2022

Genome-Wide Association Mapping and Genomic Prediction for Enhancing FHB Resistance in Hard Winter Wheat

Jinfeng Zhang
South Dakota State University, jinfeng.zhang@jacks.sdstate.edu

GENOME-WIDE ASSOCIATION MAPPING AND GENOMIC PREDICTION FOR ENHANCING FHB RESISTANCE IN HARD WINTER WHEAT

BY

JINFENG ZHANG

A dissertation submitted in partial fulfillment of the requirements for the

Doctor of Philosophy

Major in Plant Science

South Dakota State University

2022

DISSERTATION ACCEPTANCE PAGE

Jinfeng Zhang

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Sunish Sehgal
Advisor Date

David Wright
Department Head Date

Nicole Lounsbery, PhD
Director, Graduate School Date

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Dr. Sunish Sehgal, for providing this great opportunity for me to pursue an advanced degree and expand my horizon and knowledge in wheat breeding, and I also appreciate him for his valuable guidance, inspiration, and numerous supports. I am very glad to have the opportunity to work with him in the field and discussed related topics on wheat and breeding as well as other points, we enjoyed the glorious time and I learned a lot from him. Dr. Sehgal's hard work, patience, and dedication impressed and inspired me and have made a difference in my career.

I also would like to thank all my advisory committee members Dr. Ali Shaukat, Dr. Brent Turnipseed, and Graduate School Representative Dr. Yajun Wu for their advice and dedication of time and effort to review the manuscripts and dissertation. Particularly, I would like to express my special thanks to Dr. Shaukat Ali for his help and advice with the experimental materials preparation, phenotyping, and data collection. In addition, I would also like to thank Dr. Jose Gonzalez for his advice on wheat scab phenotyping techniques. I consider myself fortunate to be advised by such great mentors.

I can never forget my lab mates: Jyotirmoy Halder, Harsimardeep S. Gill, and Rami Altameemi. Without their help and encouragement, I could not have finished my program successfully. Thanks to Cody Hall for his help in harvesting and planting the materials. I am grateful to Navreet Brar for her hard work and help in field inoculation and data collection. I would also like to thank US Wheat & Barley Scab Initiative (USWBSI) and South Dakota Wheat Commission for providing the funding for my study and the South

Dakota State University Agriculture Experimental Station who provided the field plots for the experiments.

Finally, I would like to thank my family, especially my wife and my son for their love, understanding, and sacrifice, without their support and encouragement I could not have imagined accomplishing my study and achieving my dream successfully. And I also appreciate my parents for their unconditional love and encouragement to pursue my Ph.D. degree. Lastly thanks to many other people who cared and help me in my studies and research program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY TABLES

Chapter 3:

Chapter 4:

LIST OF SUPPLEMENTARY FIGURES

Chapter 3:

Chapter 4:

ABSTRACT

GENOME-WIDE ASSOCIATION MAPPING AND GENOMIC PREDICTION FOR ENHANCING FHB RESISTANCE IN HARD WINTER WHEAT

JINFENG ZHANG

2022

Wheat is one of the most important staple crops providing 20% of energy for 35% of the world population. Fusarium head blight (FHB), primarily caused by the fugal *Fusarium graminearum* Schwabe, is a damaging disease in wheat that affects global wheat production every year and causes food safety issues. The disease not only reduces the grain yield and quality but also produces mycotoxin in the diseased kernels making them unsuitable for human consumption or as livestock feeds. Breeding FHB resistant cultivar is the most effective and economical approach to managing the disease. This study combines genome-wide association study (GWAS) and genomic approaches (GS) to identify resistance loci/markers and evaluate the efficiency of genomic prediction (GP) in hard winter wheat breeding lines in the South Dakota State University (SDSU) winter wheat breeding program. In the first study, we conducted a multi-locus genome-wide association study (ML-GWAS) with 9,321 high-quality single nucleotide polymorphisms (SNPs) and a panel of 257 elite breeding lines from the South Dakota State University (SDSU) breeding program to uncover the genetic basis of native FHB resistance in the US hard winter wheat. Marker-trait associations (MTAs) were identified with eight different ML-GWAS models, the most appropriate being Fixed and random model

Circulating Probability Unification (FarmCPU) for FHB disease index (DIS) and Fusarium damaged kernels (FDK). A total of six distinct quantitative trait nucleotides (QTNs) were identified for DIS on five different chromosomes 2A, 2B, 3B, 4B, and 7A, where five were considered 'reliable QTNs' as those were identified by multiple models. For FDK, a total of eight unique QTNs were identified on six different chromosomes 3B, 5A, 6B, 6D, 7A, and 7B, where four QTNs were considered reliable. In the second study, we further evaluated the genomic prediction potential of advanced breeding lines in predicting FHB disease index (DIS), and the percentage of Fusarium damaged kernels (FDK) in early generation breeding lines. Advanced breeding lines evaluated in 2018, 2019, and 2020 were used as training populations (TP18, TP19, and TP20, respectively) for genomic prediction (GP) of FHB traits. We observed moderate prediction accuracy (PA) using univariate models for DIS (0.39 and 0.35) and FDK (0.35 and 0.37) using TP19 and TP20, respectively, and slightly higher PA (0.41 for DIS and 0.38 for FDK) when TP19 and TP20 (TP19+20) were combined to leverage the advantage of a large training population. However, GP with a multivariate approach including plant height and days to heading as covariates further did not significantly improve PA for DIS and FDK over univariate models, PA for DON increased by 20% using DIS, FDK, and DTH as covariates using the multi-trait model in 2020. Finally, we used TP19, TP20, and TP19+20 in forward prediction to calculate genomic-estimated breeding values (GEBVs) for DIS and FDK in preliminary breeding lines at an early stage of the breeding program. We observed moderate PA of up to 0.59 for DIS and 0.54 for FDK, demonstrating the promise in genomic prediction for FHB resistance in earlier generations using advanced

lines. Our results demonstrate the potential of integrating genomic selection in hard winter wheat breeding to improve FHB resistance.

Key words: Triticum aestivum, FHB, germplasm development, Phenotypic Selection, GBS, GWAS, Genomic Selection, GP, FHB resistance

Chapter 1. Introduction

Wheat is one of the oldest and most widely grown food crops that is consumed worldwide (http://www.fao.ogr/faostat/). Further wheat leads in world trade among all crops and its annual production rank second behind maize, outnumbering other crops including rice, soybean, and potatoes (Lev-Yadun et al., 2002). However, wheat production has been challenged by the increasing abiotic and biotic stresses. *Fusarium* head blight (FHB), primarily caused by fungi *Fusarium* species, stands out as the most damaging wheat disease across the world. The lack of FHB-resistance in wheat cultivars leads to significant yield loss and grain quality degradation in an epidemic year and results in food safety concerns with mycotoxins produced by the pathogen in diseased grains (Ma et al., 2020).

Applying fungicides to control FHB is effective but remains challenging because of the very short application window, potential environmental contamination, and increased cost of wheat production (Bai et al., 2018; McMullen et al., 2012). Therefore, growing resistant cultivars is the most economic and effective approach to minimizing the damage caused by the FHB (Bai & Shaner, 2004; McMullen et al., 2012). The most stable and often employed resistant germplasms are of Asia origin. However, these germplasms are usually either un-adapted to growing conditions in most of North America due to late maturity, variation in vernalization requirement, or have poor agronomics such as tall plant heights with a tendency to lodge, and low grain quality and low yield (Somers et al., 2003). Therefore, recent efforts have intensified to select and exploit native sources of resistance in locally adapted cultivars without negatively affecting other traits (Clinesmith et al., 2019; Hashimi, 2019; Jin et al., 2013; Thambugala et al., 2020). The

major native sources for FHB resistance in hard winter wheat have been cultivars Overland, Everest, Lyman, and Emerson, however, these cultivars provide only moderate control to FHB and have to be combined with fungicides under the epidemic condition for effective management of FHB.

FHB resistance is a typical quantitative trait that is controlled by multiple quantitative trait loci (QTLs), especially for native resistance that is often governed by multiple minor effect QTLs. Due to the nature of QTL, the improvement of FHB resistance is challenging. Conventional phenotypic selection for FHB resistance is still the primary approach for the development of FHB resistant varieties in hard winter wheat breeding programs. However, FHB resistance is often influenced by environmental conditions (Buerstmayr et al., 2012; Miedaner et al., 2001) thus phenotypic selection is unstable and has low efficiency. With the advance in molecular marker technology and genomic tools, molecular marker-assisted selection (MAS) and genomic selection (GS) provides opportunities to improve the efficiency of phenotypic selection for FHB resistance breeding.

Genome-Wide Association Studies (GWAS) is a mapping strategy that identifies marker-trait associations in a panel assembled using a diverse set of accessions/individuals and has been used extensively in human and animal genetic analysis where large segregating populations are not available (Varshney et al., 2007). GWAS has several advantages over linkage mapping including the potential for increased QTL resolution and without the requirement of developing a segregation population (Arruda et al., 2016). GWAS has been used to identify markers/QTLs for biotic and abiotic stress resistance and agronomic

traits that can be used in MAS to develop disease-resistant genotypes in breeding programs (Arora et al., 2017). Further, genomic selection (GS) is another alternative to conventional phenotypic selection. GS captures the total additive genetic variance using genome-wide molecular markers and predicts genomic estimated breeding values (GEBV) in an un-phenotyped population for difficult quantitative traits using molecular markers (Poland & Rutkoski, 2016), thus bypassing the identification of QTL and measurement of its effect. With the genotypic information, GS has made it possible to predict plants' performance at the early stage in the breeding program that can help advance superior lines to the next generations based on their GEBVs. Thus the prediction of phenotypes can substitute the phenotype-dependent laborious field evaluation and effort and investment in assessment for phenotypes can be substantially reallocated to enhance the efficiency of the breeding program (Kuti et al., 2012).

Thus the goal of my study was to implement advanced techniques in the South Dakota winter wheat breeding program to select hard winter wheat germplasm with enhanced FHB resistance, the specific objectives of this study were to:

To characterize the genomic regions associated with FHB resistance in hard winter wheat breeding lines using genome-wide association analysis.

To optimize genomic selection/prediction for FHB resistance in the hard winter wheat breeding program.

Reference

http://www.fao.ogr/faostat/

Arora, S., Singh, N., Kaur, S., Bains, N. S., Uauy, C., Poland, J., & Chhuneja, P. (2017). Genome-wide association study of grain architecture in wild wheat Aegilops tauschii. Frontiers in Plant Science, 8, 886.

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., Foresman, B. J., & Kolb, F. L. (2016). Genome-Wide Association Mapping of Fusarium Head Blight Resistance in Wheat using Genotyping-by-Sequencing. The Plant Genome, 9(1). https://doi.org/10.3835/plantgenome2015.04.0028

Bai, G., & Shaner, G. (2004). Management and resistance in wheat and barley to fusarium head blight. Annual Review of Phytopathology, 42, 135–161. https://doi.org/10.1146/annurev.phyto.42.040803.140340

Bai, G., Su, Z., & Cai, J. (2018). Wheat resistance to Fusarium head blight. Canadian Journal of Plant Pathology, 40(3), 336–346. https://doi.org/10.1080/07060661.2018.1476411

Buerstmayr, M., Huber, K., Heckmann, J., Steiner, B., Nelson, J. C., & Buerstmayr, H. (2012). Mapping of QTL for Fusarium head blight resistance and morphological and developmental traits in three backcross populations derived from Triticum dicoccum× Triticum durum. Theoretical and Applied Genetics, 125(8), 1751–1765.

