3
	TraesCS5B02G177400.2	23	22	7	2734	657	4	3
	TraesCS5B02G177400.3	22	21	7	2674	637	4	3
	TraesCS5B02G177400.4	20	20	7	2233	618	4	3
fis2-D	TraesCS5D02G184200.1	15	14	3	1487	393	2	2
	TraesCS5D02G184200.2	22	21	7	2521	638	2	2
mea-A	TraesCS7A02G128600.1	16	16	21	2821	801	2	33
	TraesCS7A02G128600.2	10	10	6	1548	515	2	33
mea-B	TraesCS7B02G028500.1	15	15	21	2595	792	4	33
	TraesCS7B02G028500.2	16	16	21	2625	802	4	33
	TraesCS7B02G028500.3	11	11	8	1569	522	4	33
	TraesCS7B02G028500.4	11	11	8	1611	536	4	33
mea-D	TraesCS7D02G127400.1	16	16	18	3005	802	0	0
fie-A	TraesCS7A02G308300.1	14	14	23	2092	473	0	0
fie-B	TraesCS7B02G377900LC.1	13	13	NA	1470	389	0	0
fie-D	TraesCS7D02G305100.1	14	14	23	1841	485	0	0

Table 2.2: Gene components and properties of mutant donor parents.

Pyramiding PRC2 mutations

There were four cycles in the crossing design (Figure 2.1). Primary mutants were used in pyramiding crosses. All crosses were made using the approach method (Curtis and Croy, 1958).

Cycle 1:

Cadenza mutant lines were genotyped using KASP markers. The first cycle consisted of crossing a Cadenza mutant of one homeolog with a Cadenza mutant for a different homeolog (Figure 2.1).

Cycle 2:

The Cycle 1 progeny were Cadenza double mutant lines that have two heterozygous mutations in two homeologs (Figure 2.1). The progeny was expected to follow a 100% heterozygosity rate for the two mutant homeologs. Progeny from Cycle 1 were genotyped using KASP markers to confirm heterozygosity for two mutations. At the end of the Cycle 2, double mutant lines with complementary mutations were crossed to generate Cycle 3.

Cycle 3:

Progeny from Cycle 2 were genotyped via amplicon sequencing. *MEA* and *FIE* homeologs were also genotyped using KASP markers. Genotyping identified lines heterozygous for either three heterozygous mutations in three homeologs (AaBbDd), or two heterozygous mutations in two homeologs and a homozygous mutation in one homeolog (AaBbdd, aaBbDd, or AabbDd) (Figure 2.1). AaBbDd genotypes were expected to occur at 12.5% for, and AaBbdd, aaBbDd, or AabbDd genotypes at 6.25% These target genotypes were selected and self-pollinated to generate Cycle 4.

Cycle 4:

Progeny from Cycle 3 were genotyped via amplicon sequencing. Genotyping screened for homozygous Cadenza triple mutant lines (aabbdd) and homozygous Cadenza double mutant lines (aaB_dd, A_bbdd, and aabbD_). Cycle 3 progeny lines were expected to occur at 6.25% or 3.125% homozygous triple mutants (aabbdd).

Genotyping

KASP markers were obtained through the John Innes Center (Table 2.3). The location of the SNP for Cadenza1379 mutant fie-B was not optimal for creating a KASP marker. KASP genotyping was conducted at the beginning of each cycle, prior to any crossing or self-pollination.

Amplicon sequencing

Amplicon sequencing was conducted at the beginning of cycle 3 and cycle 4. Mutation donor lines and wild type Cadenza were used as controls. Primers were created with Primer1 for each SNP and included a Fluidigm CS1/CS2 universal oligomers at their 5' end. Product sizes ranged from 200bp to 300bp in length (Table 2.4). Samples were amplified using PCR protocol: 95°C (5 min), followed by 35 cycles of 95°C (45 sec) 60°C (45 sec), 72°C (90 sec), and elongation at 72°C (10 min). Size selection and primer dimer removal was conducted with Omega Bio-Tek's Mag-Bind TotalPure NGS. Samples were then quantified, normalized, and sequenced on an Illumina MiSeq v2 in 2x250 bp paired ends. Cycle 3 samples had an average

number of read pairs per sample of approximately 58,020 and ranged from 10,858 to 98,574 reads. Cycle 4 had an average number of read pairs per sample of approximately 10,200 with 96% of samples falling within 6,750 to 13,850 read pairs.

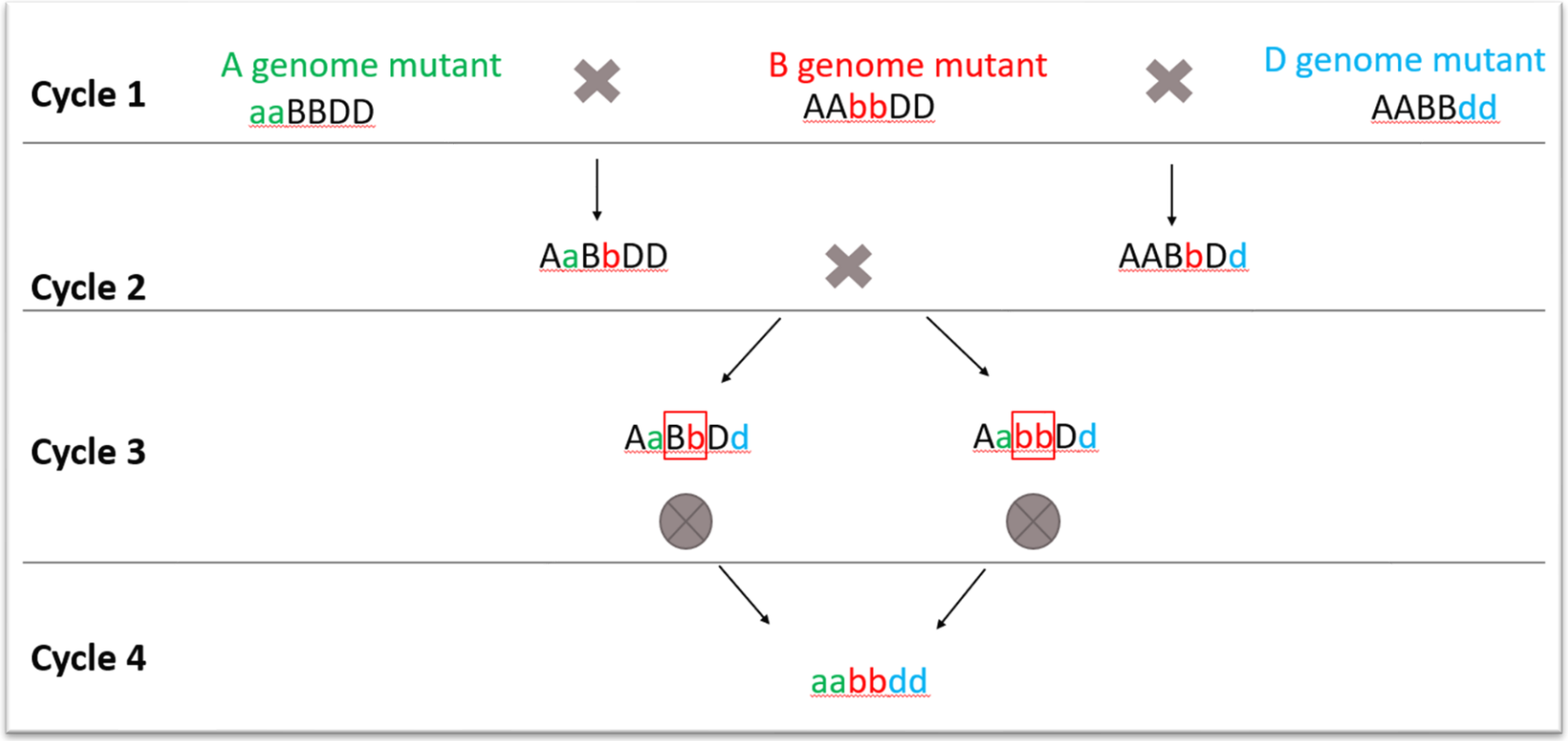


Figure 2.1: Pyramid diagram for Cycles 1-4.

Homeoallele	Wild Type Primer	Alternative primer	Common Primer
fis2-A	ACCATGTAATGATTGAGCAGGATC	ACCATGTAATGATTGAGCAGGATT	ATTTCCAGATATTTCTGCACCAC
fis2-B	GCGTTTCCTTATGATTAAACTCTGG	GCGTTTCCTTATGATTAAACTCTGA	GCCGCAGACATTCATGGT
fis2-D	ACCACGATTTTCCCCAACTTG	ACCACGATTTTCCCCAACTTA	CCATCACATTTGGTTCCCTTTTTAG
mea-A	TGCAAACTACATGGGTGTTCTC	TGCAAACTACATGGGTGTTCTT	GGTCCATCATCCATGCTGTC
mea-B	GCTGCTTCCCTATGAGAGATAAG	GCTGCTTCCCTATGAGAGATAAA	TGTACCCCAGAATGATGTGGA
mea-D	AGGTTATGTTTGTAGCAGGCG	AGGTTATGTTTGTAGCAGGCA	TGCCTGCATTAATCCGCTCA
fie-A	GCGAAGGCAAAATCTATGTGTG	GCGAAGGCAAAATCTATGTGTA	CTAGCACCACCACCACATT
fie-B	NA	NA	NA
fie-D	GTTCGCTAACCTTTCATTGACC	GTTCGCTAACCTTTCATTGACT	CCACCCTTCTGATATCTACCG

Table 2.3: KASP primer information.

A systematic grep search around the targeted SNP was used as a general guide in selecting planting desired genotypes. Grep searching allowed a fast turnaround time between harvesting and planting. The Unix command zgrep -c was used to generate the number of reads containing desired SNP and the number of reads containing the wild type sequence in cycle 3 samples. Cycle 4 samples were first cleaned using Cutadapt to remove Fluidigm CS1/CS2 adapters. Reads were then paired using software PEAR before the Unix command zgrep -c was used to generate the number of reads containing the desired SNP and the number of reads containing the wild type sequence. Tables 2.5 and 2.6 list the sequences used in Cycle 3 and Cycle 4 grep search. Some Cycle 4 grep sequences differ from Cycle 3 grep sequences due to pooled samples causing false positive calls.

Phylogenetic trees

Phylogenetic trees were created in R studio using R package phylogram (Figure 2.3). The phylogenetic trees were originally generated using EMBl-EBI tool Clustal Omega and were recreated in R studio to generate higher quality images (Madeira et al., 2019).

Homeoallele	Primer set	Left Primer	Right primer	Total bp	bp from 5' to SNP	Wild SNP	Mutant SNP
fis2-A	1	CACACGGAATTTCAGCATTG	GTGGGTGGTGAAAGATTGCT	208	165	G	A
fis2-B	1	AGGAGCTGGCGTTTCCTTAT	GGATCACTGCTGCTTCAACA	281	33	G	A
fis2-D	1	AGGTCGCACCAAAGCTACAT	CACATGGTCCACAGAGCAGT	225	178	C	T
mea-A	1	CAGCATTGGATTCCTTCGAT	CTGTCCCAAGGTGATTGCTT	268	158	C	T
mea-B	1	TGGCAAGGATGACTGTCTTG	AAACCGAGCTGACAAGGAGA	245	38	C	T
mea-D	1	TGCTTTTGCATGAGCTGAGT	GTCAGGCGCGTAATGGTAAT	289	209	G	A
fie-A	1	ACATGCTCGTGGCTAATCGT	CACCACCACCACATTCAGAG	242	143	G	A
fie-A	2	AGTTTCTGAGGATGGCCGTA	CCGGGTAATTAGCACAGGAG	234	197	G	A
fie-B	1	CGGCACTAACTGTCCTTTCA	TCTCCACGTATGGCCAAAAT	223	106	G	A
fie-D	1	GCCAGTTGTGGCATGGATA	AGCACGATGCTGAGAGGATT	291	35	G	A

Table 2.4: Amplicon sequencing primers.

Cycle 3		
Homeoallele	wild	mutant
fis2-A	GATCCTGCTCAATCATTACATGGT	AATCCTGCTCAATCATTACATGGT
fis2-B	GCGTTTCCTTATGATTAAACTCTGG	GCGTTTCCTTATGATTAAACTCTGA
fis2-D	CAAGTTGGGGAAAATCGTGGT	TAAGTTGGGGAAAATCGTGGT
fie-A	GCGAAGGCAAAATCTATGTGTG	GCGAAGGCAAAATCTATGTGTA
fie-B	TGGATAATACTGTTAAAATCTG	TGGATAATACTGTTAAAATCTA
fie-D	GGTCAATGAAAGGTTAGCGAA	AGTCAATGAAAGGTTAGCGAA
mea-A	TGCAAACTACATGGGTGTTCTC	TGCAAACTACATGGGTGTTCTT
mea-B	CTTATCTCTCATAGGGAAGCAGC	TTTATCTCTCATAGGGAAGCAGC
mea-D	AGGTTATGTTTGTAGCAGGCG	AGGTTATGTTTGTAGCAGGCA

Table 2.5: Grep sequences used for read 1 Cycle 3 samples. Reverse complement of these sequences was used for read 2.

Cycle 4		
Homeoallele	wild	mutant
fis2-A	GATCCTGCTCAATCATTACATGGT	AATCCTGCTCAATCATTACATGGT
fis2-B	GCGTTTCCTTATGATTAAACTCTGG	GCGTTTCCTTATGATTAAACTCTGA
fis2-D	CAAGTTGGGGAAAATCGTGGT	TAAGTTGGGGAAAATCGTGGT
fie-A	GCGAAGGCAAAATCTATGTGTG	GCGAAGGCAAAATCTATGTGTA
fie-B	CATTCATACTTCCGTTATATCAGGACTTCCACCCTTCTGATATCTACCGAATTGCCAGTTGTGGCATGGATAATACTGTTAAAATCTGGTCAATGAAAGGTTAGCAAACCATTTCATCAATCTTGCTTCCTATTTACATTTTATTACTGCTGATCACTGACATGGTTTGTTAA	CATTCATACTTCCGTTATATCAGGACTTCCACCCTTCTGATATCTACCGAATTGCCAGTTGTGGCATGGATAATACTGTTAAAATCTAGTCAATGAAAGGTTAGCAAACCATTTCATCAATCTTGCTTCCTATTTACATTTTATTACTGCTGATCACTGACATGGTTTGTTAA
fie-D	CGGCACTAACTGTCCTTTCATTCATACTTCTGTTATATCAGGACTTCCACCCTTCTGATATCTACCGAATTGCCAGTTGTGGCATGGATAATACTGTTAAAATCTGGTCAATGAAAGGTTAGCGAACCATTTCATCAATCTTGCTTCCTATTTACATTT	CGGCACTAACTGTCCTTTCATTCATACTTCTGTTATATCAGGACTTCCACCCTTCTGATATCTACCGAATTGCCAGTTGTGGCATGGATAATACTGTTAAAATCTAGTCAATGAAAGGTTAGCGAACCATTTCATCAATCTTGCTTCCTATTTACATTT
mea-A	TGCAAACTACATGGGTGTTCTC	TGCAAACTACATGGGTGTTCTT
mea-B	CTTATCTCTCATAGGGAAGCAGC	TTTATCTCTCATAGGGAAGCAGC
mea-D	AGGTTATGTTTGTAGCAGGCG	AGGTTATGTTTGTAGCAGGCA

Table 2.6: Grep sequences used for Cycle 4 samples.

Variant calling

The primary method used to confirm genotypes was a BWA/bcftools pipeline (Yao et al., 2020; Danecek et al., 2021). Reads were first cleaned using cutadapt (version 2.4.Py3) with parameters -a ACACTGACGACATGGTTCTACA, -A TACGGTAGCAGAGACTTGGTCT, and -m 1. The reference genome was comprised of expected amplicon products (Table 2.7). The cleaned reads were aligned to the amplicon product reference genome using bwa (version 0.7.16a) mem to generate SAM output files. The SAM files were converted to BAM format via samtools (version 1.11) view with parameters -S and -b. The bam files were then sorted via samtools sort. The sorted files were used to generate a BCF file of genotype likelihoods using bcftools (version 1.11) mpileup with parameters -d 10000000 and -a FORMAT/AD. Variants were then called using bcftools call with parameters -Oz and -m. Bcftools view with parameter -Oz was used to convert the files to VCF format. VCF files were compressed using bcftools index. Finally, VCF files from all samples were merged into a single VCF file using bcftools merge with parameter -Ov.

Cadenza mutants and Ae. tauschii hybridizations

Hybridizations between Cadenza mutants and *Ae. tauschii* was done using the approach method (Curtis and Croy, 1958). The maternal parents, Cadenza mutants, were emasculated and covered for 3-6 days prior to hybridization. Afterwards, the emasculated heads were arranged to be slightly below *Ae. tauschii* spikes. *Ae. tauschii* spikes were either currently releasing viable pollen or about to release pollen. Dialysis tubing was hydrated and dried to a hollow tube shape. The tube contained the emasculated Cadenza heads and *Ae. tauschii* heads for each hybridization

and prevented cross contamination while also allowing light to pass through. Each Cadenza mutant line, and wild type Cadenza, had 3-10 spikes, or replicates, hybridized with *Ae. tauschii*. The exception was wild type Cadenza which had 15 spikes hybridized with *Ae. tauschii*. Seeds per spike were calculated within each line. Lines where then grouped into genotypes within *MEA*, *FIE*, or *FIS2*.

Homeoallele	Primer set	Product (Wild)
fis2-A	1	CACACGGAATTTCAGCATTGAGAAGTTCTGCATGCTGTTTTGAATATAACACACCTGAAACA TTTTGATCCTTACCTTAGGTGAAGATTTATTTCTTACACGTGCTCATATCCCATTTCCGTTGT CTAAAATTTCCAGATATTTCTGCACCACGTGCTTCTGTTGATCCTGCTCAATCATTACATGGT AGCAATCTTTCACCACCCAC
fis2-B	1	AGGAGCTGGCGTTTCCTTATGATTAAACTCTGGAACCACAGTCTATTAGATGCCCGCGCCAT GAATGTCTGCGGCACAATTCTTCAAGGCTACCAAAATGAAAGCTCGGACCCCAAGAAAATG TGAGTCGAGCATAGTGCGCTCATTATAATTTAAATCAAGTGGGTGTTTGGAGGAGGCCTTAG GAGCAGGCTAAAGAGAGGTGAAAGCTGCACCTGAGGCCATGGAGGACTATTCTCGATTCTA TTTATCGATCGGTTGTGTTGAAGCAGCAGTGATCC
fis2-D	1	AGGTCGCACCAAAGCTACATTCATAATTCCAGACATCAAGAATTTATCAACCTCCCGAGCTT GCAACCTTAACATTATCCTTATCAGCTGCGGTATGATGTGCTTCATATGTGATGTTTATGGTT TTACATTTCTTCACCACCATCACATTTGGTTCCCTTTTTAGTTTCGGAAGGGCAAGTTGGGGA AAATCGTGGTGAACATAACTGCTCTGTGGACCATGTG
mea-A	1	CAGCATTGGATTCCTTCGATAATCTCTTTTGTCGGCGATGCCTAGTATGCTCATCTATCAATC ATCATGTGATATATGTAATGCCTCTTGTTGCTGATGAATATTCTTCTCATCGTGACATTGGCA GGTTTTTTATTGCAAACTACATGGGTGTTCTCAAGATTTAATATTTCCTGTAAGTTTTGTTCC AGAATTATTTCCAAAATTACGTGTTTTTTCTTATTTGTGCCTTCTTCATACAGGCAGAAAAGC AATCACCTTGGGACAG
mea-B	1	TGGCAAGGATGACTGTCTTGGGGAGTACACTGGGGAGCTTATCTCTCATAGGGAAGCAGCC AAGCGTGGACAGAGATATGACCGTGAGAATTCATCATTCCTTTTCAACTTAAATACTGAGGT ACGTGTGCAGAAATATACCAGTTACAGTATGCAAACATTCACGGTTACACGCGGGACTGAT CACCAGATTTAGCTATAGTGCTTTTGCATGAGCTGAGTTATTCTCCTTGTCAGCTCGGTTT
mea-D	1	TGCTTTTGCATGAGCTGAGTTATTCTCCTTCTCAGCTCGGTTTATTCAAGAACCTGTATTTTA TATCTGATCCAGACGCTAGCTGTTGCTTATGCATACTCTTTGGACGCAGTTTGTTCTCGATGC CTTCAGGATGGGCAACAAGCTGAAGTTTGCCAACCATTCCCCTGATCCGAATTGCTATGCCA AGGTTATGTTTGTAGCAGGCAACCATAGGGTGGGCATATTCGCCAATGAGCGGATTAATGC AGGCAAGGAGATTTTTTATGATTACCATTACGCGCCTGAC

Table 2.7: Amplicon sequencing primer products used as reference genome in variant calling pipeline.

Table 2.7 (cont'd)

Homeoallele	Primer set	Product (Wild)
fie-A	1	ACATGCTCGTGGCTAATCGTTTGCTCGTAAACGAACTCATCTACATGCACTTGATCAGGGTT AGATGAAATTCAAAAAGTATAACTCTTGTGGTTAATCGTTTGCTTGTAAACAGGCAACCGCG AAGGCAAAATCTATGTGTGGGAAGTGCAGGCGAGCCCTCCTGTGCTAATTACCCGGTAAGT TTAACCGAGCAAGTTCAGCTATTCCCAATTTCTCCTCCTCTGAATGTGGTGGTGGTG
fie-A	2	AGTTTCTGAGGATGGCCGTACTTATACACGTGAAAAAACTGATATTTGTAGAAAACATGCTC GTGGCTAATCGTTTGCTCGTAAACGAACTCATCTACATGCACTTGATCAGGGTTAGATGAAA TTCAAAAAGTATAACTCTTGTGGTTAATCGTTTGCTTGTAAACAGGCAACCGCGAAGGCAA AATCTATGTGTGGGAAGTGCAGGCGAGCCCTCCTGTGCTAATTACCCGG
fie-B	1	CGGCACTAACTGTCCTTTCATTCATACTTCCGTTATATCAGGACTTCCACCCTTCTGATATCT ACCGAATTGCCAGTTGTGGCATGGATAATACTGTTAAAATCTGGTCAATGAAAGGTTAGCA AACCATTTCATCAATCTTGCTTCCTATTTACATTTTATTACTGCTGATCACTGACATGGTTTG TTAATCTTGTCATAGAATTTTGGCCATACGTGGAGA
fie-D	1	GCCAGTTGTGGCATGGATAATACTGTTAAAATCTGGTCAATGAAAGGTTAGCGAACCATTTC ATCAATCTTGCTTCCTATTTACATTTTATTACTGCTGATCACTGACATGGTTTGCTAACGTTG TCATAGAATTTTGGCCATACGTGGAGAAATCCTTTACATGGACTGACCTTCCATCAAAATTT CCAACAAAATTTGTCCAATTTCCGGTATGTTAAGTATCTCTCATCTTTGAACCTGTTATTATG CAAACTTTTAGTATTGTTCTTAATCCTCTCAGCATCGTGCT

Results

Cycle 1, 2 and 3 genotyping with KASP markers

KASP markers were used to identify homozygous mutant, heterozygous mutant, and wild type samples for each homeoallele of *MEA*, *FIE* and, *FIS2* in Cycles 1, 2, and 3. Cycle 1 genotyping was used to either confirm homozygosity or select for homozygous individuals in heterozygous mutant donor parents. Cycle 2 genotyping identified F_1 intercrosses that were heterozygous for two mutations (Table 2.8).

Gene	Cross	Offspring Genotype	Plants
MEA	AABBdd/AAbbDD	AABbDd	4
	aaBBDD/AABBdd	AaBBDd	2
		AaBB ?	1
	aaBBDD/AAbbDD	AaBbDD	2
FIE	AABBdd/aaBBDD	AaBBDd	4
	AAbbDD/AABBdd	AABbDd	2
		AABBDd	5
FIS2	AAbbDD/AABBdd	AABbDD	2
	aaBBDD/AABBdd	AABBDd	2
	AAbbDD/aaBBDD	AABbDD	2
	aaBBDD/AABBdd	AaBBDd	4
		? BBDd	2
		AABBDd	2

Table 2.8: Cycle 2 KASP genotyping results from selected offspring. ? indicates KASP results that were labeled “unknown”.

Cycle 3 KASP genotyping identified Cadenza triple mutant lines, or lines with either three heterozygous mutations in three homeologs (AaBbDd), or two heterozygous mutations in two homeologs and a homozygous mutation in one homeolog (AaBbdd, aaBbDd, or AabbDd) (Table 2.9).

Gene	Genotype	Plants
MEA	AaBbDd	3
	AabbDd	1
	AaBb*	1
*FIE***	Aa * *	16
FIS2	AaBbDd	4
	Aa*Dd	1

Table 2.9: Cycle 3 KASP genotyping results. Genotypes missing a subgenome had inconclusive KASP results. ** indicates KASP markers were not available for *FIE* B and D genome mutations.

Homeoallele specificity, or lack thereof, in KASP markers became apparent after genotyping Cycle 2 F1 lines. The mea-A, mea-B, mea-D, fie-A, and fis2-D KASP markers demonstrated clear segregation patterns between homozygous mutant, heterozygous mutant, and wild type samples (Figure 2.2).

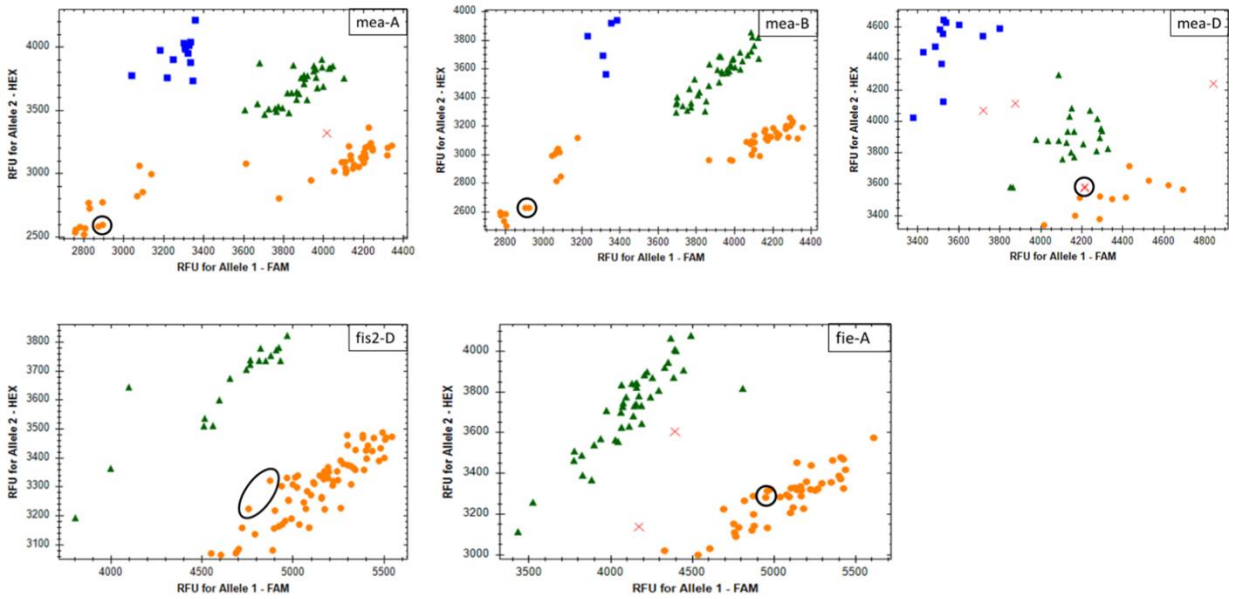


Figure 2.2: KASP segregation for mea-A, mea-B, mea-D, fie-A, and fis2-D in Cycle 3 genotyping. Blue squares represent homozygous mutant samples, green triangles represent heterozygous samples, orange circles represent wild type samples, and red X's represent unknown genotypes. Samples in black circles contain wild controls.

The fie-B, fie-D, fis2-A, and fis2-B KASP markers did not demonstrate clear and consistent segregation patterns between homozygous mutant, heterozygous mutant, and wild type samples. Wild controls were often indistinguishable from homozygous mutant and heterozygous mutant samples. A KASP marker was not available for fie-B. Attempts were made to create a homeoallele specific fie-B KASP maker without success which led to an amplicon sequencing approach for genotyping.

Amplicon sequencing Cycle 3

A subset of lines that were expected to be double or triple homozygous mutant from KASP were genotyped using amplicon sequencing. A total of 20 Cycle 3 lines were genotyped, along with the mutant donor parents and wild Cadenza. Cycle 2 progeny were expected to occur at 12.5% for heterozygous mutations in three homeologs (AaBbDd), and 6.25% for heterozygous mutations in two homeologs and a homozygous mutation in one homeolog (AaBbdd, aaBbDd, or AabbDd). Amplicon sequencing confirmed *MEA* genotypes and elucidated further information for *FIS2* and *FIE* samples (Table 2.10). The target genotype (AaBbDd, AaBbdd, aaBbDd, or AabbDd) could not be obtained for *FIS2*.

Genotype	*MEA*	*FIE*	*FIS2*
AaBbDd	3	1	0
AabbDd	1	0	0
AaBbdd	1	0	0
AaBbDD	0	7	0
AaBBDd	0	0	1
AaBBDD	0	7	1
AABBDD	0	0	3
total	5	15	5

Table 2.10: Cycle 3 amplicon sequencing genotypic distribution.

Amplicon sequencing Cycle 4

A total of 509 Cycle 4 lines were genotyped, along with the mutant donor parents and wild Cadenza (Table 2.11). Desired Cycle 3 progeny were expected to occur at 3.125% if the parental genotype was heterozygous for all three homeologs (AaBbDd) or 6.25% if the parental genotype had heterozygous mutations in two homeologs and a homozygous mutation in one homeolog (AaBbdd, aaBbDd, or AabbDd). *FIE* samples sequenced had parental genotypes of the former (AaBbDd) while *MEA* samples sequenced had parental genotypes of the latter (AaBbdd and AabbDd). Even though there were no homozygous triple mutant *FIE* lines, this was within the range of the expected distribution (P>0.05). *MEA* homozygous triple mutants occurred at a slightly lower rate than expected, but only slightly at P=0.047. (Table 2.12). Homozygous triple mutant *FIS2* lines were unable to be obtained.