Clinesmith, M. A., Fritz, A. K., da Silva, C. L., Bockus, W. W., Poland, J. A., Dowell, F. E., & Peiris, K. H. S. (2019). QTL mapping of fusarium head blight resistance in

winter wheat cultivars ‘art’ and ‘everest.’ Crop Science, 59(3), 911–924. https://doi.org/10.2135/cropsci2018.04.0276

Hashimi, A. (2019). Mapping Quantitative Trait Loci for Fusarium Head Blight Resistance in US Winter Wheat. Kansas State University.

Jin, F., Zhang, D., Bockus, W., Baenziger, P. S., Carver, B., & Bai, G. (2013). Fusarium head blight resistance in U.S. Winter wheat cultivars and elite breeding lines. Crop Science, 53(5), 2006–2013. https://doi.org/10.2135/cropsci2012.09.0531

Kuti, C., Láng, L., Gulyás, G., Karsai, I., Mészáros, K., Vida, G., & Bedő, Z. (2012). Bioinformatics tool for handling molecular data in wheat breeding. Cereal Research Communications, 40(4), 573–582.

Lev-Yadun, S., Abbo, S., & Doebley, J. (2002). Wheat, rye, and barley on the cob? Nature Biotechnology, 20(4), 337–338.

Ma, Z., Xie, Q., Li, G., Jia, H., Zhou, J., Kong, Z., Li, N., & Yuan, Y. (2020). Germplasms, genetics and genomics for better control of disastrous wheat Fusarium head blight. Theoretical and Applied Genetics, 133(5), 1541–1568. https://doi.org/10.1007/s00122-019-03525-8

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., & Van Sanford, D. (2012). A unified effort to fight an enemy of wheat and barley: Fusarium head blight. Plant Disease, 96(12), 1712–1728. https://doi.org/10.1094/PDIS-03-12-0291-FE

Miedaner, Reinbrecht, Lauber, Schollenberger, & Geiger. (2001). Effects of genotype and genotype-environment interaction on deoxynivalenol accumulation and resistance to Fusarium head blight in rye, triticale, and wheat. Plant Breeding, 120(2), 97–105. https://doi.org/10.1046/j.1439-0523.2001.00580.x

Poland, J., & Rutkoski, J. (2016). Advances and Challenges in Genomic Selection for Disease Resistance. Annual Review of Phytopathology, 54, 79–98. https://doi.org/10.1146/annurev-phyto-080615-100056

Somers, D. J., Fedak, G., & Savard, M. (2003). Molecular mapping of novel genes controlling Fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. Genome, 46(4), 555–564.

Thambugala, D., Brûlé-Babel, A. L., Blackwell, B. A., Fedak, G., Foster, A. J., MacEachern, D., Gilbert, J., Henriquez, M. A., Martin, R. A., McCallum, B. D., Spaner, D., Iqbal, M., Pozniak, C. J., N'Diaye, A., & McCartney, C. A. (2020). Genetic analyses of native Fusarium head blight resistance in two spring wheat populations identifies QTL near the B1, Ppd-D1, Rht-1, Vrn-1, Fhb1, Fhb2, and Fhb5 loci. Theoretical and Applied Genetics, 133(10), 2775–2796. https://doi.org/10.1007/s00122-020-03631-y

Varshney, R. K., Langridge, P., & Graner, A. (2007). Application of genomics to molecular breeding of wheat and barley. Advances in Genetics, 58, 121–155.

Chapter 2. Literature review

2.1 Wheat production

Wheat is a staple food crop and is consumed worldwide. It is the most widely grown crop (216 million hectares worldwide in 2020), outnumbering any other food crops in the world (https://apps.fas.usda.gov/psdonline/circulars/production.pdf). Wheat can be grown in a wide range of environments and conditions, including cool, hot, dry and wet areas (Dupont & Altenbach, 2003). Its annual production of over 700 million tons, ranking second amongst all crops behind corn, provides about 20% of the calories and protein for over 2.5 billion people in the world (Ma et al., 2020). Wheat grain is also an excellent energy source for farm animals and up to 16.7% of worldwide wheat production has been used as animal feed (Aviles, 2019). Broad adaptability, high yield potential, abundant nutrition, the simplicity of growing, and the ease of handling, transport, storage, and processing, have contributed to its widely cultivation and sustained usage throughout history (Kislev, 1984). Mainly, wheat production distributes in five major regions of the world: Asia, Europe, North America, South America, eastern Africa, and Australia.

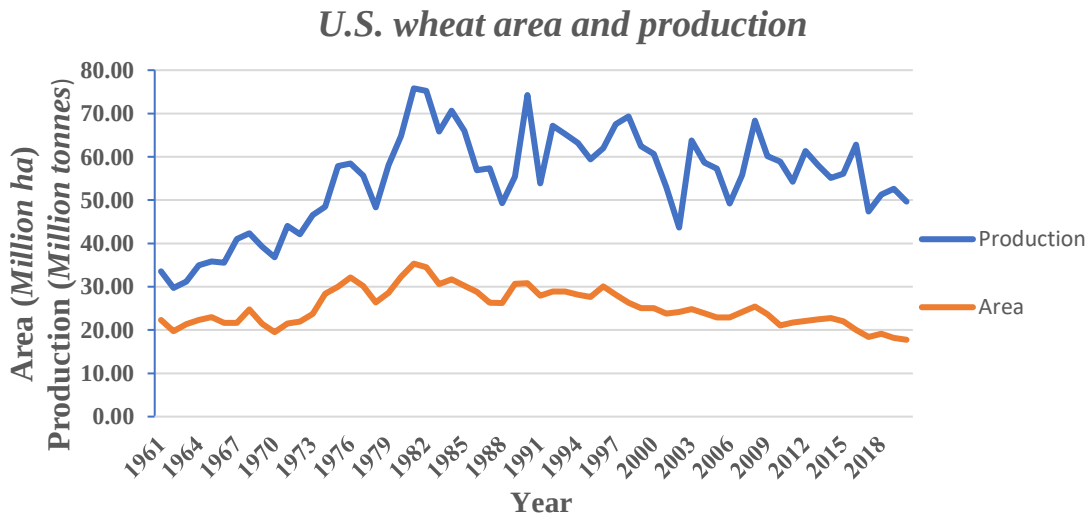


Figure 2.1 U.S. wheat production and area from 1961-2020. Data source: USDA

In the United States (U.S.), wheat ranks the third largest crop in planted acreage and production behind corn and soybean. Wheat planting in the U.S. has a long-term downward trend after peaking in1981 (https://www.ers.usda.gov/topics/crops/wheat/). The planting area reached its top peak of 35 million hectares (Mha) in 1981 with the production of 75.8 million tons then keeping downward especially in the planted area until now (Figure 2.1). Many challenges contribute to the change. International competition might be one of the major reasons for decreasing in U.S. wheat production. Due to foreign competition in the global wheat market, farmers' profitability for planting wheat in the U.S. has declined relative to other crops, which has encouraged some farmers to reduce wheat planting. In addition, the flexible policy that allows farmers to choose the crops for growing by eliminating the requirement of maintaining the base acreage of a crop also encourages the swinging of wheat planting and production. Besides, the fact that genetic improvement has been slower for wheat due to the crop's significantly more complex genetics and lower potential returns from research investments may be technically an additional reason. Although it has been challenged by many factors, the U.S. wheat production ranks the fifth after European Union, China, India, and Russia (https://apps.fas.usda.gov/psdonline/circulars/production.pdf).

2.2 Wheat classes

Wheat can be classified into different categories based on various criteria. Wheat (Triticum spp.) in fact includes several species, the economically important is common wheat (*T. aestivum* L. 2N=6X=42) with more than 90% of planting area, comprising 95%

of global wheat production, and durum (*Triticum durum* L. 2N = 4x = 28) with around 6-7% of sowing acreage (Márta Molnár-Láng, Carla Ceoloni, 2015). Common wheat can be further separated into spring and winter wheat regarding the ecotypes. Winter wheat typically require a relatively cold vernalization process. Vernalization is a period of low temperatures necessary in order to induce flowering and complete the life cycle in certain varieties. Winter wheat varieties are sown in the fall and get established before cold weather arrives and complete vernalization during winter. It is necessary to expose to a low temperature near 5°C for 2 to 10 weeks for vernalization depending on different varieties. Based on the difference in duration of vernalization requirements, winter wheat cultivars are categorized into three types: a weak winter type which requires brief exposure to low temperature, a semi-winter type that requires 2-4 weeks of cold exposure to induce flowering, and a strong winter type that needs more than 4 weeks of cold exposure (Crofts, 1989). Studies have shown that there are 4 groups of *Vrn* genes (*Vrn1-Vrn4*) associated with the vernalization response (Crofts, 1989; Guedira et al., 2014). Vernalization is considered an evolutionarily adaptive mechanism for winter wheat to delay flowering for survival in the harsh winter climate. The variation of winter and spring ecotype enable wheat to adapt to a wide range of growing environment from tropical zone to within the Arctic cycle (Guedira et al., 2014).

Wheat can also be categorized depending on the seed color, texture and growing habits. In the U.S., wheat varities grown are divided into six classes: hard red winter wheat (HRWW), hard red spring wheat (HRSW), soft red winter wheat (SRWW), soft white, hard white wheat, and durum (https://www.ers.usda.gov/topics/crops/wheat/). Compared with red color grain, white wheat generally has a higher milling yield and produces whole

wheat flour without the needs for discoloration. The differentiation of grain texture (soft or hard) results in a very pronounced difference in the uses of wheat. Hard wheat contains a higher level of gluten and protein than soft wheat (https://flour.com/types/). The high level of gluten, an elastic substance, is capable of retaining gas in the dough, thus causing the baked product to expand or rise to form a stronger structure, thus suitable for bread baking. On the other hand, soft wheat (with relatively less protein) is usually used for products requiring minimal structure, such as cakes, crackers and piecrusts. Wheat varieties grown in dry climates are generally hard types, while the wheat types in humid areas are softer, with weak gluten. HRWW in the U.S. accounts for about 40 percent of total production and is usually grown in the Great Plains (Colorado, Kansas, Oklahoma, Nebraska, Texas, South Dakota, and Montana), whereas the HRSW is mainly grown in Northern Plains (South and North Dakota, Montana, Minnesota), contributing about 25 percent of production. Soft red winter (SRW) wheat accounts for about 15 percent of total production and is grown primarily in States along the Mississippi River and in eastern States (https://www.ers.usda.gov/topics/crops/wheat/wheat-sector-at-a-glance/#classes).

2.3 Wheat breeding

Since the “Green Revolution” in the 20th century, wheat has experienced a spectacular yield increase in total production from 303 million tons in 1966 to 766 million tons in 2019 (Paux, 2012). From the early 1960s, there has been little increase in the area sown to wheat, but over the same period, yields have increased almost 3-fold (Marshall et al., 2001) (Figure 2.2). Much of the yield increase has been through improved agricultural

practice, largely due to the release of new improved varieties (Marshall et al., 2001). Therefore, the breeding of new varieties has been crucial.

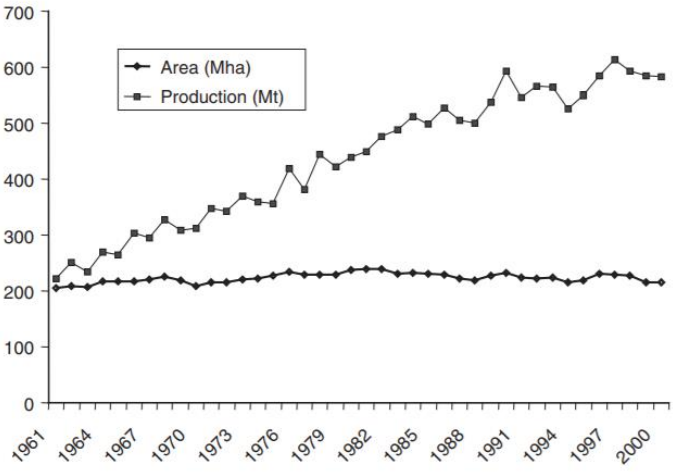


Figure 2.2 Wheat production worldwide compared with the area from 1961 to 2000. (From Marshall et al., 2001)

However, growth in wheat yields has stagnated at around 0.9% per year over the past decade — by contrast, maize (corn) yields grow by almost double that at approximately 1.6% per year. Compared with corn, wheat gets less investment. The total global spend on wheat breeding and research (around US$500 million per year) is one-quarter of that spent to improve maize (anonymity, 2014). One reason for this discrepancy is that wheat is a self-pollinating crop thus farmers can replant seeds from several successive harvests and seed companies get no such annual income as corn, therefore they do not have much motivation to increase the investment in wheat breeding.

Wheat breeding relies long termly on conventional approaches and mostly on “classical” phenotypic selection. Currently, breeding for improved varieties still relies on a 10-year cycle (Table 2.1). From the selection of crossing parent and do cross at the first year,

flowing F2, F3, head row, early observation trial (EOT), Preliminary yield trial (PYT), Advanced yield trial (AYT), elite yield trial (EYT), and Crop performance trial (CPT), each trial needs one year. The cycle repeats year by year. And different trials require the tests in different locations in candidate planting areas.

Table 2.1: Wheat breeding pipeline and cycles

Year	**Gen.**	**Entry**	**Cycle**			
			1	2	3	4
0	P	Parent selection	P			
1	F1	Crossing (~700)	F1	P		
2	F2	F2 (~600)	F2	F1	P	
3	F3	F3 (~350)	F3	F2	F1	P
4	F4	Head row (15000)	F4	F3	F2	F1
5	F5	EOT*(~2000)	F5	F4	F3	F2
6	F6	PYT*(~700)	F6	F5	F4	F3
7	F7	AYT*(126）	F7	F6	F5	F4
8	F8	EYT*(36)	F8	F7	F6	F5
9	F9	CPT*(10)	F9	F8	F7	F6
10	F10	…	…	…	…	…

*EOT: Early observation trial; PYT: Preliminary yield trial; AYT: Advanced yield trial; EYT: Elite yield trial; CPT: Crop performance trial; Number in the bracket is the combination number or lines number.

On the other hand, with the advance in molecular marker technology and genomic technology, high throughput SNP markers can be available and affordable, and marker-assisted selection (MAS) can improve the efficiency of breeding by permitting the selection of target traits or pyramiding various effective genes using associated or linked

molecular markers. Genomic selection (GS) provides an alternative option to improve the efficiency of MAS by capturing small-effect loci that might be missed in MAS and is considered a promising method to increase the selection accuracy and breeding efficiency for the complex traits in wheat breeding (Poland & Rutkoski, 2016).

2.4 *Fusarium* Head Blight in wheat

Fusarium head blight (FHB) or scab is a fungal disease that occurs on wheat, barley, oats, and other small-grain crops and corn. The disease mainly affects heads and grains, leading to bleached or pink color spikelets in premature heads. Diseased grains appear discolored or pinkish, shrived, and are lightweight. These grains sometimes are called "tombstones" because of their chalky, lifeless appearance. FHB causes not only considerable yield losses but also quality concern that is associated with damaged kernels and mycotoxins accumulated in the infected kernels (Bai & Shaner, 2004).

Grains contaminated with mycotoxins, particularly deoxynivalenol (DON), can be a serious safety issue for the health of human and domestic animals for consumption (Ferrigo et al., 2016; Pestka, 2010). 1-3 ppm of DON can cause reduced feed intake and lower weight gain in animals, especially in swine. Vomiting and feed refusal may occur when levels of intaking DON exceed 10 ppm (https://extension.psu.edu/downloadable/download/sample/sample_id/180/). Therefore, the strict upper limits on levels of DON allowed in food and feed have been set in many countries. For example, the maximum DON content in European Union (EU) is 1.25 ppm in unprocessed bread wheat, 0.5 ppm in bread and bakeries, and 0.2 ppm in baby food (Anonymous 2005). In 1993, the Food and Drug Administration (FDA) of the USA

announced the acceptable levels of 1 ppm DON on finished wheat products (http://www.cfsan.fda.gov/~dms/graingui.html). Health Canada has set limits for DON in soft wheat of 2 and 1 ppm for non-staple foods and baby foods, respectively (http://www.hc-sc.gc.ca/fn-an/securit/chem-chim/contaminants-guidelines-directives_e.html).

Epidemics of FHB have had a significantly deleterious impact on global wheat production. In the USA, after firstly being reported by Arthur (1891) in Indiana, FHB epidemics have spread to all major wheat-growing states, including North Dakota, South Dakota, Minnesota, Ohio, Michigan, Missouri, Kansas, and Arkansas (McMullen et al., 2012). Recently, the frequency of FHB epidemics has increased because of the change in climate conditions (frequent summer rains) and farming practices change (less or no tillage and the increased acreages of corn providing rich sources of inoculum) (Bai et al., 2018). Epidemics of FHB result in severe losses through directly reducing the grain yield and increasing grain cleaning costs. Estimated losses to growers, grain handlers and industries that utilize wheat-related products exceeded $ 1 billion in North Dakota, Minnesota, and South Dakota during 1993 alone (Mcmullen et al., 1997).

Several *Fusarium* species such as *F. graminearum*, *F. avenaceum*, *F. culmorum* and *F.poae*. can cause FHB, but predominated species for FHB is *F. graminearum* (sexual stage – *Gibberella zeae*) in the U.S. (H. Buerstmayr et al., 2009). In warm (75 to 85 F) and humid environments, flowering spikelets can be easily infected. For effective infection, warm and moist conditions need to prolong to 72 hours. However, in cooler temperatures when high humidity persists for longer than 72 hours, the infection does also occur (Friskop, 2018). Because FHB development depends on favorable

environmental conditions, disease occurrence and severity may vary depending on the years. A combination of factors that may lead to an FHB outbreak are: abundant inoculum, repeated periods of rain and high humidity during flowering through kernel development, and the use of susceptible cultivars.

2.4.1 Life cycle of Fusarium species

Fusarium species have a complex life cycle including sexual and asexual stages. And basically, there are two types of spores produced for *Fusarium* species. In the asexual stage, conidia are produced, whereas in the sexual stage ascospores are produced (Figure 2.3). Conidia are typically transported by raindrops. Because these conidia are usually entrapped in raindrops, they can not be easily delivered by wind. On the other hand, ascospores are light and can be easily picked up by the wind and delivered over long distances. Studies have reported that spores of *G. zeae* (Ascospores) can be transported not only significant vertical distances from 50m to 1 kilometer in the air but also significant horizontal distances nearly 3 km over the surface of the earth (Maldonado-Ramirez et al., 2005; SA Isard, 2001). Wheat crops are susceptible to infection from the flowering period up to the hard dough stage of kernel development but are most vulnerable at flowering. During moist weather, spores of the fungi are windblown or splashed onto the flowering spikelets, infecting susceptible wheat heads. After the colonization of the wheat heads with *Fusarium*, mycelia spread to other spikelets through vascular bundles of the rachilla in certain susceptible varieties under wet conditions (Bai & Shaner, 2004). Infected florets often fail to produce grain, or the grain is poorly filled and shriveled. The fungus starts to produce and accumulate DON as soon as the infection occurs, which helps *F. graminearum* enter wheat florets (Hernandez Nopsa, 2010). After

harvesting the crops, the mycelium of *Fusarium* species overwinters in crop residues including wheat, barley, corn, and rice, or on the grasses or in the soil. FHB has commonly one cycle per year. However, early infections may produce air-borne spores, which can incite the secondary spread of the disease, especially if the crop has uneven flowering due to late tillers (Figure 2.3).

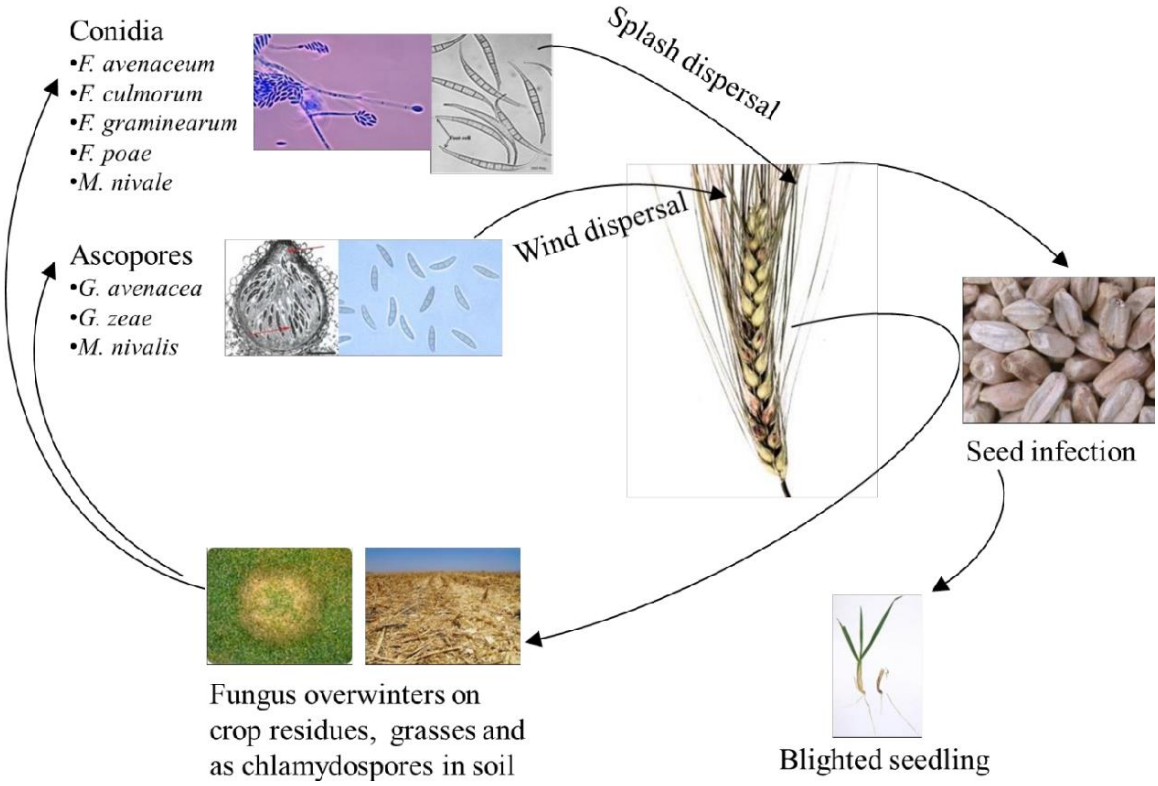


Figure 2.3 Life cycle of Fusarium species (From Eeckhout et al., 2013)

2.4.2 Management of FHB

FHB management is achieved using integrated pest management practices that combine the use of multiple tools. Using a single strategy often fails when the environment favors disease development. Practical measurements include: applying resistant cultivars, seed treatment, crop rotation, tillage if possible, and using fungicide (Friskop, 2018).