	MEA	*FIE*	*FIS2*
aabbdd	2	0	0
AaBbDd	0	1	0
AabbDd	10	1	0
aaBbDd	0	5	0
aabbDd	9	0	0
aaBBdd	0	0	0
aaBBDd	0	0	1
AaBBDd	0	1	7
AaBBdd	1	0	0
AAbbdd	0	0	0
AABbDd	0	3	0
AABbdd	0	0	0
AAbbDd	6	0	0
aabbDD	16	0	0
AaBbDD	9	94	0
AabbDD	23	24	0
aaBbDD	5	46	0
AaBBDD	8	15	62
aaBBDD	3	6	17
AABbDD	7	58	0

Table 2.11: Cycle 4 genotyping results.

Table 2.11 (cont'd)

	MEA	*FIE*	*FIS2*
AAbbDD	10	18	0
AABBDd	0	0	3
AABBdd	0	0	0
wild	5	7	26
total	114	279	116

	Genotype	Observed occurrence	Expected occurrence	chi square value	P-value
MEA	aabbdd	2	7.125	3.93	0.047
FIE	aabbdd	0	1.656	1.709	0.191

Table 2.12: Cycle 4 chi square results.

Hybridizing single, double, and/or triple mutant homeoallele(s) to Ae. tauschii

In Cycle 1, single homeoallele mutant hybridizations yielded no viable endosperm for any homeoallele in any gene. After Cycle 4, single, double and/or triple mutant Cadenza lines were hybridized with *Ae. tauschii* (Table 2.13). Wild type endosperm development did not occur for any of these hybridizations, but caryopses were still able to be harvested. Due to lines having a variable number of spikes hybridized with *Ae. tauschii* (3-10 spikes per line), seeds per spike was calculated for each line (Table 2.13). Means were generated for all genotypes within each PRC2 gene (Table 2.14). No genotype had a higher seeds per spike count than the self-crossed wild Cadenza, however all averages were greater than the wild Cadenza hybridized with *Ae. tauschii*. *FIS2* genotypes in particular had more seeds per spike than *MEA* and *FIS2* genotypes. FIS2 genotype aaBBDd average seeds per spike was close to the self-crossed wild Cadenza seeds per spike (Table 2.14).

PRC2 Gene	Genotype	Seeds per Spike
MEA	AabbDd	1.3
MEA	aabb	3.333333
MEA	aabb	1.833333
MEA	aabb	1.166667
MEA	aabb	6.666667
MEA	aabbdd	5.666667
MEA	aabb	1.333333
MEA	bb	4.166667
MEA	aabb	1
MEA	aabb	0.666667
MEA	bb	0.666667
MEA	aabb	1.333333
MEA	aabb	0.166667
MEA	bb	1.666667
MEA	aa	1.5
MEA	aabb	4.5
MEA	aabb	1.714286
MEA	aabb	2
MEA	bb	0.166667
MEA	aa	4
MEA	aabbdd	3.166667
MEA	bb	0.857143
MEA	aabb	2.333333
FIE	bb	1.5
FIE	aa	0.833333
FIE	aaBb	1
FIE	bb	2.333333
FIE	aa	1.5
FIE	aa	1.5
FIE	aaBb	4
FIE	aaBb	1
FIE	aa	2.666667
FIE	aaBb	2.5
FIE	aaBb	0.5
FIE	aaBb	3.571429
FIE	bb	1.166667
FIE	bb	1.333333

Table 2.13: Caryopses yield from single, double, and/or triple mutant Cadenza lines hybridized with *Ae. tauschii*. Wild Cadenza x wild Cadenza and wild Cadenza x *Ae. tauschii* hybridizations were made as controls.

Table 2.13 (cont'd)

PRC2 Gene	Genotype	Seeds per Spike
FIS2	AaBBDd	2.833333
FIS2	aaBBDd	6.666667
FIS2	aa	2.75
FIS2	AaBBDd	8
FIS2	aa	3.166667
FIS2	aa	7.166667
FIS2	aa	5.333333
	Wild x Wild	8.666667
	Wild x *Ae. tauschii*	0.6

PRC2 Gene	Genotype	Seeds per spike
MEA	aaBBDD	2.7
	aabbDD	2.157
	aabbdd	4.416
	AabbDd	1.3
	AAbbDD	1.5
FIE	aaBBDD	1.625
	aaBbDD	2.095
	AAbbDD	1.58
FIS2	aaBBDD	4.604
	aaBBDd	6.66
	AaBBDd	5.416

Table 2.14: Average seeds per spike of the different genotypes of *MEA*, *FIE*, and *FIS2* resulting from hybridizations between *Ae. tauschii*.

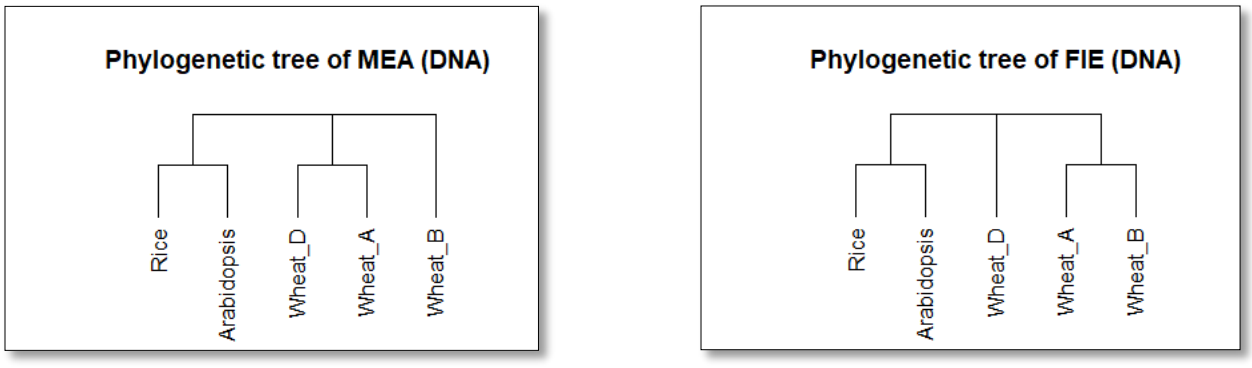
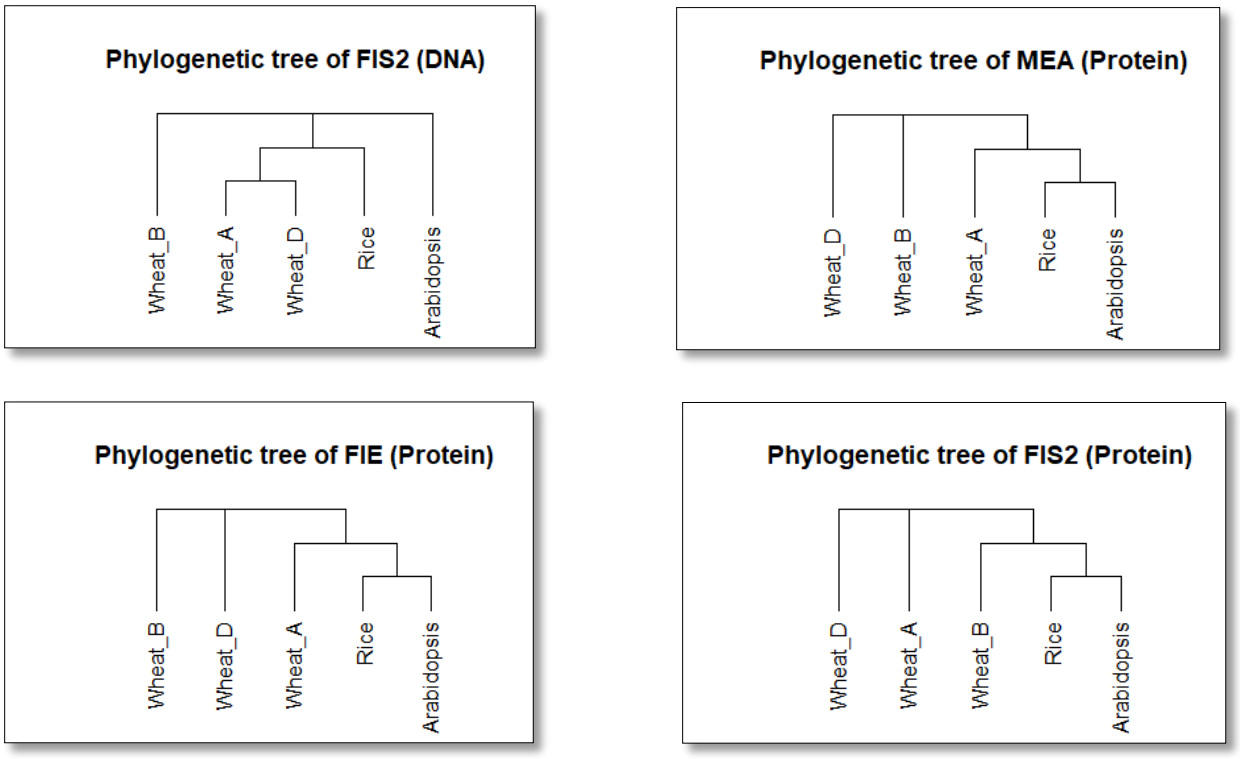


Figure 2.3: Phylogenetic trees of *MEA*, *FIE*, and *FIS2* genes.

Figure 2.3 (cont'd)



Discussion

KASP marker results can be split into two categories:

1. Clear and consistent segregation between wild, heterozygous, and homozygous mutant samples.
2. Little to no segregation between wild, heterozygous, and homozygous mutant samples.

Markers in the first category have homeoallele specificity. Homeoallele specificity is consistent amplification of wild type and mutant type sequences in the homeoallele that marker is targeting. Markers in the first category most likely do not have homeoallele specificity and have off-target amplification. Polyploid species are particularly susceptible to off-target amplification with SNP markers due to similar homoeologous or paralogous sequences in the genome (Edwards et al.,

2009). Similar issues arose when converting SNP chip arrays to KASP markers because of the presence of paralogs and homologs in polyploid crops like hexaploid wheat (Makhoul et al., 2020). Marker development and improvement was attempted for fie-B. A KASP marker was not available at John Innes Center for the targeted SNP of fie-B. Two attempts were made to create a KASP marker that was homeoallele-specific to the SNP of fie-B, but both attempts were unsuccessful.

Accurate genotyping is necessary to distinguish homozygous and heterozygous mutant homeoalleles in lines after pyramiding cycles. Therefore, amplicon sequencing was employed to overcome the off-target amplification in fie-B, fie-D, fis2-A, and fis2-B KASP markers. *MEA* KASP results had consistent homeoallele specificity for all 3 homeoalleles. Due to this consistency, mea-A, mea-B, and mea-D were used to compare the effectiveness of amplicon sequencing in Cycle 3. Genotyping results were similar between KASP markers and the amplicon sequencing pipeline for *MEA* homeoalleles.

The Unix command grep was used as an approximate guide for determining which lines would be planted in Cycle 4, and which lines would be hybridized with *Ae. tauschii* after Cycle 4. These grep searches contained the desired homeoallele SNP or wild type sequence. Grep searches closely matched the KASP and snpcalling pipeline genotyping results for *MEA* samples in Cycle 3. Unfortunately, the grep method had less accurate results for all other PRC2 genes in Cycle 3, and all PRC2 genes in Cycle 4. This, combined with inaccurate KASP markers, resulted in several lines being incorrectly labeled double and triple mutants. Most notably, the parent for fis2-B was mis-genotyped as homozygous mutant for fis2-B. The parental line Cadenza1722 is heterozygous for the desired SNP of homeoallele fis2-B, and accurate KASP genotyping wasn't available for this line to determine segregating homozygous mutant fis2-B plants. The amplicon

sequencing pipeline was successful in genotyping pooled (Cycle 4) and non-pooled (Cycle 3) samples. It is necessary to have pooled samples to decrease sequencing costs, while maximizing the number of samples sequenced.

Single, double, and triple mutant lines generated and identified in Cycle 4 were hybridized with *Ae. tauschii* (Table 2.13, Table 2.14). Although endosperm development did not resemble wild type endosperm development, caryopses did form from these hybridizations. It is possible that with continued pyramiding and genotyping using the amplicon sequencing pipeline, more triple and double homozygous mutants can be generated and potentially form viable wild type resembling endosperm when hybridized with *Ae. tauschii*. *FIS2* is particularly promising. *FIS2* mutants with the genotype aaBBDd produced an average 6.66 seeds per spike. *FIS2* genotypes with more non-functional alleles of *FIS2* (AaBbDd, aabbdd, or even aaBBdd) may out produce self-hybridized wild Cadenza. Work is underway to develop homozygous mutant double and triple mutants of *FIE*, *FIS2*, and *MEA*.

Crossability, instead of endosperm development, is a potential phenotypic observation in this study. Crossability genes have been known to exist in wheat, most notably Kr1, Kr2, Kr3, and Kr4 (Manickavelu et al., 2009). Dominant crossability alleles have been known to repress production of fertile seeds from hybridizations between wheat and some *Poaceae* family members such as barley and rye (Alfares et al., 2009; Manickavelu et al., 2009). Most of these crossability genes are on the same chromosomes as the *FIS2* homeoalleles: Kr1 on 5B, Kr2 on 5A, and Kr3 on 5D (Alfares et al., 2009). The mode of action is likely to be different than crossability genes, as dominant Kr alleles prevent the entry of pollen tubes into maternal tissue (Fedak and Jui, 1982). Hybridizations between wheat and *Ae. tauschii* can still create viable embryos, and therefore functioning pollen tubes to create those embryos.

The phenotype of these mutations may be influenced by the background mutations from the EMS lines. For example, the triple homozygous mutant lines of *MEA* may have a stronger impact on the phenotype of endosperm development when hybridized with *Ae. tauschii*, but the phenotype may be inhibited by the background mutations of the EMS lines. These lines will be backcrossed to the non-EMS parent and selected for heterozygous PRC2 mutations via the amplicon sequencing pipeline. The amplicon sequencing pipeline developed will have future applications in direct hybridization work including introgression into elite wheat cultivars. Research in other allopolyploid crops may also adapt this amplicon sequencing pipeline when experiencing off target amplification with markers.

BIBLIOGRAPHY

BIBLIOGRAPHY

Alfares W, Bouguennec A, Balfourier F, Gay G, Bergès H, Vautrin S, Sourdille P, Bernard M, Feuillet C (2009) Fine Mapping and Marker Development for the Crossability Gene SKr on Chromosome 5BS of Hexaploid Wheat (Triticum aestivum L.). Genetics **183**: 469–81, 1SI

Cox TS, Wu J, Wang S, Cai J, Zhong Q, Fu B (2017) Comparing two approaches for introgression of germplasm from Aegilops tauschii into common wheat. Crop J **5**: 355–362

Curtis BC, Croy LI (1958) The Approach Method of Making Crosses in Small Grains1. Agron J **50**: 49–51

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al (2021) Twelve years of SAMtools and BCFtools. GigaScience. doi: 10.1093/gigascience/giab008

Edwards KJ, Reid AL, Coghill JA, Berry ST, Barker GLA (2009) Multiplex single nucleotide polymorphism (SNP)-based genotyping in allohexaploid wheat using padlock probes. Plant Biotechnol J **7**: 375–390

FAOSTAT (2020) Food and agriculture organization of the United Nations. Food Agric data. http://www.fao.org/faostat/en/#data

Faure J.E. (2011) Double Fertilization in Flowering Plants: Origin, Mechanisms and New Information from in vitro Fertilization. Fertil High Plants Mol Cytol Asp 79

Fedak G, Jui PY (1982) Chromosomes of chinese spring wheat carrying genes for crossability with betzes barley. Can J Genet Cytol **24**: 227–233

Gertz EM, Yu Y-K, Agarwala R, Schäffer AA, Altschul SF (2006) Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. BMC Biol **4**: 41

Holec S, Berger F (2012) Polycomb Group Complexes Mediate Developmental Transitions in Plants. Plant Physiol **158**: 35

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum Aegilops complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci **99**: 8133

Jullien PE, Mosquna A, Ingouff M, Sakata T, Ohad N, Berger F (2008) Retinoblastoma and Its Binding Partner MSI1 Control Imprinting in Arabidopsis. PLOS Biol **6**: e194

Köhler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W (2003) Arabidopsis MSI1 is a component of the *MEA*/*FIE* Polycomb group complex and required for seed development. EMBO J **22**: 4804–4814

Kradolfer D, Hennig L, Köhler C (2013) Increased Maternal Genome Dosage Bypasses the Requirement of the FIS Polycomb Repressive Complex 2 in Arabidopsis Seed Development. PLOS Genet **9**: e1003163

Lagudah ES, Appels R, Brown AHD, McNeil D (1991) The molecular–genetic analysis of Triticum tauschii, the D-genome donor to hexaploid wheat. Genome **34**: 375–386

Lanning SP, Blake NK, Sherman JD, Talbert LE (2008) Variable Production of Tetraploid and Hexaploid Progeny Lines from Spring Wheat by Durum Wheat Crosses. Crop Sci **48**: 199–202

Lewis EB (2004) A Gene Complex Controlling Segmentation in Drosophila. *In* HD Lipshitz, ed, Genes Dev. Cancer Life Work Edw. B Lewis. Springer US, Boston, MA, pp 205–217

Li A, Liu D, Yang W, Kishii M, Mao L (2018) Synthetic Hexaploid Wheat: Yesterday, Today, and Tomorrow. Engineering **4**: 552–558

Ling H-Q, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, et al (2013) Draft genome of the wheat A-genome progenitor Triticum urartu. Nature **496**: 87

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, et al (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res **47**: W636–W641

Makhoul M, Rambla C, Voss-Fels KP, Hickey LT, Snowdon RJ, Obermeier C (2020) Overcoming polyploidy pitfalls: a user guide for effective SNP conversion into KASP markers in wheat. Theor Appl Genet **133**: 2413–2430

Manickavelu A, Koba T, Mishina K, Sassa H (2009) Molecular characterization of crossability gene Kr1 for intergeneric hybridization in Triticum aestivum (Poaceae: Triticeae). Plant Syst Evol **278**: 125–131

Mcfadden ES, Sears ER (1946) The origin of triticum spelta and its free-threshing hexaploid relatives. J Hered **37**: 81–89

Miranda LM, Murphy JP, Marshall D, Leath S (2006) Pm34: a new powdery mildew resistance gene transferred from Aegilops tauschii Coss. to common wheat (Triticum aestivum L.). Theor Appl Genet **113**: 1497–1504

Nallamilli BRR, Zhang J, Mujahid H, Malone BM, Bridges SM, Peng Z (2013) Polycomb group gene Os*FIE*2 regulates rice (Oryza sativa) seed development and grain filling via a mechanism distinct from Arabidopsis. PLoS Genet **9**: e1003322–e1003322

Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res **31**: 3812–3814

Olson EL, Rouse MN, Pumphrey MO, Bowden RL, Gill BS, Poland JA (2013) Simultaneous transfer, introgression, and genomic localization of genes for resistance to stem rust race TTKSK (Ug99) from Aegilops tauschii to wheat. Theor Appl Genet **126**: 1179–1188

Saluja M, Kaur S, Bansal U, Subhash Chander Bhardwaj, Chhuneja P (2018) Molecular mapping of linked leaf rust resistance and non-glaucousness gene introgressed from Aegilops tauschii Coss. in hexaploid wheat Triticum aestivum L. Plant Genet Resour **16**: 82–88

Schmidt A, Wöhrmann HJP, Raissig MT, Arand J, Gheyselinck J, Gagliardini V, Heichinger C, Walter J, Grossniklaus U (2013) The Polycomb group protein MEDEA and the DNA methyltransferase MET1 interact to repress autonomous endosperm development in Arabidopsis. Plant J **73**: 776–787

Shen X, Gmitter FG, Grosser JW (2011) Immature Embryo Rescue and Culture. *In* TA Thorpe, EC Yeung, eds, Plant Embryo Cult. Methods Protoc. Humana Press, Totowa, NJ, pp 75–92

Velu G, Guzman C, Mondal S, Autrique JE, Huerta J, Singh RP (2016) Effect of drought and elevated temperature on grain zinc and iron concentrations in CIMMYT spring wheat. J Cereal Sci **69**: 182–186

Weinhofer I, Hehenberger E, Roszak P, Hennig L, Köhler C (2010) H3K27me3 Profiling of the Endosperm Implies Exclusion of Polycomb Group Protein Targeting by DNA Methylation. PLoS Genet **6**: e1001152

Wiersma AT, Pulman JA, Brown LK, Cowger C, Olson EL (2017) Identification of Pm58 from Aegilops tauschii. Theor Appl Genet **130**: 1123–1133

Yadegari R, Kinoshita T, Lotan O, Cohen G, Katz A, Choi Y, Katz A, Nakashima K, Harada JJ, Goldberg RB, et al (2000) Mutations in the *FIE* and *MEA* Genes That Encode Interacting Polycomb Proteins Cause Parent-of-Origin Effects on Seed Development by Distinct Mechanisms. Plant Cell **12**: 2367

Yao Z, You FM, N'Diaye A, Knox RE, McCartney C, Hiebert CW, Pozniak C, Xu W (2020) Evaluation of variant calling tools for large plant genome re-sequencing. BMC Bioinformatics **21**: 360

CHAPTER III

DEVELOPMENTAL CHANGES IN WINTER AND SPRING WHEAT IN VARIOUS PLANTING DENSITIES

Abstract

The wheat breeding program at Michigan State University utilizes a modified bulk breeding method to rapidly advance populations from the F2 to F4 stage under greenhouse conditions. To determine the influence of planting density on yield components and plant architecture in greenhouse settings, two experiments were carried out using two spring and winter wheat. Planting densities ranged from 2 to 39.5 seeds per 10 cm^2. All traits measured increased as planting density increased, decreased as planting density increased, or followed no trend as planting density increased. Experiments with both spring and winter wheat demonstrated that increased planting density reduces plant height and reduces the number spikelets per spike. Grain dimensions are also influenced with reduced grain length and width observed at higher planting densities.

Introduction

Wheat yield components demonstrate plasticity in response to the environment. Nutrients, water, light, soil volume, and other plants are influential environmental factors for wheat growth habit and yield. Agronomic practices that provide an environment without limitations to nitrogen and micronutrients result in higher grain yields (Black et al., 1947). However, modern agronomic practices require high density monocultures that may create competition for resources.

Competitive high density planting environments can change plant architecture and in turn affect components of grain yield. Competition is often less severe between relatives, or kin, within a species due to kin recognition (Dudley and File, 2007; Caffaro et al., 2013). Some species, such as *Lupinus angustifolius*, do not have kin recognition and have similar competition

between kin and non-kin lines (Milla et al., 2009). Kin recognition has other limits, specifically under high density plantings in commercial crop production. High density plantings force roots to be in close proximity and compete for water and nutrients (Nord et al., 2011). This competition may lead to overallocation of resources to roots, or root over-proliferation, to the detriment of above ground biomass contributing to grain yield (Nord et al., 2011; Zhu et al., 2019). Gersani et al. (2001) demonstrated that non-competing *Glycine max* produce 30% more yield than *Glycine max* competing with kin *Glycine max*.

Few grain planting density studies in greenhouse conditions have been conducted. Cisar et al. (1982) observed the impacts of planting density of oats in greenhouse settings. Cisar et al. (1982) found that as planting density increased, so did the number of spikes containing 1 or 0 seeds. These oats were grown in stressed conditions which intensified the competition induced changes in yield components (Cisar et al., 1982). Recently, data gathered from Olson et al. (2021) suggested that some yield components may be genotype dependent, while others genotype independent. As planting densities increase, it is expected that yield components, such as number of seeds per spike, will be altered in ways that may or may not be genotypic dependent.

Characterizing the impact of planting density on wheat phenotypes, especially yield components, will allow growers and breeders to know the full extent of risks and benefits of a range of planting densities. In this study, two experiments were carried out, one using spring wheat and one using winter wheat varieties, to characterize the impacts of planting density on yield components and plant architecture under greenhouse conditions. Planting densities ranged from 2 to 39.5 seeds per 10 cm^2. Phenotypes evaluated included plant height, spikelets per spike,

mature spikelet count, spikes per pot, total seed weight, average seed weight, seed count, seed length, seed width, seeds per spike, and seeds planted vs number of spikes.

Materials and methods

Experimental design

Winter wheat and spring wheat studies were conducted in a Randomized Complete Block Design (RCBD) created with PRISM (Central Software Solutions, Orange City, IA). 11 planting density treatments were applied (Table 3.1) increasing from 2 to 39.5 seeds per 10 cm^2, or 32 to 640 seeds per 413 cm^2 pot. Each planting density treatment was replicated four times across two different winter and spring wheat varieties. Seeds for each density level were counted using a Seed Counter R-25 (DATA Technologies, Baku, Yasamal AZ).

Seeds per square 10 cm^2	Seeds per pot
2	32
4	64
8	128
12	192
16	256
20	320
24	384
28	448
31.5	512
35.5	576
39.5	640

Table 3.1: Seed density treatments.

Winter wheat

The winter wheat lines, Whitetail and DF-112-R, were vernalized in and placed in a Dart 22.9 cm x 24.1 cm x 7.6 cm ClearSeal® Hinged Lid Plastic Container (Dart Container

Corporation, Mason, MI) at 4°C for 8 weeks. Each box contained an 18.4 cm x 19.1 cm section of 200 M weight White Verigood Blotter paper (Anchor Paper Company, St. Paul, Minnesota) saturated with 1% Thiram (T24201, Sigma Aldrich, St. Louis, MO) and DI water solution. 32, 64, 128, 192, 256, 320, 384, 448, 512, 576, or 640 seeds were added to each salad box. A light layer of Thiram solution was added over the seeds to prevent fungal growth. Thiram solution was added as needed throughout vernalization to ensure the Blotter paper stayed saturated, but not over saturated. The salad boxes were under LED lighting from WLFA2 series Fixtures (Super Bright LEDs Inc., Earth City, MO) with an 8-hour photoperiod.

After vernalization, the seed mat of each salad box was transplanted in 20.3 cm x 20.3 cm x 19.1 cm Gro Pro® Premium Square Black Plastic Pots (Hawthorne Gardening Co., Vancouver, WA) with saturated Sure Mix potting media (Michigan Grower Products Inc., Galesburg, MI). The seed mats were placed on top of the soil and were lightly watered from above within 20 minutes of transplanting. Overhead watering occurred every 20-30 minutes for at least 5 hours after transplanting. Subdue, a systemic fungicide, was applied to the soil after transplanting to prevent downy mildew. 2'x4' greenhouse trays (Greenhouse Megastore, Danville, IL) were used with two 1/16" holes drilled for drainage. Flood irrigation was used for both watering and water-soluble fertilizer. Fertilizer (20-20-20 NPK) was used one week after planting (188g / 5-gal H2O at 1:50 dilution) and again once weekly until Feekes 11. Plants were watered as needed and progressed to daily as plants matured. Water was cutoff when 75% of all plants in each pot reached physiological maturity (Feekes 11). The experiment was carried out in early spring and plants were under a 22-hour photoperiod.

Spring wheat

The spring wheat lines, Melba and Seahawk, were germinated in clear salad boxes at 40F for 5 days. Each box contained a section of 200 M weight White Verigood Blotter paper (Anchor Paper Company, St. Paul, Minnesota) saturated with 1% Thiram (T24201, Sigma Aldrich, St. Louis, MO) and DI water solution. 32, 64, 128, 192, 256, 320, 384, 448, 512, 576, or 640 seeds were added to each salad box. A light layer of Thiram solution was added over the seeds to prevent fungal growth. After 5 days, the germinated seeds were transplanted to 8" x 8" x 7.5" Gro Pro® Premium Square Black Plastic Pots (Hawthorne Gardening Co., Vancouver, WA) with saturated Sure Mix potting media (Michigan Grower Products Inc., Galesburg, MI). The seeds were lightly covered with saturated soil and watered every 20-40 minutes for several hours after planting. Subdue was applied to the soil after transplanting. Greenhouse conditions were similar to winter wheat however the experiment was carried out during mid-winter and the lighting was switched to Phillips LED lighting for a 22-hour photoperiod. Watering and fertilizer were carried out the same as winter wheat.

Data collection

Data was collected for several yield components. Grain length and grain width were measured using a caliper. Number of spikes per pot were counted after senescence. Seed count was collected using DATA Technologies Seed Counter R-25 plus. Spikelet count was collected for mature and total spikelets. Winter genotypes had several spikes with immature spikelets, or spikelets that do not contain seeds. 10 subsamples were taken for grain length, grain width,

individual grain weight, mature and total spikelets per spike, and height. Seeds per spike was calculated by dividing total number of spikes per pot over total seed count.

Data analysis

Data analysis including ANOVA, LSD, and adjusted R squared values were generated in R studio. All traits were initially screened with the full model version of ANOVA (planting density and genotype and interaction of planting density and genotype) (Table 3.2). Factors in the models that were indicated to be significant at $P<0.05$ were used to create the final model used for LSD. LSD was calculated using R package agricolae. All graphs were created using R package ggplot2.

Results

Spring			
	Planting density	Line	Interaction
Height	4.11E-12**	0.9032	0.0136*
Spikelets per spike	3.48E-14**	0.5362	0.6223
Mature spikelets per spike	2.00E-16**	0.03241*	0.87185
Spikes per pot	2.00E-16**	0.147	0.9462
Total seed weight	0.3365	0.05197	0.77064
Average seed weight	2.41E-05**	6.07E-09**	0.6085
Seed count	0.9588	1.73E-06**	0.481
Seed length	0.0509	0.000692**	0.4538
Seed width	1.12E-05**	0.02347*	0.06101
Seeds per spike	2.20E-16**	3.66E-06**	0.5636
Seeds planted vs. Number of spikes	2.00E-16**	0.1441	0.41
Spikelets per spike	3.47E-16**	0.0286*	0.468

Table 3.2: P-values of independent variables planting density, line, and the interaction between planting density and line. * indicates significance at $P<0.05$, ** indicates significance at $P<0.01$.