The use of resistant cultivars is the most effective and economical method to reduce the FHB damage (Bai & Shaner, 2004; McMullen et al., 2012). Although none of the available commercial cultivars are completely resistant to Fusarium infection, differences in reaction to FHB do occur. Certain varieties have moderate levels of resistance, and the selection of these less susceptible varieties can have a significant impact on disease severity and grain quality. Additionally, plan to plant several different varieties that vary in flowering date (maturity) or stagger planting time. The variation in flowering date decreases the risk that the entire wheat crop will be at vulnerable growth stages when weather conditions favor disease development.

Crop rotation is an effective practice in reducing FHB levels. The key to success for crop rotation is to plant nonhost crops. Thus, planting the wheat crop in a field that previously was planted with a broadleaf crop is a good option. In some regions, a wheat-corn rotation is practiced with zero or reduced tillage together, which has a high risk of FHB by providing rich sources of inoculum for initial infection (Dill-Macky & Jones, 2000; Vogelgsang et al., 2011). Even in this case, the selected wheat variety with resistance had the greatest effect on FHB, as well as the treatment of maize residues with a field shredder (Vogelgsang et al., 2011). Rotation to a legume crop between corn and wheat crops is a suggestion that will provide time for the residues to break down and the pathogen population to decline.

Applying the fungicide is necessary when the forecasting model suggests a high risk of FHB during the growing season. The most effective fungicides labeled for FHB are triazole and pydiflumetofen, active ingredients containing prothioconazole, metconazole, tebuconazole and pydiflumetofen (Á. Mesterházy et al., 2003; Paul et al., 2008; Saldago

et al., 2018). All of these active ingredients, except pydiflumetofen, are in the sterol biosynthesis inhibitor class of fungicides and are site-specific to inhibit the biosynthesis of ergosterol, which is a component of fungal cell walls, but absent in animal and plant cell membranes (Kuck & Scheinpflug, 1986). Studies showed that applying fungicides containing prothioconazole at the beginning of anthesis greatly suppressed FHB disease (Haidukowski et al., 2012; Willyerd et al., 2012). Applying a fungicide four to seven days after early flowering in wheat (full head in barley) still can suppress FHB.

In addition, seed treatment by using fungicide will not protect against FHB, but it can reduce seedling blight if the scabby seeds are used. If a scabby seed source is used, consider using a fungicide seed treatment and adjusting planting density. And biological control has been explored as an additional strategy to manage FHB. Biocontrol agents can be antibiosis, mycoparasitism, and competition (Legrand et al., 2017).

2.5 Host plant resistance to FHB

Host resistance is the most economically effective and environmentally sustainable approach to managing FHB. The success of breeding programs aiming at improving host resistance is largely dependent on the availability of resistant germplasm, and genetic variation in breeding lines. So far, no complete resistance to FHB in wheat is reported (Ma et al., 2020; Parry et al., 1995), therefore, the current strategy for developing resistant cultivars is to accumulate or pyramid of available resistance genes or major quantitative trait loci (QTL) that confer moderate resistance into one genotype without detriment of agronomic traits. The challenge of developing resistant cultivars is the resistance to FHB in wheat is a quantitative trait that is controlled by a large number of

genes or QTLs with moderate heritability, which is mostly influenced significantly by environmental conditions (M. Buerstmayr et al., 2012; Miedaner et al., 2001).

There are several resistance types proposed: type I (resistance to initial infection), type II (resistance to fungal spread in the spike), type III (resistance to DON accumulation) (A Mesterhazy, 1995; Miller et al., 1985); Morphological and phenological traits such as tallness, absence of awns, anther retention/extrusion, loose spikelet distribution and staggered flowering time could contribute to Type I resistance (A. Mesterházy, 1995). Type II resistance is extensively studied because it is the most stable and easy to be evaluated in greenhouse (Bai & Shaner, 2004). Recent studies showed that the Type II resistance might be attributed to the cell wall thickening of rachis nodes and mycotoxin decomposition (Jansen et al., 2005; X. Li et al., 2017). Mesterházy (1995) proposed Type IV resistance to kernel damage and Type V is the tolerance to yield loss because some cultivars had significant differences in kernel infection rate and yield performance even at almost the same levels of FHB. Type IV resistance can be measured using the percentage of Fusarium damaged kernels (FDK) (Rudd et al., 2001). Recently, type III and type IV resistance received more attention in wheat breeding practice because they are types of resistance that end-use or growers are mostly concerned (Mesterhazy, 2020; Verges et al., 2020). The distinction of resistance type or component is to facilitate the evaluation of FHB resistance and may help to understand the mechanism of FHB resistance in wheat (Venske et al., 2019).

Apart from the complexity of resistance evaluation and the underlying genetic basis of resistance, disease phenotyping is a process of labor and time consuming, and resources requirement. In addition, combining the different types of resistance into one cultivar,

which breeders expected, is another challenge. Therefore, with the advance in molecular and genomic technology, researchers are working to use molecular or genomic methods for assisting the selection and identification of resistance sources. The mapping of QTLs and the use of marker-assisted selection (MAS) provide researchers with tools that assist to identify and combine of major QTLs or resistance genes in a shorter period of time (Snijders, 2004). Applying molecular markers to breed for FHB resistance helps breeders to select target genes/QTL in earlier generations on large sets of populations and make sure not losing some lines with elite FHB resistance (Snijders, 2004). Molecular markers can also be used to identify and pyramid several QTLs in advanced lines.

2.6 Resistance sources and QTLs

In common (hexaploidy) wheat, genetic variation for FHB resistance is large and the resistance sources can be generally categorized as exotic and native germplasm based on their origins (Steiner et al., 2017). Exotic germplasms especially from Chinese sources such as Sumai-3 and its derivative Ning7840 and Wangshuibai have been widely used as resistance donors. From these resistance sources, several major QTLs have been identified and formally named: *Fhb1*, *Fhb2* from Sumai-3, *Fhb4*, and *Fhb5* in Wangshuibai (Bai et al., 1999; Cuthbert et al., 2006, 2007; Waldron et al., 1999; Xue et al., 2010, 2011). *Fhb1*, the well-known and the strongest QTL, has been cloned independently by several groups, even arguably (G. Li, Zhou, Jia, Gao, Fan, Luo, Zhao, Xue, Li, & Yuan, 2019; Rawat et al., 2016; Su et al., 2019). According to Rawat et al. (2016), a gene encoding pore-forming toxin-like (PFT) was responsible for the *Fhb1*-led resistance. However, the conclusion was disputed by other researchers based on the cases in which many lines from a large collection of wheat samples with the functional *PFT*

allele are susceptible. Li et al. (2019) and Su et al. (2019) independently reported that the *his* or *HRC* gene encoding a histidine-rich calcium-binding protein, around 50 kb away from *PFT*, might be the candidate for *Fhb1*. *Fhb1* located on the short arm of chromosome 3B, confers stable mainly type II (resistant to spread) resistance. *Fhb1* has been widely used in wheat breeding programs worldwide, particularly in spring wheat (Bai et al., 2018; H. Buerstmayr et al., 2009; Ma et al., 2020; McMullen et al., 2012). To remove the poor agronomics of linkage drag in *Fhb1* carrying Chinese landraces, recently, the US Department of Agriculture Central Small Grain Genotyping Laboratory has transferred *Fhb1* to 16 locally adapted hard winter wheat cultivars or breeding lines from five hard winter wheat states using marker-assisted backcross. One of the selected lines, 'OverlandFHB-10', has been promoted to the 2017 and further Northern Regional Performance Nursery for yield trails (Bai et al., 2018). Diagnostic markers either gel-based or Kompetitive allele specific PCR (KASP) protocols have been developed and available (Su et al., 2018), which facilitates marker-assisted selection (MAS) in the breeding program.

Fhb2 is also a major QTL identified in Chinese spring wheat germplasm Sumai-3 (Cuthbert et al., 2007). This QTL, on chromosome 6BS, explained a wide range of the phenotypic variation (4.4-23%) for type II and III resistance (low DON accumulation) (Bai et al., 2018; T. Li et al., 2011; Waldron et al., 1999). *Fhb2* QTL is also present in other germplasm from different regions, including Arina and Apache from Europe, Patton from USA, and DH181 from Canada (Bai et al., 2018; Semagn et al., 2007).

Fhb4 is one major-effect QTL first mapped in the Chinese landrace Wangshuibai (Lin et al., 2006), which confers type I resistance. *Fhb4* is located on the long arm of

chromosome 4B, tightly linked to SSR markers *Xgwm149* and *Xgwm6* (Cai et al., 2016; Xue et al., 2010). Several other studies have been repeatedly mapping QTL on the *Fhb4* interval, indicating that this QTL is present at a relatively high frequency in wheat germplasms from different sources (Cai, 2016; Clinesmith et al., 2019; Liu et al., 2009).

Fhb5 is another large-effect QTL found in Wangshuibai and is associated with type I resistance (Xue et al., 2011), but also associated with type II resistance in some populations (Liu et al., 2009). According to Liu et al. (2009), *Fhb5* is present in many populations including the resistance germplasm from China ('Wangshuibai'), Japan ('Nyu Bai'), Europe ('F201R') and America ('Frontana' and 'Ernie'). This QTL is located in the close centromeric region of 5AS flanked by the SSR markers *Xgwm304* and *Xgwm415* (Xue et al., 2011).

Besides, some alien species such as tetraploid or hexaploidy wild relatives of wheat *Ae. ventricosa, Ae. Speltoides, Thinopyrum ponticum, Th. elongatum, Th. intermedium, Dasypyrum villosum, Leymous racemosus*, and *Elymus tsukushiensis*, are also used as resistance sources of FHB (Bai et al., 2018; Cainong et al., 2015; Oliver et al., 2005; Qi et al., 2008). Cytogenetic approaches have been used to transfer resistance genes from these alien sources to generate substitution, translocation or recombinant lines by backcrossing with adapted common wheat varieties. Three alien FHB resistance fragments successfully introgressed into common wheat, are denominated as *Fhb3* (from species *Leymus racemosus* (Qi et al., 2008), *Fhb6* (from species *Elymus tsukushiensis* (Cainong et al., 2015), and *Fhb7* (from *Thinopyrum ponticum* (Guo et al., 2015), respectively, all of which showed a high level of FHB resistance in the wheat background (Bai et al., 2018). The germplasm KS14WGRC61 (containing *Fhb6* in Chinese Spring

background) (TA5660) and TA5093 (*Fhb6* in Kansas winter wheat cultivar Everest background) were released and available (Cainong et al., 2015; Friebe et al., 2013). Recently, *Fhb7*, on chromosome 7D, has been cloned as a glutathione S-transferase (GST) gene, which confers broad resistance to Fusarium species by detoxifying trichothecenes through de-epoxidation. *Fhb7* introgression in wheat confers resistance to both FHB and crown rot in diverse wheat backgrounds without yield penalty, providing germplasm for Fusarium resistance breeding (H. Wang et al., 2020).

In addition, many native germplasms were used as FHB resistance sources. Due to the linkage drag of undesirable traits that comes from *Fhb1*-containing cultivar, Sumai 3, and other exotic sources of resistance, incorporating major-effect QTL into the regional breeding program is not always easy or successful. Many studies pay attention to discovering and utilizing native FHB resistance. Fortunately, there are still some local cultivars showing moderate resistance to FHB in many regions (Aviles, 2019; Bai et al., 2018; Clinesmith et al., 2019). For example, several moderately FHB resistant cultivars without *Fhb1* have been released in US soft winter wheat (SWW) regions, such as ‘Truman’, ‘Massy’, ‘Bess’, ‘Ernie’, ‘Roane’ and ‘Freedom’ (Bai et al., 2018; Rudd et al., 2001). And some regional hard winter wheat (HWW) varieties: ‘Everest’, ‘Overland’, ‘Lyman’, ‘Heyne’, ‘Century’, and ‘Hondo’ have been reported of possessing moderately FHB resistance (Bai et al., 2018; Clinesmith et al., 2019; Jin et al., 2013) and native resistance QTLs have been identified in cultivar Art, Everest and Lyman (Clinesmith et al., 2019; Hashimi, 2019). Resistance QTLs reported include: 2B, 3B, 4B and 5A from Ernie (S. Liu et al., 2007); 2A from Freedom (Gupta et al., 2001); 1A, 2B, 3B, 4D, and 5D from Bess (Petersen, 2015); 1A, 1B, 3A, 4A and 6A from Lyman; 1A, 1B, 5A, and

6A from Overland (Eckard et al., 2015); 2D, 4B, 4D from ‘Art’ (Clinesmith et al., 2019); 4A, 4D, 5B, and 4DL from ‘Overland’(Fatima, 2016). Those QTLs can be easily incorporated into U.S. HWW cultivars due to the good adaptation of these donor parents. Besides, some cultivars from Europe including Arina, Renan, and Dream show moderate resistance to FHB (H. Buerstmayr et al., 2009). A spring wheat variety Frontana from Brazil shows a high-level type I and type II resistance, in which the major QTL on chromosome 3A and 5A explained collectively 25% of the phenotypic variation were reported (Steiner et al., 2004).

2.7 QTL mapping strategy

Basically, there are two strategies for QTL mapping: linkage and association mapping. Genetic mapping by linkage is based on genetic recombination events for a specific trait in a segregation population (F2, Double Haploid, or Recombinant inbred lines), using statistical analysis to locate all possible loci responsible for the trait variation (Gómez et al., 2011). The development of segregating population needs various times or resources depending on population type. An F2 population is easy and quick to produce, but each individual is a single genotype and phenotype and cannot be repeated in different environments. Recombinant inbred lines (RILs) consisting of homozygous lines derived from continuously selfing from F2 require several growing seasons of selfing and can be used repeatedly but are time-consuming. Doubled haploid (DH) population by regenerating plants from haploid tissue through chromosome doubling can significantly reduce the time of population development while requiring specific skills (Chen et al., 2006; Hashimi, 2019). Molecular markers are used to separate the segregation population into different groups depending on the genotype of a specific marker. Then QTL analysis

determines whether there is a significant difference between the genotypic groups based on the phenotypic trait of interest (Paterson, 1996). If there is a significant difference in the phenotypic means between the genotypic groups, indicating that the marker used to partition the population into different groups is linked to a QTL affecting the trait. A test of likelihood is used to determine the linkage between a marker and the QTL (Aviles, 2019). Single marker analysis (SMA), simple interval mapping (SIM), composite interval mapping (CIM), and multiple interval mapping (MIM) can be used for mapping analysis. The statistical methods applied for single-marker analysis include t-tests, analysis of variance (ANOVA), and linear regression. CIM is more precise and commonly utilized because it considers other QTL effects thus reducing the background "noise" (Zeng, 1994).

Although the linkage mapping provides a simplistic and powerful tool for QTL identification, it has obvious shortcomings, including limited genetic variation and the presence of only one or few meiotic generations resulting in low resolution of QTL mapping, and segregation population development is time-consuming and resources requirement sometimes deviated from breeding efforts for cultivar development (Arruda et al., 2016; Eckard et al., 2015). Association mapping, also called genome-wide association study (GWAS), appeared as an alternative to surpass the drawbacks of linkage mapping and is a promising method particularly due to the advance in high throughput SNP genotyping.

2.8 Genome Wide Association Study (GWAS)

GWAS, initially developed for human disease studies in the early 2000s, was generally applied in plant genetics studies (Scherer & Christensen, 2016). GWAS has the advantage of exploiting naturally accumulated recombination events from diverse panels, without the additional cost of time, and resources associated with population development (Korte & Farlow, 2013). Standard GWAS test statistics assume that all samples in the analysis are unrelated and selected from a uniform, random-mating population. Therefore, unbalanced populations and relatedness among individuals can lead to false marker-trait associations, so it is necessary to use statistical models that fit the population structure and the kinship matrix of genetic effects as covariates to reduce the false discovery rate of genetic markers (Yu & Buckler, 2006). Related statistical methods have been developed and tested to minimize these confounding effects while optimizing computing speed. Among them, the mixed linear model (MLM) incorporating both the fixed effects and the random effects to control population structure and family relatedness, is the most popular method used in GWAS and also is a conservative model because of the stringent Bonferroni correction for multiple tests (Wang et al., 2016; Yu & Buckler, 2006).

Several studies have been reported to identify significant marker-trait associations with FHB resistance in wheat using GWAS (Arruda et al., 2016; Kollers et al., 2013; R. Wang et al., 2017; Zhu et al., 2020). Arruda et al. (2016) identified significant QTL on chromosomes 1D, 3B, 4A, 4D, 6A, 7A, and 7D in US Midwest and eastern winter wheat breeding lines using Genotyping by sequencing (GBS) SNP markers. Wang et al. (2017) and Zhu et al. (2020) respectively studied the Spring wheat lines from Pacific Northwest

and CIMMYT and the Chinese Elite wheat lines using 90K SNP array by GWAS, identifying significant SNP trait association on chromosome 1B, 2B, 4B, 5A, 5B, 6A and 1AS, 2DL, 5AS, 5AL, 7DS, correspondingly.

Although with the increasing application, GWAS still has some limitations such as spurious associations and missing some rare variants. MAS based on GWAS also has the problem that any QTL mapping method has, only considering partially the effect of markers associated with QTLs and missing some heritability. Fortunately, there is one alternative technology that utilizes genome-wide predictions to overcome these limitations (Silva, 2018).

2.9 Genomic selection/prediction

Genomic selection (GS) can be considered as an extension form of marker-assisted selection. Meuwissen et al. (2001) developed a method of using genome-wide dense markers to predict the total breeding value of animals or plants population and selection on predicted breeding value could substantially increase the rate of genetic gain in simulated animal and plant breeding populations. The intent was not to detect QTL but to predict the breeding value that counts the total genetic additive variance from all markers across the whole genome for selection, a modified version of marker-assisted selection. With genome-wide markers, every trait locus is likely to be in linkage disequilibrium (LD) with at least one marker in the target population (Dreisigacker et al., 2016). As result, GS is more capable of capturing small-effect loci associated with complex quantitative traits such as FHB and yield (Poland & Rutkoski, 2016).

GS uses a population that has been both genotyped and phenotyped as a ‘training set’ for traits of interest in a target environment to develop a model to predict the phenotypic performance of a genetically related ‘testing set’ that only has been genotyped (Silva, 2018). GS estimates marker effects across the entire genome of the breeding population (BP) based on the prediction model developed with the training population (Xu, 2013). This procedure produces genomic estimated breeding values (GEBV), which can be used as a selection criterion.

2.9.1 Genomic selection model

Many statistical models can be used to calculate GEBV. Here we introduce several classical models. First, rrBLUP and a software package based on R, have been developed primarily for genomic prediction with a mixed linear model that considers random effects for markers and assumes that each marker locus contributes equally to the phenotypic value (each locus explains an equal amount of variance) (Meuwissen et al., 2001; Endelman, 2011). The rrBLUP model was often chosen because it was more efficient (lower computational time and equal or higher accuracies) than other models (Hoffstetter et al., 2016). Bayesian estimation: Bayes A means that all markers have effect variances and are distributed normally, but different markers have different effect variances that follow the scaled inverted chi-square distribution. Bayes B is partly labeled with effect variance, but also with different effect variances that have the inverted chi-square distribution (Meuwissen et al., 2001). In reality, there are many loci with no genetic variance and some of them with genetic variance. Bayes B is a model that is more likely to meet the realistic situation. The Random Forest (RF) is a machine-learning algorithm capable of capturing non-additive effects. Predictions from this model are based on a

multiple decision trees where the tree structure allows the effect of markers to vary depending on other markers' genotypes, indicating capturing nonadditive effects (Poland & Rutkoski, 2016). Although RF sometimes generates high prediction accuracies, it is not clear how much of the additive genetic variation is captured relative to nonadditive genetic information (Rutkoski et al., 2012).

In addition, A Multi-trait (MT) model is developed and implemented to improve the prediction ability for complex traits when other traits correlated to the target trait are available (Jia & Jannink, 2012; Larkin et al., 2020; Schulthess et al., 2018). A few studies showed that the multi-trait model did significantly improve prediction accuracy for complex traits like FHB especially when there was a strong correlation between the covariate traits and the predicted trait and correlated traits had higher heritability than the predicted trait (Larkin et al., 2020; Schulthess et al., 2018). In multiple trait models, correlated traits are as covariates conjoined into mixed linear modes (Jia & Jannink, 2012).

2.9.2 Training population (TP) optimization

The composition of the TP, its size, and its relatedness to the breeding population (BP) are crucial to determining the prediction accuracy of GS (Bassi et al., 2015). Isidro et al. (2015) also described the most important factors that affect prediction accuracy are the TP size, the number of molecular markers of training the model, and the relationship between the training and validating populations.

The size of TP has been investigated in many studies and there is a general agreement that larger TP increases the prediction accuracy (PA) within a special range and further

increases in size do not affect PA again. Asoro et al. (2011) evaluated a maximum of 300 individuals in Oats and found that PA for five different traits increased with the increases in TP size. In wheat, Isidro et al. (2015) evaluated TP size from 25 to 300 individuals and observed that maximum PA for five traits was reached with the larger TP (300). Similar results were found in one evaluation of TP size from 50-350 for grain yield, test weight, heading date, plant height, and powdery mildew resistance, and the highest PA was achieved at TP size from 300 to 350 for five different traits (Sarinelli et al., 2019). However, for complex traits including FHB, increases in TP size do not always improve the PA significantly due to the composition of TP and the relatedness between TP and the validation population (Adeyemo et al., 2020; Verges et al., 2020).

Higher PAs were obtained when closely related individuals were found in TP and VP both in animal and plant breeding (Clark et al., 2012; Herter et al., 2019b; Hoffstetter et al., 2016). In animals, simulated and real data showed that PAs are closely associated to the maximum level of relatedness between the training set and the particular predicted individual (Clark et al., 2012). In wheat, Herter et al. (2019) reported two experimental winter wheat populations derived from six and eight bi-parental families, the prediction between the two half-sib families (with one common parent) resulted in the highest prediction accuracies for all four traits, while prediction accuracies were lowest between unrelated families. The study also suggests that the composition of the training population is of utmost importance in genomic selection for FHB and Septoria tritici blotch resistance (Herter et al., 2019).

2.9.3 Cross prediction

It is certainly that forward prediction of non-phenotyped elite lines using a built GP model from a training set is an important application of GS. The GP tool could be also applied for cross prediction to help predict which patents or combinations would produce the most promising progeny when crossed (Silva, 2018). Given a typical wheat breeding program that includes ~100 elite lines in the crossing list from where around 700 crosses are made every year. This crossing number represents only 14% of the total number of the 4950 possible combinations, indicating that the majority of possible combinations are not tested or missing.

Cross prediction refers to the ability to estimate or predict a cross from a set of elite parents aiming to identify which ones are more likely to generate a superior progeny. Although cross prediction does not guarantee sufficient genetic variance is involved, it could estimate the desired population mean, and simulating study had been proposed to predict the performance of the progeny population (Bernardo, 2014). Following this idea, the related R package ‘PopVar’ that is capable of estimating the genetic variance in simulated populations depending on phenotypic and genotypic data from a list of potential parents was developed (Mohammadi, 2015). Cross-predictions have been applied in wheat crosses aiming for superior grain yield and baking quality in INIA Uruguay and CIMMYT (Lado, 2017), as well as FHB resistance in barley (Mohammadi, 2015).