Table 3.2 (cont'd)

Winter			
	Planting density	Line	Interaction
Height	2.00E-16**	0.277	0.056
Spikelets per spike	3.47E-16**	0.0286*	0.468
Seed width	8.50E-12**	0.03839*	0.85565
Seeds per spike	2.20E-16**	0.002596**	0.27199
Seeds planted vs. number of spikes	2.00E-16**	0.1236	0.6642
Mature spikelets per spike	2.00E-16**	0.4623	0.987
Spikes per pot	2.20E-16**	0.04797*	0.00254**
Total seed weight	0.0008054**	0.7317	0.91644
Average seed weight	6.15E-07**	0.002308**	0.4944
Seed count	0.3298	0.01902*	0.9495
Seed length	8.47E-06**	5.12E-09**	0.4446

Plant height

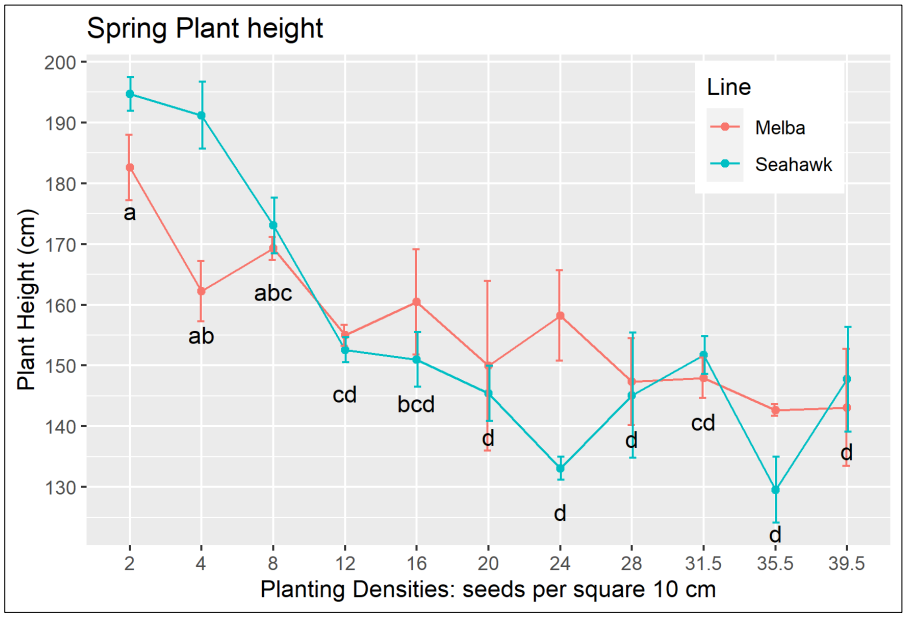
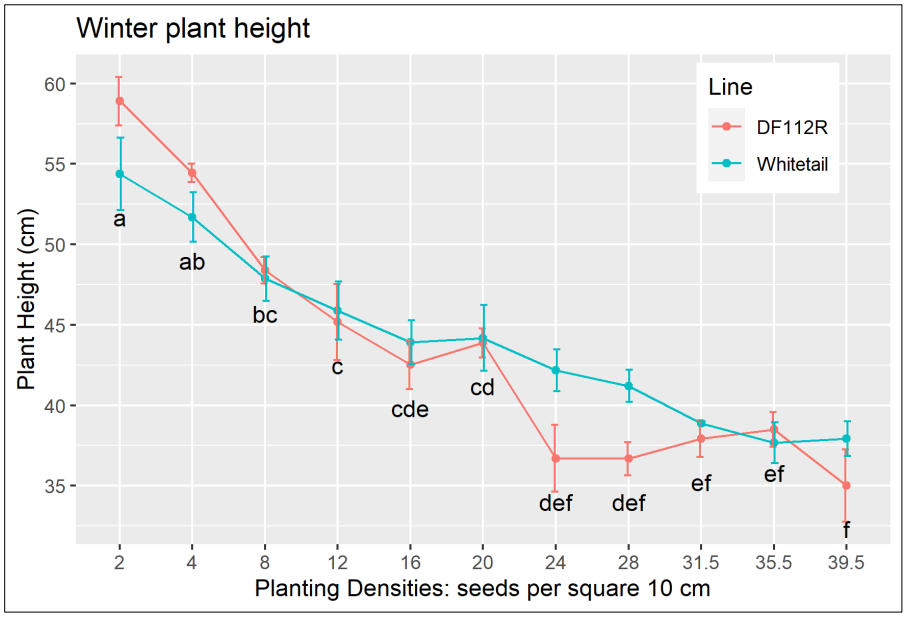


Figure 3.1: Average plant height of spring wheat and winter wheat for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for best fitted model. Bars are standard error.

Figure 3.1 (cont'd)



Both spring wheat and winter wheat plant height decreased as planting density increased. In spring wheat, height decreased significantly from 2 to 16 seeds/10cm^2 and from 16 to 24 seeds/10cm^2 ($P<0.01$). Mean plant height was highest at 188.6 cm at 2 seeds/10cm^2 and dropped to 136.1 cm at 35.5 seeds/10cm^2. Height was not significantly different for planting densities 16 to 39.5 seeds/10cm^2. From 16 to 39.5 seeds/10cm^2, mean plant height ranged from 155.75 cm to 136.11 cm. Interaction between planting density and line was significant for height in spring wheat ($P<0.05$).

In winter wheat, height decreased significantly from 2 to 20 seeds/10cm^2 ($P<0.01$). Mean plant height was highest at 56.63 cm at 2 seeds/10cm^2 and dropped to 43.22 cm at 16 seeds/10cm^2. Height was not significantly different in planting densities 24 to 39.5 seeds/10cm^2. From 24 to 39.5 seeds/10cm^2, mean plant height ranged from 39.43 cm to 36.36 cm.

Spikelets per spike

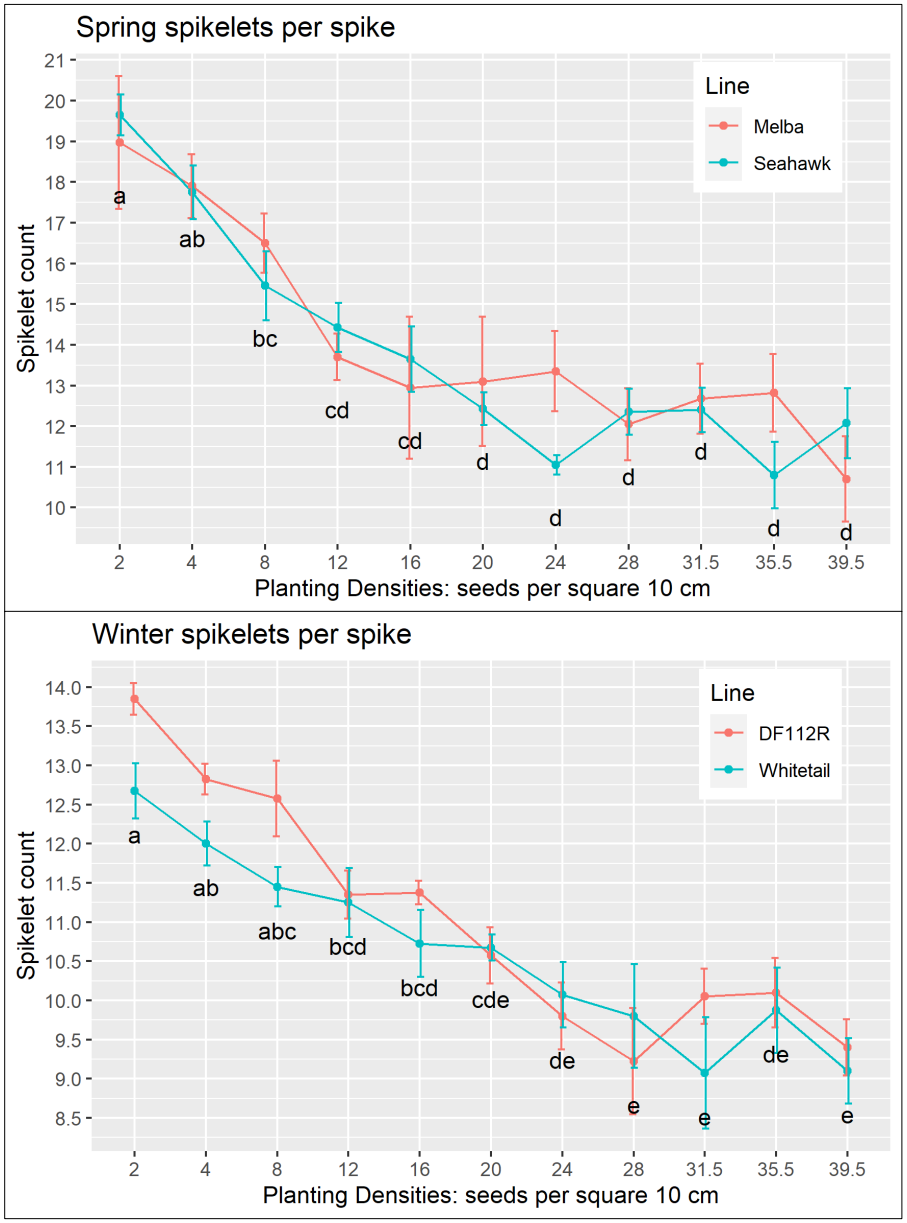


Figure 3.2: Spikelets per spike of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Spikelets per spike include both mature spikelets with seeds and immature spikelets without seeds. Letters are LSD values for best fitted model. Bars are standard error.

Both spring wheat and winter wheat spikelets per spike decreased as planting density increased. In spring wheat, spikelets per spike decreased significantly from 2 seeds/10cm^2 to 8 seeds/10cm^2 ($P<0.01$). Mean spikelets per spike ranged from 19.312 spikelets at 2 seeds/10cm^2

to 11.38 spikelets at 39.5 seeds/10cm^2. Spikelets per spike were not significantly different in planting densities 12 to 39.5 seeds/10cm^2.

In winter wheat, spikelets per spike decreased significantly from 2 to 16 seeds/10cm^2 ($P<0.01$). Mean spikelets per spike ranged from 13.26 spikelets at 2 seeds/10cm^2 to 9.25 spikelets at 39.5 seeds/10cm^2. Spikelets per spike were not significantly different in planting densities 20 to 39.5 seeds/10cm^2. Winter wheat lines DF112R and Whitetail were significantly different from one another for spikelets per spike ($P<0.05$).

Mature spikelets per spike

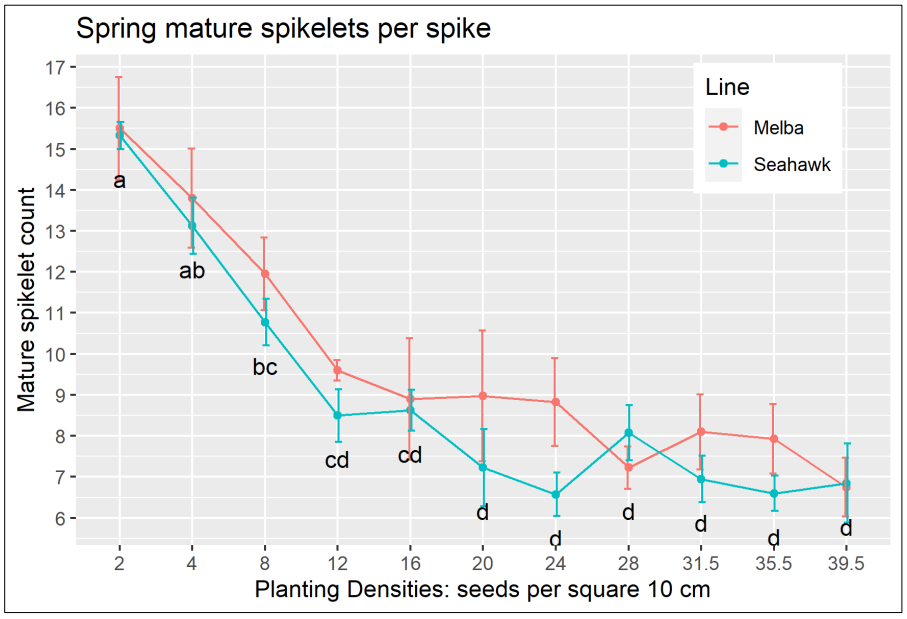
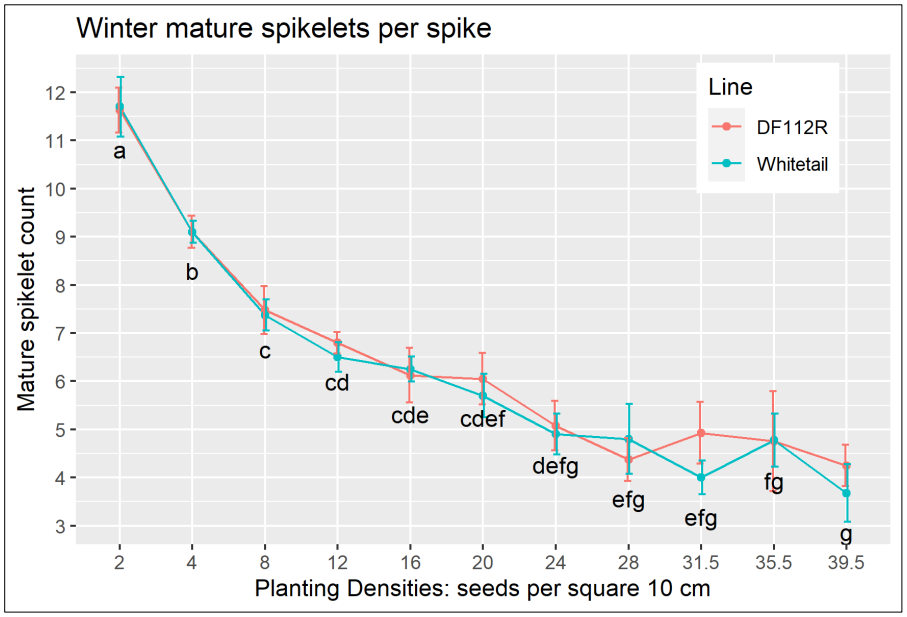


Figure 3.3: Mature spikelets per spike of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for best fitted model. Bars are standard error.

Figure 3.3 (cont'd)



Both spring wheat and winter wheat mature spikelets per spike decreased as planting density increased. In spring wheat, mature spikelets per spike decreased significantly from 2 seeds/10cm^2 to 8 seeds/10cm^2 ($P<0.01$). Mean mature spikelets per spike peaked at 15.4 spikelets at 2 seeds/10cm^2 and were lowest at 6.8 spikelets at 39.5 seeds/10cm^2. Spikelets per spike were not significantly different for planting densities 12 to 39.5 seeds/10cm^2. Spring wheat lines Melba and Seahawk were significantly different from one another for mature spikelets per spike ($P<0.05$).

In winter wheat, mature spikelets per spike decreased significantly from 2 seeds/10cm^2 to 20 seeds/10cm^2 ($P<0.01$). Mean mature spikelets per spike were highest at 11.66 spikelets at 2 seeds/10cm^2 and lowest at 3.96 spikelets at 39.5 seeds/10cm^2. Mature spikelets per spike were not significantly different for planting densities 24 to 39.5 seeds/10cm^2.