Generally, although there have been some progress cited in the literature, more validations are needed for cross predictions. In addition, a more diverse and complex cross design should be considered because the majority of cross combinations in wheat

breeding programs is three-way or four-way other than the biparental cross on which the current prediction model focuses (Silva, 2018).

2.10 Reference

https://www.britannica.com/plant/wheat

https://www.ers.usda.gov/topics/crops/wheat/

https://flour.com/types/

http://www.cfsan.fda.gov/~dms/graingui.html

Arthur, J. C. 1891. Wheat scab. Indiana Agric. Exp. Sta. Bull. 36:129-138

Adeyemo, E., Bajgain, P., Conley, E., Sallam, A., & Anderson, J. (2020). Optimizing Training Population Size and Content to Improve Prediction Accuracy of FHB-Related Traits in Wheat. *Agronomy*, *10*(4), 543. https://doi.org/10.3390/agronomy10040543

Ágnes, S. H., Szabolcs, L. K., Mónika, V., László, P., János, P., Csaba, L., & Ákos, M. (2014). Differential influence of QTL linked to Fusarium head blight, Fusarium-damaged kernel, deoxynivalenol contents and associated morphological traits in a Frontana-derived wheat population. *Euphytica*, *200*(1), 9–26. https://doi.org/10.1007/s10681-014-1124-2

Anderson, J. A., Stack, R. W., Liu, S., Waldron, B. L., Fjeld, A. D., Coyne, C., Moreno-Sevilla, B., Fetch, J. M., Song, Q. J., Cregan, P. B., & Frohberg, R. C. (2001). DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theoretical and Applied Genetics*, *102*(8), 1164–1168. https://doi.org/10.1007/s001220000509

anonymity. (2014). Wheat lag. *Nature*, *507*(7493), 399–400. https://doi.org/10.1038/507399b

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., Foresman, B. J., & Kolb, F. L. (2016). Genome-Wide Association Mapping of Fusarium Head Blight Resistance in Wheat using Genotyping-by-Sequencing. *The Plant Genome*, *9*(1). https://doi.org/10.3835/plantgenome2015.04.0028

Asoro, F. G. (n.d.). *Asoro 2011 Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats.pdf* (p. 2011).

Aviles, A. C. (2019). LSU Digital Commons Identification of Quantitative Trait Loci (QTLs) for Resistance to Fusarium Head Blight (FHB) in Wheat Variety AGS 2060 and Evaluation of the Effect of Fhb1 Resistance Gene on FHB Reaction in a Louisiana Wheat Breeding Population R. *PhD DISSERTATION*.

Bai, G., Kolb, F. L., Shaner, G., & Domier, L. L. (1999). Amplified fragment length polymorphism markers linked to a major quantitative trait locus controlling scab resistance in wheat. *Phytopathology*, *89*(4), 343–348. https://doi.org/10.1094/PHYTO.1999.89.4.343

Bai, G., & Shaner, G. (2004). Management and resistance in wheat and barley to fusarium head blight. *Annual Review of Phytopathology*, *42*, 135–161. https://doi.org/10.1146/annurev.phyto.42.040803.140340

Bai, G., Su, Z., & Cai, J. (2018). Wheat resistance to Fusarium head blight. *Canadian Journal of Plant Pathology*, *40*(3), 336–346. https://doi.org/10.1080/07060661.2018.1476411

Bai Guihua, & Shaner, G. (1994). Scab of wheat: Prospects for control. In *Plant Disease* (Vol. 78, Issue 8, pp. 760–766). https://doi.org/10.1094/pd-78-0760

Bassi, F. M., Bentley, A. R., Charmet, G., Ortiz, R., & Crossa, J. (2015). Breeding schemes for the implementation of genomic selection in wheat (Triticum spp.). *Plant Science*, *242*, 23–36. https://doi.org/10.1016/j.plantsci.2015.08.021

Bernardo. (2014). *Bernardo 2014 Genomewide Selection of Parental Inbreds_ Classes of Loci and Virtual Biparental Populations.pdf*.

Berzonsky, W. (2009). *South Dakota Agricultural Experiment Station , Brookings*. *393*, 141–145.

Bourdoncle, W., & Ohm, H. W. (2003). Quantitative trait loci for resistance to Fusarium head blight in recombinant inbred wheat lines from the cross huapei 57-2/Patterson. *Euphytica*, *131*(1), 131–136. https://doi.org/10.1023/A:1023056207513

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse

samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Breseghello, F., & Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. *Genetics*, *172*(2), 1165–1177. https://doi.org/10.1534/genetics.105.044586

Buerstmayr, H., Ban, T., & Anderson, J. A. (2009). QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: A review. *Plant Breeding*, *128*(1), 1–26. https://doi.org/10.1111/j.1439-0523.2008.01550.x

Buerstmayr, Hermann, Steiner, B., Lemmens, M., & Rucjenbauer, P. (2000). Heritability and Trait Associations. *Crop Science*, *40*(August), 662–664.

Buerstmayr, M., & Buerstmayr, H. (2015). Comparative mapping of quantitative trait loci for Fusarium head blight resistance and anther retention in the winter wheat population Capo × Arina. *Theoretical and Applied Genetics*, *128*(8), 1519–1530. https://doi.org/10.1007/s00122-015-2527-8

Buerstmayr, M., & Buerstmayr, H. (2016). The semidwarfing alleles Rht-D1b and Rht-B1b show marked differences in their associations with anther-retention in wheat heads and with fusarium head blight susceptibility. *Phytopathology*, *106*(12), 1544–1552. https://doi.org/10.1094/PHYTO-05-16-0200-R

Buerstmayr, M., Huber, K., Heckmann, J., Steiner, B., Nelson, J. C., & Buerstmayr, H. (2012). Mapping of QTL for Fusarium head blight resistance and morphological and developmental traits in three backcross populations derived from Triticum

dicoccum × Triticum durum. *Theoretical and Applied Genetics*, *125*(8), 1751–1765. https://doi.org/10.1007/s00122-012-1951-2

Cai, J. (2016). *Meta-analysis of QTL for Fusarium head blight resistance in Chinese wheat landraces using genotyping by sequencing*. Kansas State University.

Cai, J., Wang, S., Li, T., Zhang, G., & Bai, G. (2016). Multiple minor QTLs are responsible for Fusarium head blight resistance in Chinese wheat landrace Haiyanzhong. *PLoS ONE*, *11*(9), 1–15. https://doi.org/10.1371/journal.pone.0163292

Cainong, J. C., Bockus, W. W., Feng, Y., Chen, P., Qi, L., Sehgal, S. K., Danilova, T. V., Koo, D. H., Friebe, B., & Gill, B. S. (2015). Chromosome engineering, mapping, and transferring of resistance to Fusarium head blight disease from Elymus tsukushiensis into wheat. *Theoretical and Applied Genetics*, *128*(6), 1019–1027. https://doi.org/10.1007/s00122-015-2485-1

Chen, J., Griffey, C. A., Saghai Maroof, M. A., Stromberg, E. L., Biyashev, R. M., Zhao, W., Chappell, M. R., Pridgen, T. H., Dong, Y., & Zeng, Z. (2006). Validation of two major quantitative trait loci for fusarium head blight resistance in Chinese wheat line W14. *Plant Breeding*, *125*(1), 99–101.

Clark, S. A., Hickey, J. M., Daetwyler, H. D., & van der Werf, J. H. J. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. In *Genetics, selection, evolution : GSE* (Vol. 44, p. 4). https://doi.org/10.1186/1297-9686-44-4

Clinesmith, M. A., Fritz, A. K., da Silva, C. L., Bockus, W. W., Poland, J. A., Dowell, F. E., & Peiris, K. H. S. (2019). QTL mapping of fusarium head blight resistance in winter wheat cultivars 'art' and 'everest.' *Crop Science*, *59*(3), 911–924. https://doi.org/10.2135/cropsci2018.04.0276

Craig, A., Ewan, R., Mesmar, J., Gudipati, V., & Sadanandom, A. (2009). E3 ubiquitin ligases and plant innate immunity. *Journal of Experimental Botany*, *60*(4), 1123–1132. https://doi.org/10.1093/jxb/erp059

Crepieux, S., Lebreton, C., Flament, P., & Charmet, G. (2005). Application of a new IBD-based QTL mapping method to common wheat breeding population: Analysis of kernel hardness and dough strength. *Theoretical and Applied Genetics*, *111*(7), 1409–1419. https://doi.org/10.1007/s00122-005-0073-5

Crofts, H. J. (1989). On defining a winter wheat. *Euphytica*, *44*(3), 225–234. https://doi.org/10.1007/BF00037529

Cuthbert, P. A., Somers, D. J., & Brulé-Babel, A. (2007). Mapping of Fhb2 on chromosome 6BS: a gene controlling Fusarium head blight field resistance in bread wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *114*(3), 429–437.

Cuthbert, P. A., Somers, D. J., Thomas, J., Cloutier, S., & Brulé-Babel, A. (2006). Fine mapping Fhb1, a major gene controlling fusarium head blight resistance in bread wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *112*(8), 1465.

Daudi, A., Cheng, Z., O'Brien, J. A., Mammarella, N., Khan, S., Ausubel, F. M., & Paul Bolwell, G. (2012). The apoplastic oxidative burst peroxidase in Arabidopsis is a

major component of pattern-triggered immunity. *Plant Cell*, *24*(1), 275–287. https://doi.org/10.1105/tpc.111.093039

DeYoung, B. J., & Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*, *7*(12), 1243–1249. https://doi.org/10.1038/ni1410

Dill-Macky, R., & Jones, R. K. (2000). The effect of previous crop residues and tillage on Fusarium head blight of wheat. *Plant Disease*, *84*(1), 71–76.

Dreisigacker, S., Sukumaran, S., Guzmán, C., He, X., Lan, C., Bonnett, D., & Crossa, J. (2016). *Molecular Marker-Based Selection Tools in Spring Bread Wheat Improvement: CIMMYT Experience and Prospects BT - Molecular Breeding for Sustainable Crop Improvement: Volume 2* (V. R. Rajpal, S. R. Rao, & S. N. Raina (eds.); pp. 421–474). Springer International Publishing. https://doi.org/10.1007/978-3-319-27090-6_16

Dupont, F. M., & Altenbach, S. B. (2003). Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *Journal of Cereal Science*, *38*(2), 133–146. https://doi.org/10.1016/S0733-5210(03)00030-4

Eckard, J. T., Gonzalez-Hernandez, J. L., Caffe, M., Berzonsky, W., Bockus, W. W., Marais, G. F., & Baenziger, P. S. (2015). Native Fusarium head blight resistance from winter wheat cultivars ‘Lyman,’ ‘Overland,’ ‘Ernie,’ and ‘Freedom’ mapped and pyramided onto ‘Wesley’-Fhb1 backgrounds. *Molecular Breeding*, *35*(1). https://doi.org/10.1007/s11032-015-0200-1

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome*, *4*(3), 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Fatima, N. (2016). *Identification and deployment of QTL for Fusarium head blight resistance in US hard winter wheat*. Kansas State University.