Spikes per pot

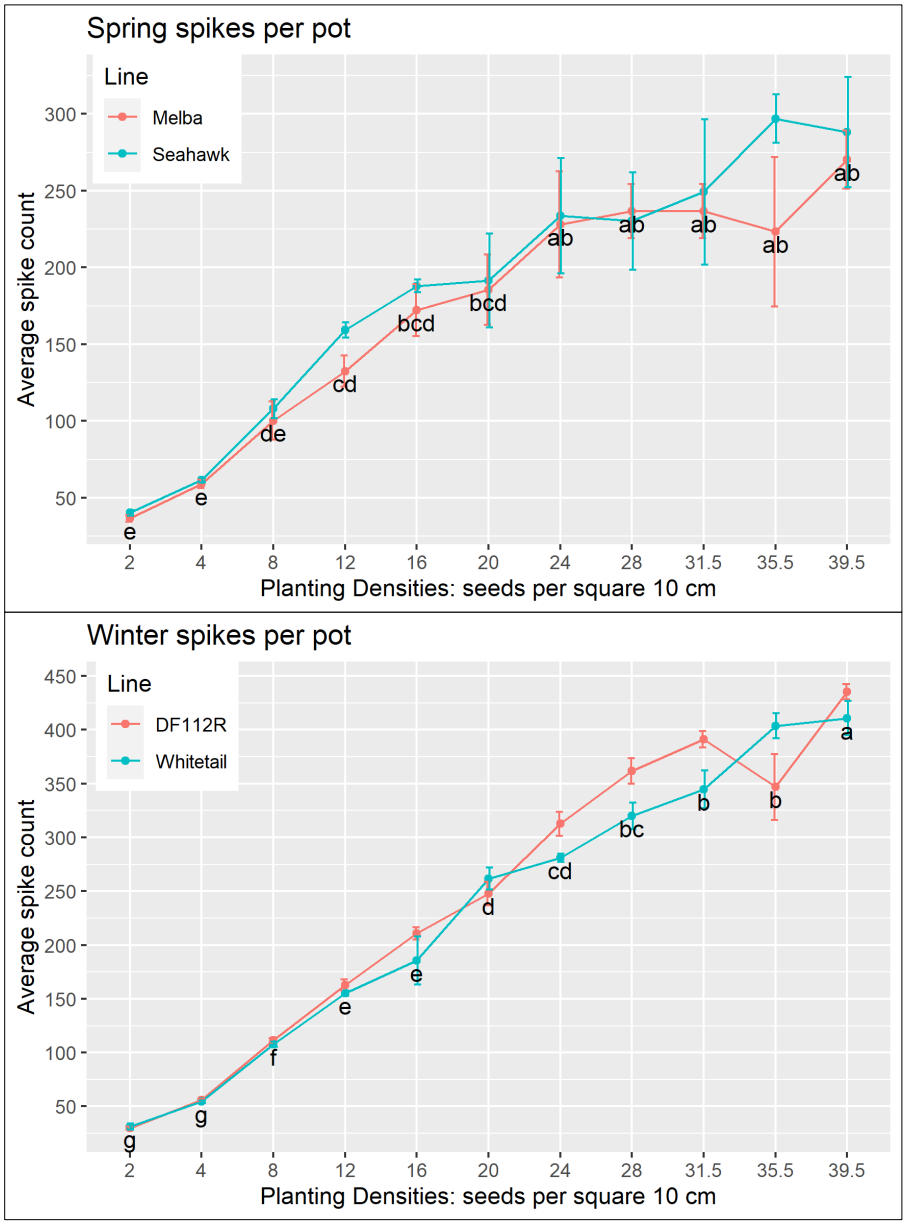


Figure 3.4: Spikes per pot of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for the best fitted model. Bars are standard error.

Spikes per pot increased as planting density increased for both spring wheat and winter wheat. In spring wheat, spikes per pot increased significantly from 2 seeds/10cm^2 to 20 seeds/10cm^2 ($P<0.01$). Mean spikes per pot ranged from 38.37 spikes at 2 seeds/10cm^2 to 279.25

spikes at 39.5 seeds/10cm^2. Spikes per pot were not significantly different for planting densities 24 to 39.5 seeds/10cm^2.

In winter wheat, spikes per pot increased significantly throughout the range of the planting densities 2 seeds/10cm^2 to 39.5 ($P<0.01$). Mean spikes per pot started at 30.5 spikes at 2 seeds/10cm^2 and increased to 423.125 spikes at 39.5 seeds/10cm^2. Winter wheat lines DF112R and Whitetail were significantly different from one another for spikes per pot ($P<0.05$). There was also significant interaction between planting density and line ($P<0.01$).

Total seed weight

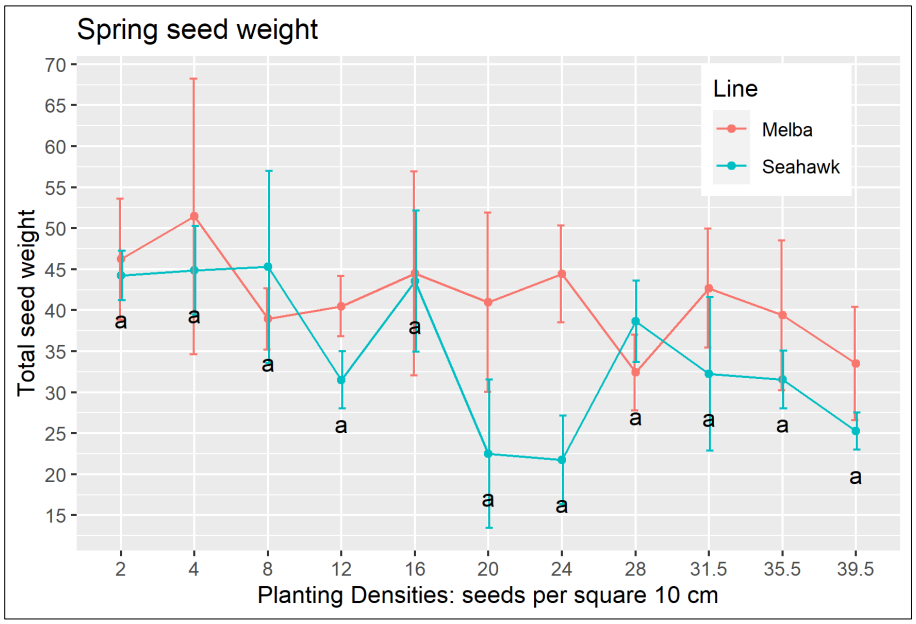
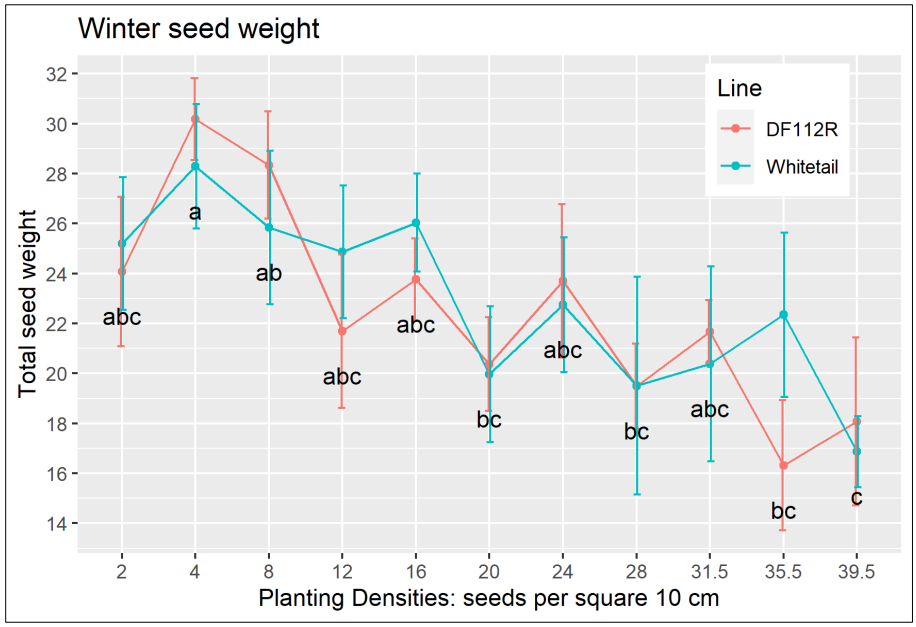


Figure 3.5: Total seed weight of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for best fitted model. Bars are standard error.

Figure 3.5 (cont'd)



Spring wheat and winter wheat experiments exhibited different trends for total seed weight. In spring wheat, total seed weight did not significantly change as planting density increased. Spring seed weight was almost significantly different between lines (P=0.51). The lack of significance among planting densities is likely due to unusually low total seed weights of Seahawk at 20 and 24 seeds/10cm^2. Mean total seed weight ranged from 29.38 g. at 39.5 seeds/10cm^2 to 48.15 g. at 4 seeds/10cm^2.

In winter wheat, total seed weight decreased significantly from 4 seeds/10cm^2 to 39.5 seeds/10cm^2 (P<0.01). Several planting densities were not significantly different from one another: 2, 12, 16, 24, and 31.5 seeds/10cm^2. Mean total seed weight ranged from 17.47 g. at 39.5 seeds/10cm^2 to 29.23 g. at 4 seeds/10cm^2.

Average seed weight

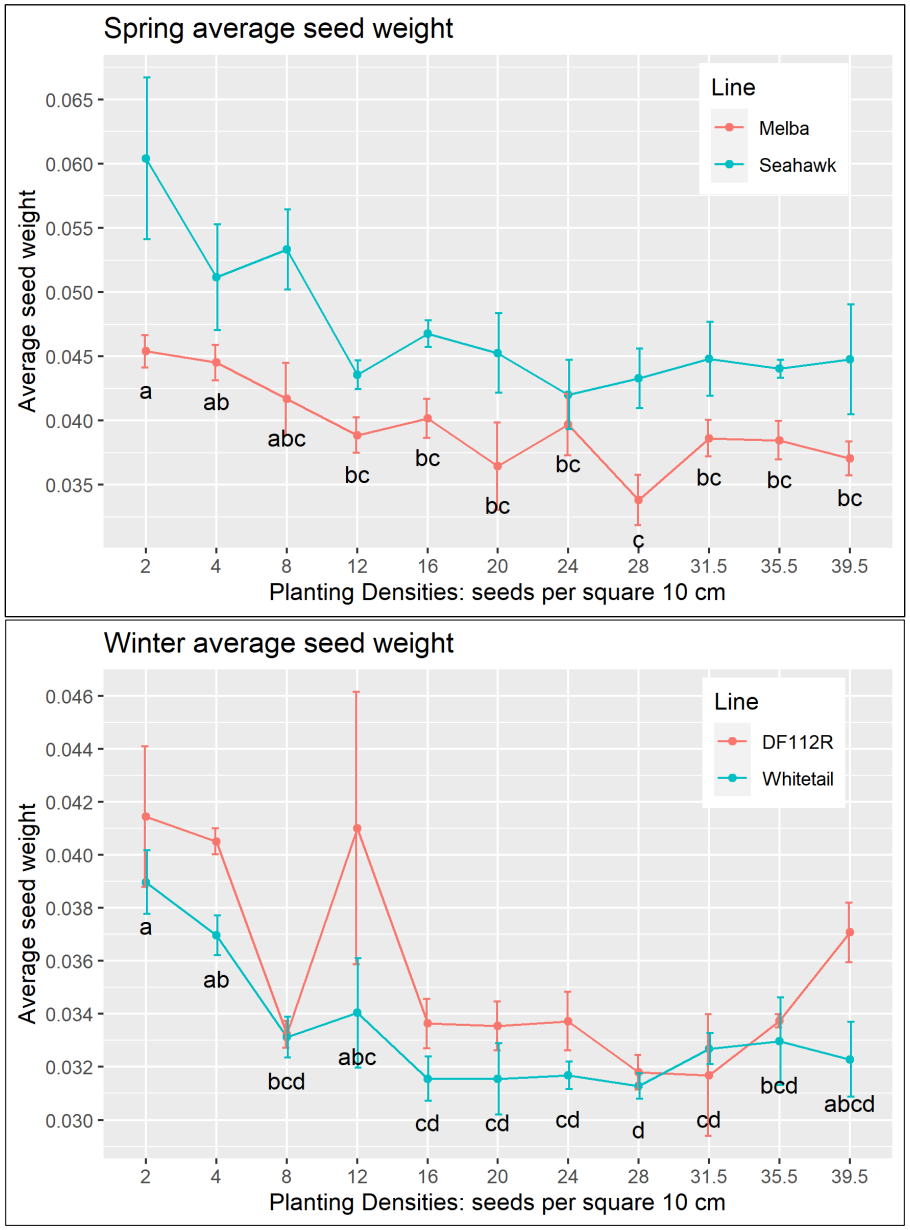


Figure 3.6: Average seed weight of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for best fitted model. Bars are standard error.

Both spring wheat and winter average seed weight decreased as planting density increased. In spring wheat, average seed weight decreased significantly from 2 seeds/10cm^2 to 12 seeds/10cm^2 ($P<0.01$). Average seed weight was highest at 0.052g at 2 seeds/10cm^2 and decreased to 0.038 g. at 28 seeds/10cm^2. Average seed weight was not significantly different for

planting densities 16 to 39.5 seeds/10cm^2. Spring wheat lines Melba and Seahawk were significantly different from each other (P<0.01).

In winter wheat, average seed weight decreased significantly from 2 seeds/10cm^2 to 12 seeds/10cm^2 (P<0.01). Average seed weight was highest at 0.04 g. at 2 seeds/10cm^2 and lowest at 0.031 g. at 28 seeds/10cm^2. Average seed weight was not significantly different for planting densities 16 to 39.5 seeds/10cm^2. Winter wheat lines DF112R and Whitetail were significantly different from each other (P<0.01).

Seed count

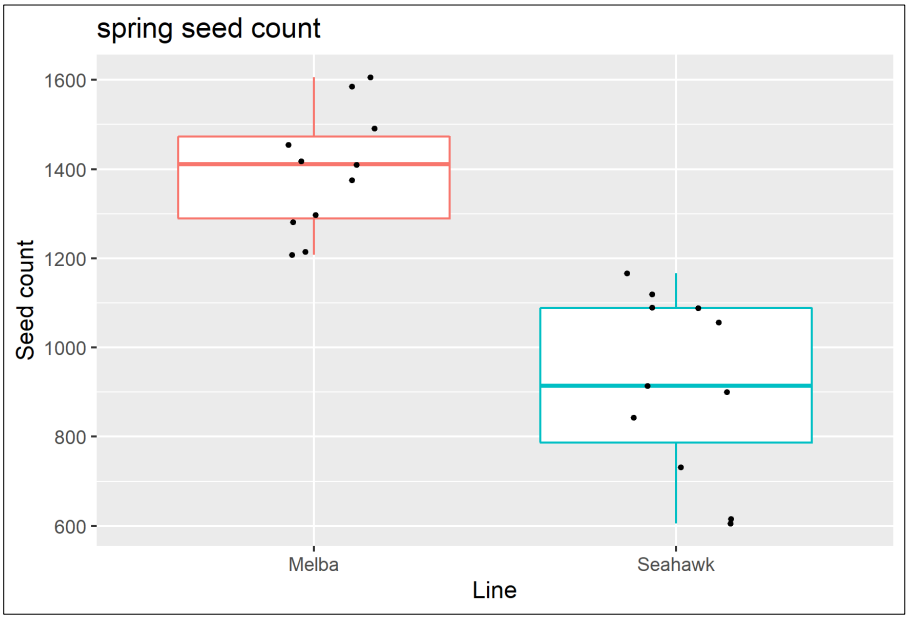
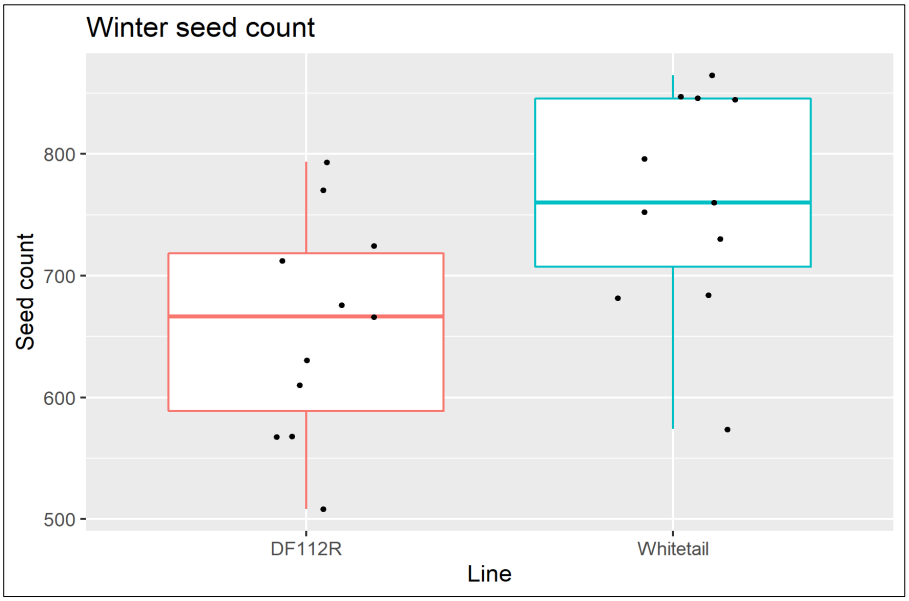


Figure 3.7: Seed count of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Black dots on plot represent mean seed count for each planting density.

Figure 3.7 (cont'd)



Both spring wheat and winter seed count did not have significant changes as planting density increased. However, lines within spring and winter wheat were significantly different from each other ($P<0.01$ for spring and $P<0.05$ for winter). In spring wheat, mean seed count was higher for Melba at 1394.93 seeds where Seahawk's mean seed count was 920.84 seeds. Winter wheat seed count was slightly higher for Whitetail than for DF112R. Mean seed count for Whitetail was 761.977 seeds and for DF112R was 657.068 seeds.

Seed length

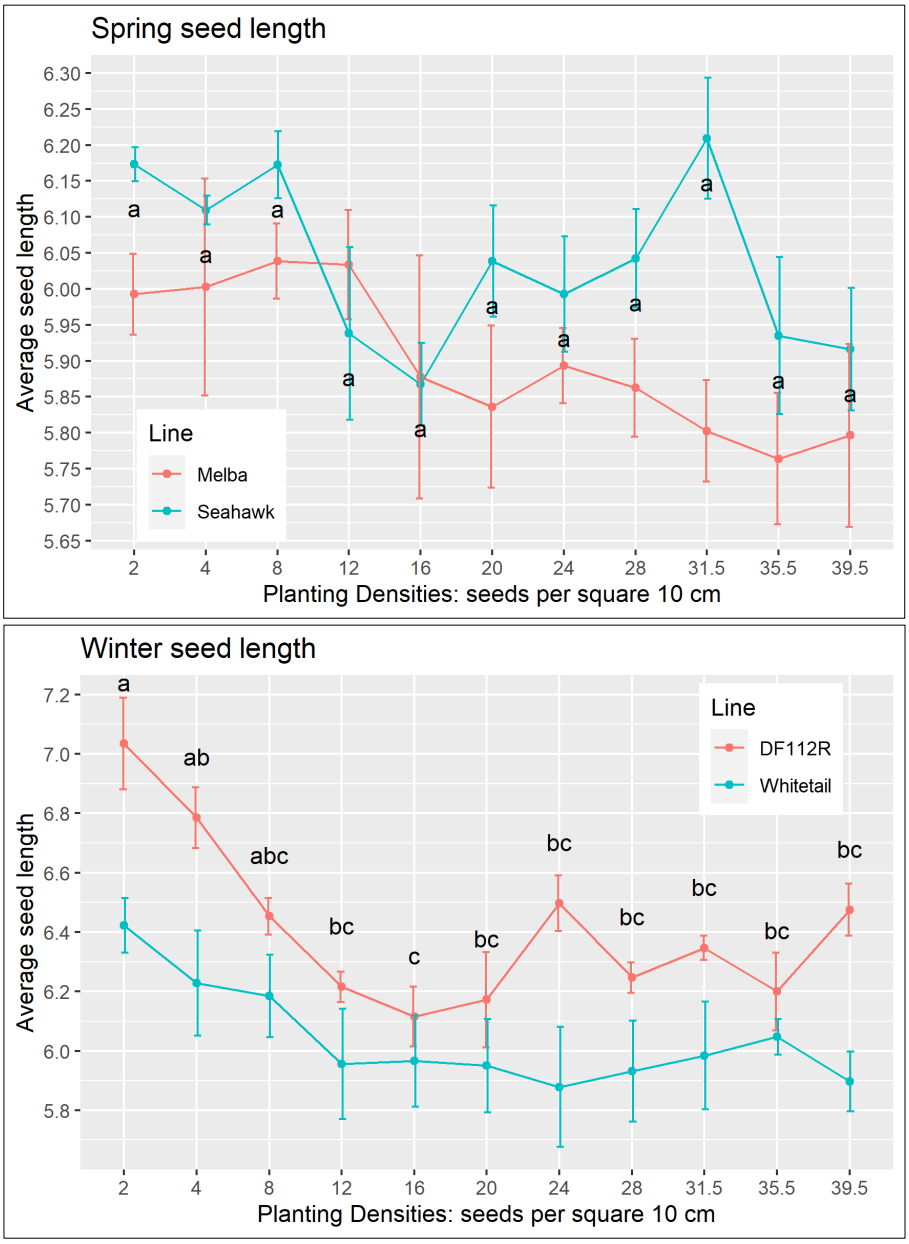


Figure 3.8: Seed length of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for best fitted model. Bars are standard error.

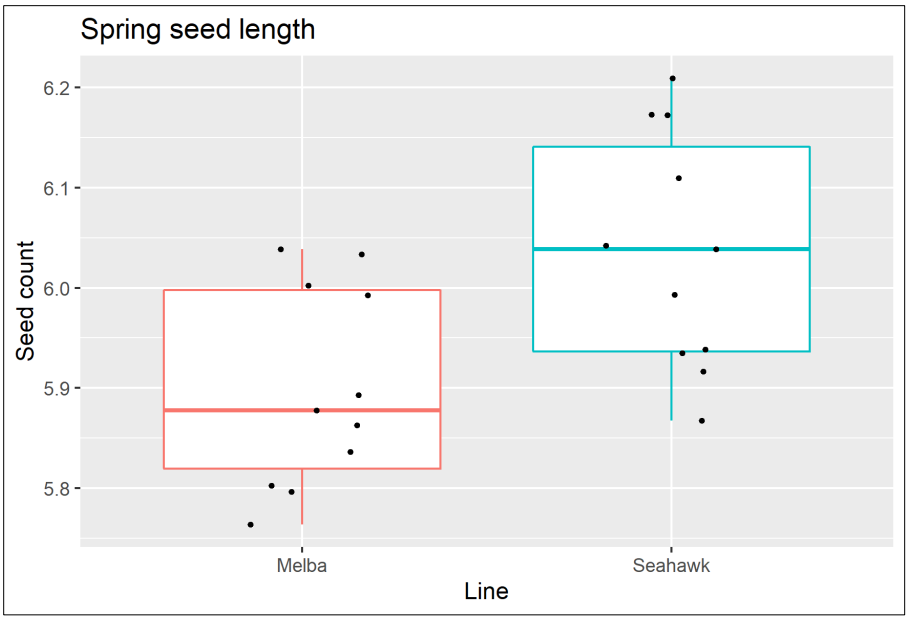


Figure 3.9: Seed length of spring wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Black dots on plot represent mean seed count for each planting density.

Spring wheat and winter wheat had different trends for seed length. In spring wheat, seed length did not have any significant changes as planting density increased. However, spring wheat lines Melba and Seahawk were significantly different from each other ($P<0.01$). Seahawk seed length was on average larger than Melba. Seahawk seed length was on average 5.89 mm. Melba was on average of 6.035 mm.

Winter seed length decreased slightly as planting density increased. In winter wheat, seed length decreased significantly from 2 seeds/10cm^2 to 16 seeds/10cm^2 ($P<0.01$). Mean seed length ranged from 6.04 mm. at 16 seeds/10cm^2 to 6.728 mm. at 2 seeds/10cm^2. Winter wheat lines DF112R and Whitetail were significantly different from each other for seed length ($P<0.01$).

Seed width

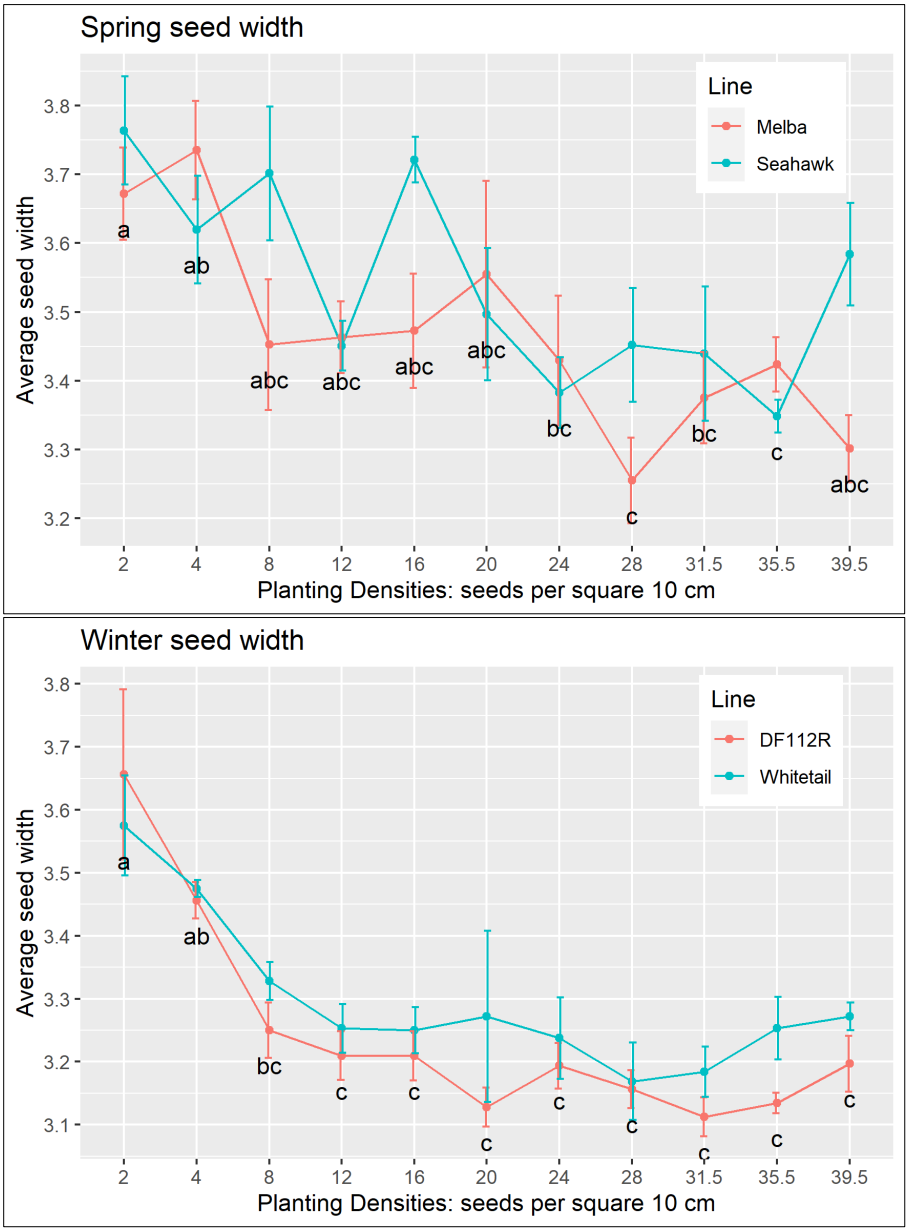


Figure 3.10: Seed width of spring wheat and winter wheat lines for planting densities 2 to 39.5 seeds/10cm^2. Letters are LSD values for the best fitted model. Bars are standard error.

Both spring wheat and winter seed width decreased as planting density increased. In spring wheat, seed width decreased significantly from 2 seeds/10cm^2 to 28 seeds/10cm^2 ($P<0.01$). Seed width ranged from 3.35 mm. at 28 seeds/10cm^2 to 3.717 mm. at 2 seeds/10cm^2.

Spring wheat lines Melba and Seahawk were significantly different from each other for seed width ($P<0.05$).

In winter wheat, seed width decreased significantly from 2 seeds/10cm^2 to 8 seeds/10cm^2 ($P<0.01$). Seed width ranged from 3.148 mm. at 31.5 seeds/10cm^2 to 3.61 mm. at 2 seeds/10cm^2. Winter wheat lines DF112R and Whitetail were significantly different from each other for seed width ($P<0.05$).