Ferrigo, D., Raiola, A., & Causin, R. (2016). Fusarium toxins in cereals: Occurrence, legislation, factors promoting the appearance and their management. *Molecules*, *21*(5). https://doi.org/10.3390/molecules21050627

Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., Duijn, C. M. Van, Wijmenga, C., Ommen, G. Van, Slagboom, P. E., Dorret, I., Ye, K., Guryev, V., Arndt, P. F., Kloosterman, W. P., Paul, I., Bakker, W. De, Sunyaev, S. R., Hospital, W., & Biology, C. M. (2016). *HHS Public Access*. *47*(7), 822–826. https://doi.org/10.1038/ng.3292.Genome-wide

Friebe, B., Bockus, W., Chen, P. D., Qi, L. L., Cainong, J., Wilson, D. L., Raupp, W. J., Poland, J., Bowden, R. L., & Fritz, A. K. (2013). Notice of release of KS14WGRC61 Fusarium head blight-resistant wheat germ plasm. *Annu Wheat Newslett*, *59*, 137.

Friskop, A. (2018). *Scab Plant Disease Management NDSU Extension Fusarium Head Blight (Scab) of Small Grains Fusarium Head Blight (Scab) of Small Grains Fusarium Head Blight PP804 (Revised)*. *804*. www.wheatscab.ndsu.edu

Gilbert, J., & Tekauz, A. (2000). Review: Recent developments in research on fusarium head blight of wheat in Canada. *Canadian Journal of Plant Pathology*, *22*(1), 1–8. https://doi.org/10.1080/07060660009501155

Gilbert, Jeannie, & Haber, S. (2013). Overview of some recent research developments in fusarium head blight of wheat. *Canadian Journal of Plant Pathology*, *35*(2), 149–174. https://doi.org/10.1080/07060661.2013.772921

Gómez, G., Álvarez, M. F., & Mosquera, T. (2011). Association mapping, a method to detect quantitative trait loci: statistical bases Mapeo por asociación, un método para la detección de loci de rasgos cuantitativos: bases estadísticas. *Agronomía Colombiana*, *29*(3), 367–376.

Graf, R. J., Beres, B. L., Laroche, A., Gaudet, D. A., Eudes, F., Pandeya, R. S., Badea, A., & Randhawa, H. S. (2013). Emerson hard red winter wheat. *Canadian Journal of Plant Science*, *93*(4), 741–748. https://doi.org/10.4141/CJPS2012-262

Guedira, M., Maloney, P., Xiong, M., Petersen, S., Murphy, J. P., Marshall, D., Johnson, J., Harrison, S., & Brown-Guedira, G. (2014). Vernalization duration requirement in soft winter wheat is associated with variation at the VRN-B1 locus. *Crop Science, 54*(5), 1960–1971. https://doi.org/10.2135/cropsci2013.12.0833

Guo, J., Zhang, X., Hou, Y., Cai, J., Shen, X., Zhou, T., Xu, H., Ohm, H. W., Wang, H., & Li, A. (2015). High-density mapping of the major FHB resistance gene Fhb7 derived from Thinopyrum ponticum and its pyramiding with Fhb1 by marker-assisted selection. *Theoretical and Applied Genetics*, *128*(11), 2301–2316.

Gupta, A., Lipps, P. E., Campbell, K. G., & Sneller, C. H. (2001). Identification of QTL associated with resistance to FHB in Ning 7840 and Freedom. *2001 National Fusarium Head Blight Forum Proceedings*, 180.

Haidukowski, M., Visconti, A., Perrone, G., Vanadia, S., Pancaldi, D., Covarelli, L., Balestrazzi, R., & Pascale, M. (2012). Effect of prothioconazole-based fungicides on Fusarium head blight, grain yield and deoxynivalenol accumulation in wheat under field conditions. *Phytopathologia Mediterranea*, 236–246.

Halder, J., Zhang, J., Ali, S., Sidhu, J. S., Gill, H. S., Talukder, S. K., Kleinjan, J., Turnipseed, B., & Sehgal, S. K. (2019). Mining and genomic characterization of resistance to tan spot, Stagonospora nodorum blotch (SNB), and Fusarium head blight in Watkins core collection of wheat landraces. *BMC Plant Biology*, *19*(1), 1–15. https://doi.org/10.1186/s12870-019-2093-3

Hashimi, A. (2019). *Mapping Quantitative Trait Loci for Fusarium Head Blight Resistance in US Winter Wheat*. Kansas State University.

He, X., Dreisigacker, S., Singh, R. P., & Singh, P. K. (2019). Genetics for low correlation between Fusarium head blight disease and deoxynivalenol (DON) content in a bread wheat mapping population. *Theoretical and Applied Genetics*, *132*(8), 2401–2411. https://doi.org/10.1007/s00122-019-03362-9

Hernandez Nopsa, J. F. (2010). *Fusarium Head Blight: Winter Wheat Cultivar responses and characterization of pathogen isolates*. https://www-proquest-com.libproxy.unl.edu/docview/822234151/fulltextPDF/73937C073EE147EBPQ/1?accountid=8116

Herter, C. P., Ebmeyer, E., Kollers, S., Korzun, V., Würschum, T., & Miedaner, T. (2019a). Accuracy of within- and among-family genomic prediction for Fusarium head blight and Septoria tritici blotch in winter wheat. *Theoretical and Applied Genetics*, *132*(4), 1121–1135. https://doi.org/10.1007/s00122-018-3264-6

Herter, C. P., Ebmeyer, E., Kollers, S., Korzun, V., Würschum, T., & Miedaner, T. (2019b). Accuracy of within- and among-family genomic prediction for Fusarium head blight and Septoria tritici blotch in winter wheat. *Theoretical and Applied Genetics*, *132*(4), 1121–1135. https://doi.org/10.1007/s00122-018-3264-6

Hoffstetter, A., Cabrera, A., Huang, M., & Sneller, C. (2016). Optimizing training population data and validation of genomic selection for economic traits in soft winter wheat. *G3: Genes, Genomes, Genetics*, *6*(9), 2919–2928. https://doi.org/10.1534/g3.116.032532

Imathiu, S. M., Edwards, S. G., Ray, R. V., & Back, M. (2014). Review article: Artificial inoculum and inoculation techniques commonly used in the investigation of fusarium head blight in cereals. *Acta Phytopathologica et Entomologica Hungarica*, *49*(2), 129–139. https://doi.org/10.1556/APhyt.49.2014.2.1

Isidro, J., Jannink, J.-L., Akdemir, D., Poland, J., Heslot, N., & Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *128*(1), 145–158. https://doi.org/10.1007/s00122-014-2418-4

Jansen, C., Von Wettstein, D., Schäfer, W., Kogel, K.-H., Felk, A., & Maier, F. J. (2005). Infection patterns in barley and wheat spikes inoculated with wild-type and

trichodiene synthase gene disrupted Fusarium graminearum. *Proceedings of the National Academy of Sciences*, *102*(46), 16892–16897.

Jia, Y., & Jannink, J.-L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics*, *192*(4), 1513–1522. https://doi.org/10.1534/genetics.112.144246

Jin, F., Zhang, D., Bockus, W., Baenziger, P. S., Carver, B., & Bai, G. (2013). Fusarium head blight resistance in U.S. Winter wheat cultivars and elite breeding lines. *Crop Science*, *53*(5), 2006–2013. https://doi.org/10.2135/cropsci2012.09.0531

Kislev. (1984). *EMERGENCE OF WHEAT AGRICULTURE Author (s): M . E . KISLEV Published by : Paleorient and CNRS Editions and CNRS Editions Stable URL : https://www.jstor.org/stable/41489606 EMERGENCE OF WHEAT AGRICULTURE*. *10*(2), 61–70.

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Plieske, J., Kulosa, D., Ganal, M. W., & Röder, M. S. (2013). Whole Genome Association Mapping of Fusarium Head Blight Resistance in European Winter Wheat (Triticum aestivum L.). *PLoS ONE*, *8*(2). https://doi.org/10.1371/journal.pone.0057500

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, *9*(1), 1–9.

Kotari. (2019). What is Host-Pathogen Initial Interaction Telling Us? an Essential Component of Biotic Stress Response Mechanism! *Advances in Biotechnology & Microbiology*, *12*(4), 4–11. https://doi.org/10.19080/aibm.2019.12.555843

Kourelis, J., & Van Der Hoorn, R. A. L. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell*, *30*(2), 285–299. https://doi.org/10.1105/tpc.17.00579

Kuck, K. H., & Scheinpflug, H. (1986). Biology of sterol-biosynthesis inhibiting fungicides. *Biology of Sterol-Biosynthesis Inhibiting Fungicides.*, 65–96.

Lado. (2017). *Lado 2017 Strategies for Selecting Crosses Using Genomic Prediction in Two Wheat Breeding Programs.pdf*.

Larkin, D. L., Holder, A. L., Mason, R. E., Moon, D. E., Brown-Guedira, G., Price, P. P., Harrison, S. A., & Dong, Y. (2020). Genome-wide analysis and prediction of Fusarium head blight resistance in soft red winter wheat. *Crop Science*, *60*(6), 2882–2900. https://doi.org/10.1002/csc2.20273

Legrand, F., Picot, A., Cobo-Díaz, J. F., Chen, W., & Le Floch, G. (2017). Challenges facing the biological control strategies for the management of Fusarium Head Blight of cereals caused by F. graminearum. *Biological Control*, *113*, 26–38.

Li, G., Zhou, J., Jia, H., Gao, Z., Fan, M., Luo, Y., Zhao, P., Xue, S., Li, N., & Yuan, Y. (2019). Mutation of a histidine-rich calcium-binding-protein gene in wheat confers resistance to Fusarium head blight. *Nature Genetics*, *51*(7), 1106–1112.

Li, G., Zhou, J., Jia, H., Gao, Z., Fan, M., Luo, Y., Zhao, P., Xue, S., Li, N., Yuan, Y., Ma, S., Kong, Z., Jia, L., An, X., Jiang, G., Liu, W., Cao, W., Zhang, R., Fan, J., … Ma, Z. (2019). Mutation of a histidine-rich calcium-binding-protein gene in wheat confers resistance to Fusarium head blight. *Nature Genetics, 51*(7), 1106–1112. https://doi.org/10.1038/s41588-019-0426-7

Li, T., Bai, G., Wu, S., & Gu, S. (2011). Quantitative trait loci for resistance to fusarium head blight in a Chinese wheat landrace Haiyanzhong. *Theoretical and Applied Genetics*, *122*(8), 1497–1502.

Li, T., Bai, G., Wu, S., & Gu, S. (2012). Quantitative trait loci for resistance to Fusarium head blight in the Chinese wheat landrace Huangfangzhu. *Euphytica, 185*(1), 93–102. https://doi.org/10.1007/s10681-012-0631-2

Li, X., Michlmayr, H., Schweiger, W., Malachova, A., Shin, S., Huang, Y., Dong, Y., Wiesenberger, G., McCormick, S., Lemmens, M., Fruhmann, P., Hametner, C., Berthiller, F., Adam, G., & Muehlbauer, G. J. (2017). A barley UDP-glucosyltransferase inactivates nivalenol and provides Fusarium Head Blight resistance in transgenic wheat. *Journal of Experimental Botany*, *68*(9), 2187–2197. https://doi.org/10.1093/jxb/erx109

Li, Y., Xiong, H., Guo, H., Zhou, C., Xie, Y., Zhao, L., Gu, J., Zhao, S., Ding, Y., & Liu, L. (2020). Identification of the vernalization gene VRN-B1 responsible for heading date variation by QTL mapping using a RIL population in wheat. *BMC Plant Biology*, *20*(1), 1–15. https://doi.org/10.1186/s12870-020-02539-5

Lin, F., Xue, S. L., Zhang, Z. Z., Zhang, C. Q., Kong, Z. X., Yao, G. Q., Tian, D. G., Zhu, H. L., Li, C. J., Cao, Y., Wei, J. B., Luo, Q. Y., & Ma, Z. Q. (2006). Mapping QTL associated with resistance to Fusarium head blight in the Nanda2419 x Wangshuibai population. II: type I resistance. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *112*(3), 528–535. https://doi.org/10.1007/s00122-005-0156-3

Liu, J., He, Z., Rasheed, A., Wen, W., Yan, J., Zhang, P., Wan, Y., Zhang, Y., Xie, C., & Xia, X. (2017). Genome-wide association mapping of black point reaction in common wheat (Triticum aestivum L.). *BMC Plant Biology*, *17*(1), 1–12. https://doi.org/10.1186/s12870-017-1167-3

Liu, S., Abate, Z. A., Lu, H., Musket, T., Davis, G. L., & McKendry, A. L. (2007). QTL associated with Fusarium head blight resistance in the soft red winter wheat Ernie. *Theoretical and Applied Genetics*, *115*(3), 417–427. https://doi.org/10.1007/s00122-007-0577-2

Liu, Shuyu, Griffey, C. A., Hall, M. D., McKendry, A. L., Chen, J., Brooks, W. S., Brown-Guedira, G., Van Sanford, D., & Schmale, D. G. (2013). Molecular characterization of field resistance to Fusarium head blight in two US soft red winter wheat cultivars. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *126*(10), 2485–2498. https://doi.org/10.1007/s00122-013-2149-y

Liu, Shuyu, Hall, M. D., Griffey, C. A., & McKendry, A. L. (2009). Meta-analysis of QTL associated with Fusarium head blight resistance in wheat. *Crop Science*, *49*(6), 1955–1968.