Seeds per spike

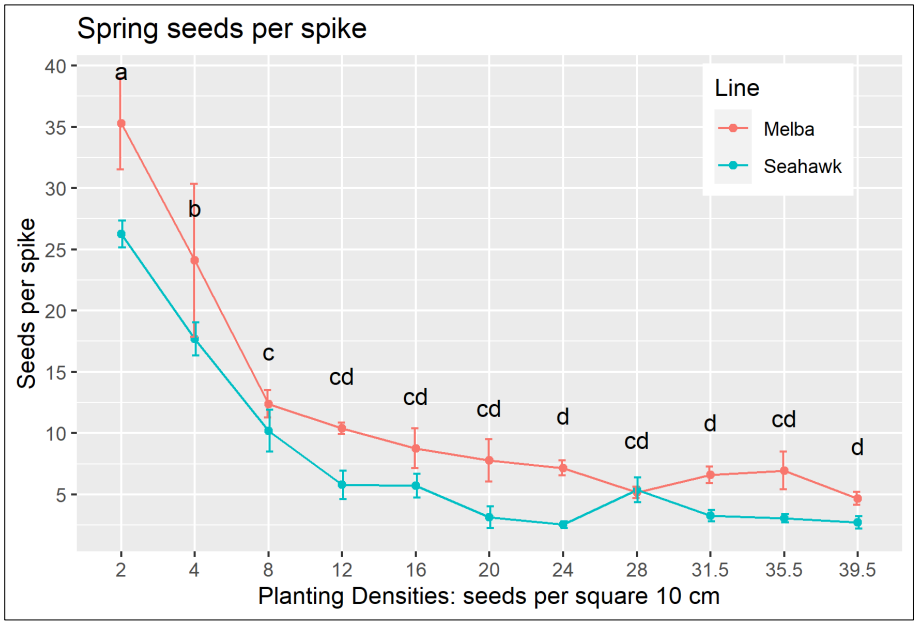
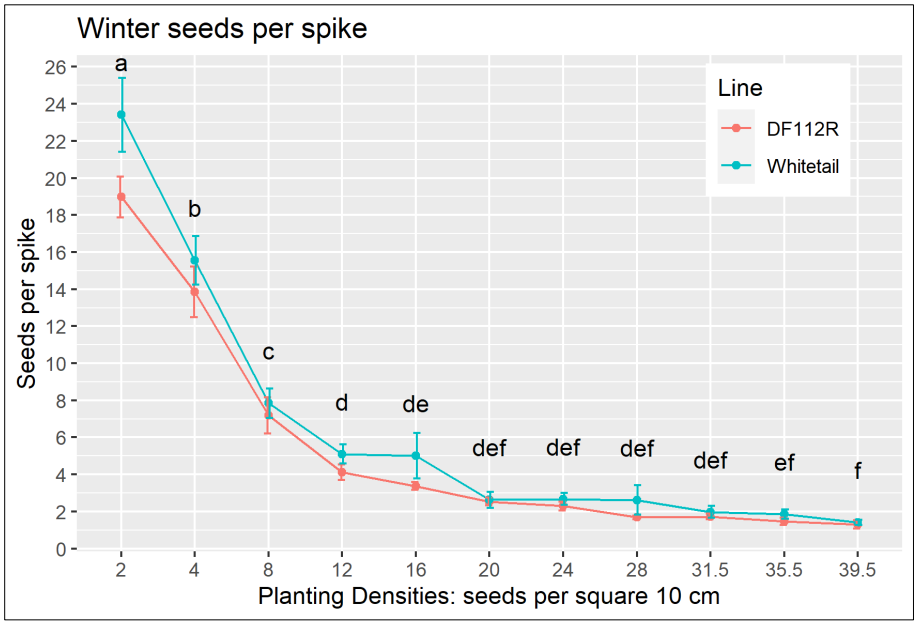


Figure 3.11: Seeds per spike of spring wheat and winter wheat lines for planting densities of 2 seeds/10cm^2 to 39.5 seeds/10cm^2. Letters are LSD values for best fitted model. Bars are standard error.

Figure 3.11 (cont'd)



Both spring wheat and winter seeds per spike decreased as planting density increased. In spring wheat, seeds per spike decreased significantly from 2 seeds/10cm^2 to 12 seeds/10cm^2 ($P<0.01$). Mean seeds per spike ranged from 3.68 seeds at 39.5 seeds/10cm^2 to 30.76 seeds at 2 seeds/10cm^2. Spring wheat lines Melba and Seahawk were significantly different from each other for seeds per spike ($P<0.01$).

In winter wheat, seeds per spike decreased significantly from 2 seeds/10cm^2 to 16 seeds/10cm^2 ($P<0.01$). Mean seeds per spike ranged from 1.34 seeds at 39.5 seeds/10cm^2 to 21.18 seeds at 2 seeds/10cm^2. Winter wheat lines DF112R and Whitetail were significantly different from each other for seeds per spike ($P<0.01$).

Seeds planted vs. number of spikes

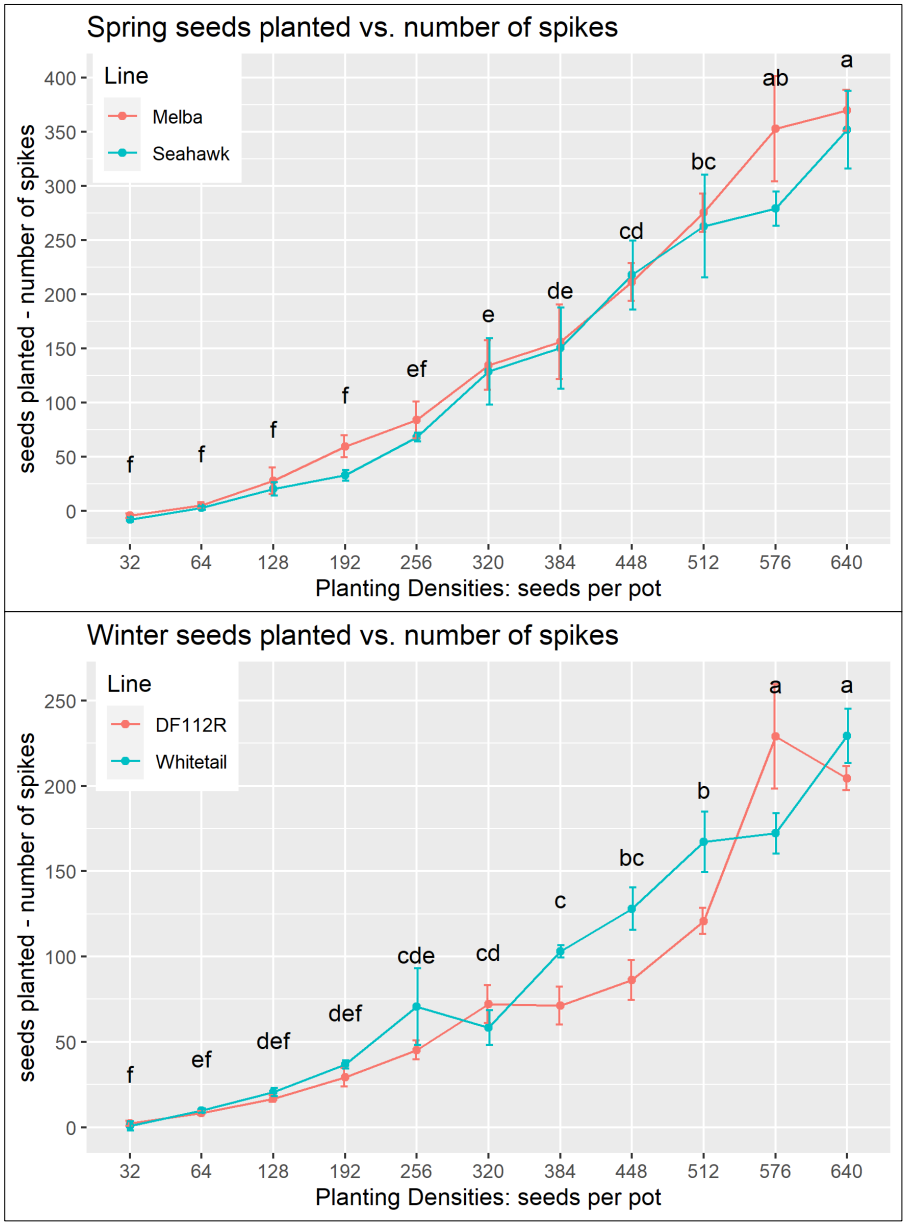


Figure 3.12: Seeds planted vs number of spikes for spring wheat and winter wheat lines for planting densities 32 to 640 seeds per pot. Letters are LSD values for best fitted model. Bars are standard error.

Both spring wheat and winter seeds planted vs number of spikes increased as planting density increased. In spring wheat, seeds planted vs number of spikes increased significantly

from 192 to 576 seeds per pot ($P<0.01$). Seeds planted vs number of spikes ranged from -6.37 to 360.75.

In winter wheat, seeds planted vs number of spikes increased significantly from 32 to 576 seeds per pot ($P<0.01$). Seeds planted vs number of spikes ranged from 1.5 to 216.875.

Planting density vs line

Most of the traits evaluated are significantly different ($P<0.05$) between the genotypes used in spring and winter wheat studies (Table 3.2). Adjusted R square values were generated to determine the proportion of variance for each trait that is explained by planting density and line (Table 3.3). For traits that were indicated to be significant for both planting density and line, only three had adjusted R square values that were greater for line than planting density (Table 3.3). The traits were mature spikelet count and average seed weight in spring wheat, and seed length for winter wheat.

Traits that were indicated to be significant for just line had larger adjusted R squared values than planting density. The traits were seed count and seed length for spring wheat, and seed count for winter wheat. However, all adjusted R squared values were less than 0.24.

Spring		
	Planting density	Line
Height	0.53200	-0.01100
Spikelets per spike	0.64200	-0.00990
Mature spikelets per spike	**0.00528**	**0.69960**
Spikes per pot	0.73000	-0.00440
Total seed weight	0.01862	0.03300
Average seed weight	**0.20300**	**0.25600**
Seed count	-0.09160	0.24900
Seed length	0.07550	0.10670
Seed width	0.28700	0.02630
Seeds per spike	0.78570	0.03720
Seeds planted vs. number of spikes	0.86630	-0.00828
Winter		
	Planting density	Line
Height	0.79300	-0.00906
Spikelets per spike	0.67080	0.00660
Mature spikelets per spike	0.83770	-0.01040
Spikes per pot	0.95900	-0.01010
Total seed weight	0.24940	-0.01040
Average seed weight	0.35970	0.05570
Seed count	0.02100	0.05537
Seed length	**0.21800**	**0.24890**
Seed width	0.57180	0.01090
Seeds per spike	0.92880	-0.00456
Seeds planted vs. number of spikes	0.83950	-0.00714

Table 3.3: Adjusted R squared values for planting density and line. Bolded values indicate adjusted R square values that were greater for line than planting density for traits that were significant for both planting density and line. Underlined values indicate traits that were significant for line only and had larger adjusted R squared values than planting density.

Discussion

The traits evaluated for spring and winter planting density studies followed one of three trends: decreasing as planting density increases, increasing as planting density increases, or followed no trend.

Traits that decreased as planting density increased

Height, spikelets per spike, mature spikelets per spike, average seed weight, seed width, and seeds per spike all decreased as planting density increased for both spring and winter wheat. Total seed weight and seed length for winter wheat also decreased as planting density increased. These findings would be useful for breeders or growers if their goal is to maximize one, or several, of these traits. To maximize these traits, such as generating a higher number of seeds per spike, a lower planting density should be used in greenhouse conditions.

Traits that increased as planting density increased

Spikes per pot and seeds planted vs number of spikes increased as planting density increased for spring and winter wheat. As planting density increases, more spikes are produced but less of the seeds planted reach maturity. If the goal of the breeder is to maximize the number of spikes that reach maturity, the breeder must also be prepared to lose more seeds that would have otherwise reached maturity at a lesser planting density. The limiting factor would then be the seed count from the previous generation.

Traits that followed no trend as planting density increased

Total seed count followed no trend as planting density increased for both spring and winter wheat. Total seed weight for spring wheat also had no trend, however this was likely due to the extremely low seed weights of seahawk at 5 and 6 seeds/in^2. The low seed weights for those samples were potentially due to the placement of those pots in the greenhouse. Increasing

or decreasing planting density will not produce significantly more or less seeds. If the goal of a breeder or grower is to maximize the number of seeds produced in greenhouse conditions, then they should use larger pots or increase the number of replications used.

Planting density vs line

Seed count does not significantly change as planting density increases, however seed count changes significantly between genotypes. This finding may have implications for breeding programs that utilizes a modified bulk breeding method to rapidly advance populations in early generations under greenhouse conditions. The wheat breeding program at Michigan State University (MSU) uses this method to rapidly advance populations from the F2 to F4 stage. The lowest seed counts occurred in DF112R, some of which were less than the maximum number of seeds used in MSU's modified bulk breeding system (500 seeds at F3). This system may have to be adjusted for genotypes that produce less seed in greenhouse conditions, such as planting less seed in the F3 generation. These low seed counts occurred in the highest planting densities, 9 and 10 seeds per square inch or 576 and 640 seeds. The low seed counts may not occur often if the planting densities are lower.

Adjusted R squared values were used to determine the proportion of variance for each trait that is explained by planting density and line. This was particularly useful in determining the impact of planting density and line for traits that were significant ($P<0.05$) for both independent variables. Of those traits, only three had adjusted R squared values greater for line than planting density. Average seed weight in spring wheat and seed length in winter wheat had slightly larger adjusted R square values for line than planting density. However, mature spikelets per spike in spring wheat had a far larger adjusted square value for line than planting density. This suggests that mature spikelets per spike in spring wheat may vary depending on the genotype.

Future implications

Although the current use of the information gathered in these experiments are useful to plant breeding, there may be future applications towards vertical farming. Asseng et al. (2020) described how up to 1940 +/- 230 t/ha of grain can be harvested annually in a hectare indoor vertical facility. Using vertical farming for growing grain crops is currently less feasible due to energy inputs but may be a viable practice in the future (Asseng et al., 2020). Additional information regarding the impacts of genotypes and planting density should be gathered in future studies that include more lines in the experimental design. These experiments could have also been more standardized by being conducted in the same greenhouse during the same time of year. Understanding the impacts of planting density on plant architecture and yield will allow breeders and growers to make informed decisions for indoor growing practices.

BIBLIOGRAPHY

BIBLIOGRAPHY

Asseng S, Guarin JR, Raman M, Monje O, Kiss G, Despommier DD, Meggers FM, Gauthier PPG (2020) Wheat yield potential in controlled-environment vertical farms. Proc Natl Acad Sci USA **117**: 19131

Black CA, Nelson LB, Pritchett WL (1947) Nitrogen Utilization by Wheat as Affected by Rate of Fertilization. Soil Science Society of America Journal **11**: 393–396

Caffaro MM, Vivanco JM, Botto J, Rubio G (2013) Root architecture of Arabidopsis is affected by competition with neighbouring plants. Plant Growth Regulation **70**: 141–147

Cisar G, Howey AE, Brown CM (1982) Optimal Population Density and Random vs. Selective Elimination of Genotypes Under a Modified Single-Seed Descent Method with Spring Oats1. Crop Science **22**: cropsci1982.0011183X002200030034x

Dudley SA, File AL (2007) Kin recognition in an annual plant. Biology Letters **3**: 435–438

Gersani M, Brown J s., O'Brien EE, Maina GM, Abramsky Z (2001) Tragedy of the commons as a result of root competition. Journal of Ecology **89**: 660–669

Milla R, Forero DM, Escudero A, Iriondo JM (2009) Growing with siblings: a common ground for cooperation or for fiercer competition among plants? Proc Biol Sci **276**: 2531–2540

Nord EA, Zhang C, Lynch JP (2011) Root responses to neighbouring plants in common bean are mediated by nutrient concentration rather than self/non-self recognition. Functional Plant Biol **38**: 941

Olson EL, Noble A, Forsberg L, Brisco-McCann E (2021) The 'Minibulk' system for advancing winter cereal breeding programs. Manuscript in preparation.

Zhu Y-H, Weiner J, Li F-M (2019) Root proliferation in response to neighbouring roots in wheat (Triticum aestivum). Basic and Applied Ecology **39**: 10–14