Lu, Q., Lillemo, M., Skinnes, H., He, X., Shi, J., Ji, F., Dong, Y., & Bjørnstad, Å. (2013). Anther extrusion and plant height are associated with Type I resistance to Fusarium head blight in bread wheat line "Shanghai-3/Catbird." *Theoretical and Applied Genetics*, *126*(2), 317–334. https://doi.org/10.1007/s00122-012-1981-9

Ma, Z., Xie, Q., Li, G., Jia, H., Zhou, J., Kong, Z., Li, N., & Yuan, Y. (2020). Germplasms, genetics and genomics for better control of disastrous wheat Fusarium head blight. *Theoretical and Applied Genetics*, *133*(5), 1541–1568. https://doi.org/10.1007/s00122-019-03525-8

Maldonado-Ramirez, S. L., Schmale, D. G., Shields, E. J., & Bergstrom, G. C. (2005). The relative abundance of viable spores of Gibberella zeae in the planetary boundary layer suggests the role of long-distance transport in regional epidemics of Fusarium head blight. *Agricultural and Forest Meteorology*, *132*(1–2), 20–27. https://doi.org/10.1016/j.agrformet.2005.06.007

Mammarella, N. D., Cheng, Z., Fu, Z. Q., Daudi, A., Bolwell, G. P., Dong, X., & Ausubel, F. M. (2015). Apoplastic peroxidases are required for salicylic acid-mediated defense against Pseudomonas syringae. *Phytochemistry*, *112*(1), 110–121. https://doi.org/10.1016/j.phytochem.2014.07.010

Mao, S. L., Wei, Y. M., Cao, W., Lan, X. J., Yu, M., Chen, Z. M., Chen, G. Y., & Zheng, Y. L. (2010). Confirmation of the relationship between plant height and Fusarium

head blight resistance in wheat (Triticum aestivum L.) by QTL meta-analysis. *Euphytica*, *174*(3), 343–356. https://doi.org/10.1007/s10681-010-0128-9

Marshall. (2001). *Journal of*. *17*.

Márta Molnár-Láng, Carla Ceoloni, J. D. (2015). *(eds.)-Alien Introgression in Wheat_ Cytogenetics, Molecular Biology, and Genomics-Springer International Publishing (2015).pdf*.

McCartney, C. A., Brûlé-Babel, A. L., Fedak, G., Martin, R. A., McCallum, B. D., Gilbert, J., Hiebert, C. W., & Pozniak, C. J. (2016). Fusarium head blight resistance QTL in the spring wheat cross Kenyon/86ISMN 2137. *Frontiers in Microbiology*, *7*(OCT). https://doi.org/10.3389/fmicb.2016.01542

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., & Van Sanford, D. (2012). Fusarium Head Blight Disease Cycle , Symptoms , and Impact on Grain Yield and Quality Frequency and Magnitude of Epidemics Since 1997. *Plant Disease*, *96*(12), 1712–1728.

Mcmullen, M., Jones, R., Gallenberg, D., & America, S. (1997). Scab of wheat and barley: A re-emerging disease of devastating Impact. *Plant Disease*, *81*(12).

Mesterhazy, A. (2020a). *Updating the Breeding Philosophy of Wheat to* (Issue Di).

Mesterhazy, A. (2020b). Updating the Breeding Philosophy of Wheat to Fusarium head blight (FHB): resistance components, QTL identification, and phenotyping—a review. *Plants*, *9*(12), 1702.

Mesterházy, A. (1995). Types and components of resistance to Fusarium head blight of wheat. *Plant Breeding*, *114*(5), 377–386. https://doi.org/10.1111/j.1439-0523.1995.tb00816.x

Mesterházy, Á., Bartók, T., Kászonyi, G., Varga, M., Tóth, B., & Varga, J. (2005). Common resistance to different Fusarium spp. causing Fusarium head blight in wheat. *European Journal of Plant Pathology*, *112*(3), 267–281. https://doi.org/10.1007/s10658-005-2853-9

Mesterházy, Á., Bartók, T., & Lamper, C. (2003). Influence of wheat cultivar, species of Fusarium, and isolate aggressiveness on the efficacy of fungicides for control of Fusarium head blight. In *Plant Disease* (Vol. 87, Issue 9, pp. 1107–1115). https://doi.org/10.1094/PDIS.2003.87.9.1107

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. In *Genetics Soc America*.

Miedaner, Reinbrecht, Lauber, Schollenberger, & Geiger. (2001). Effects of genotype and genotype-environment interaction on deoxynivalenol accumulation and resistance to Fusarium head blight in rye, triticale, and wheat. *Plant Breeding*, *120*(2), 97–105. https://doi.org/10.1046/j.1439-0523.2001.00580.x

Miller, J. D., Young, J. C., & Sampson, D. R. (1985). Deoxynivalenol and Fusarium head blight resistance in spring cereals. *Journal of Phytopathology*, *113*(4), 359–367.

Mohammadi. (2015). *Mohammadi 2015 PopVar_ A Genome-Wide Procedure for Predicting Genetic Variance and Correlated Response in Biparental Breeding Populations.pdf*.

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., & Buckler, E. (2009). Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell*, *21*(8), 2194–2202. https://doi.org/10.1105/tpc.109.068437

Nganje, W. E., Kaitibie, S., Wilson, W. W., Leistritz, L. F., & Bangsund, D. A. (2004). Economic impacts of fusarium head blight in wheat and barley: 1993-2001. In *Agribusiness & Applied Economics* (Issue 538).

Oliver, R. E., Cai, X., Xu, S. S., Chen, X., & Stack, R. W. (2005). Wheat-alien species derivatives: A novel source of resistance to Fusarium head blight in wheat. *Crop Science*, *45*(4), 1353–1360.

Ollier, M., Talle, V., Brisset, A. L., Le Bihan, Z., Duerr, S., Lemmens, M., Goudemand, E., Robert, O., Hilbert, J. L., & Buerstmayr, H. (2019). Whitened kernel surface: A fast and reliable method for assessing Fusarium severity on cereal grains by digital picture analysis. *Plant Breeding*, *138*(1), 69–81. https://doi.org/10.1111/pbr.12667

Parry, D. W., Jenkinson, P., & McLeod, L. (1995). Fusarium ear blight (scab) in small grain cereals—a review. *Plant Pathology*, *44*(2), 207–238.

Paterson, A. H. (1996). *Making genetic maps, pp. 23–39 in Genome Mapping in Plants, edited by Paterson AH*. Academic Press, San Diego.

Paul, P. A., Lipps, P. E., Hershman, D. E., McMullen, M. P., Draper, M. A., & Madden, L. V. (2008). Efficacy of Triazole-Based Fungicides for Fusarium Head Blight and Deoxynivalenol Control in Wheat: A Multivariate Meta-Analysis. *Phytopathology®*, *98*(9), 999–1011. https://doi.org/10.1094/PHYTO-98-9-0999

Paul, P. A., Lipps, P. E., & Madden, L. V. (2005). Relationship between visual estimates of Fusarium head blight intensity and deoxynivalenol accumulation in harvested wheat grain: A meta-analysis. *Phytopathology*, *95*(10), 1225–1236. https://doi.org/10.1094/PHYTO-95-1225

Paux. (2012). *Paux 2012 Sequence-based marker development in wheat_ Advances and applications to breeding.pdf*.

Pestka, J. J. (2010). Toxicological mechanisms and potential health effects of deoxynivalenol and nivalenol. *World Mycotoxin Journal*, *3*(4), 323–347. https://doi.org/10.3920/WMJ2010.1247

Petersen, S. (2015). *Advancing marker-assisted selection for resistance to powdery mildew and Fusarium head blight in wheat*. North Carolina State University.

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., Dreisigacker, S., Crossa, J., Sánchez-Villeda, H., Sorrells, M., & Jannink, J. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *The Plant Genome*, *5*(3). https://doi.org/10.3835/plantgenome2012.06.0006

Poland, J., & Rutkoski, J. (2016). Advances and Challenges in Genomic Selection for Disease Resistance. *Annual Review of Phytopathology*, *54*, 79–98. https://doi.org/10.1146/annurev-phyto-080615-100056

Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glössl, J., Luschnig, C., & Adam, G. (2003). Detoxification of the Fusarium Mycotoxin Deoxynivalenol by a UDP-glucosyltransferase from Arabidopsis thaliana. *Journal of Biological Chemistry*, *278*(48), 47905–47914. https://doi.org/10.1074/jbc.M307552200

Qi, L. L., Pumphrey, M. O., Friebe, B., Chen, P. D., & Gill, B. S. (2008). Molecular cytogenetic characterization of alien introgressions with gene Fhb3 for resistance to Fusarium head blight disease of wheat. *Theoretical and Applied Genetics*, *117*(7), 1155–1166.

Rawat, N., Pumphrey, M. O., Liu, S., Zhang, X., Tiwari, V. K., Ando, K., Trick, H. N., Bockus, W. W., Akhunov, E., Anderson, J. A., & Gill, B. S. (2016). Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nature Genetics*, *48*(12), 1576–1580. https://doi.org/10.1038/ng.3706

Reif, J. C., Maurer, H. P., Korzun, V., Ebmeyer, E., Miedaner, T., & Würschum, T. (2011). Mapping QTLs with main and epistatic effects underlying grain yield and heading time in soft winter wheat. *Theoretical and Applied Genetics*, *123*(2), 283–292. https://doi.org/10.1007/s00122-011-1583-y

Ringli, C., Keller, B., & Ryser, U. (2001). Glycine-rich proteins as structural components of plant cell walls. *Cellular and Molecular Life Sciences*, *58*(10), 1430–1441. https://doi.org/10.1007/PL00000786

Rudd, J. C., Horsley, R. D., McKendry, A. L., & Elias, E. M. (2001). Host plant resistance genes for Fusarium head blight: sources, mechanisms, and utility in conventional breeding systems. *Crop Science*, *41*(3), 620–627.

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J.-L., & Sorrells, M. (2012). Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *The Plant Genome*, *5*(2), 51–61. https://doi.org/10.3835/plantgenome2012.02.0001

SA Isard, S. G. (2001). *Flow of Life in the Atmosphere*. Michigan State University Press.

Saldago, J. D., Bergstrom, G., Bradley, C., Bowen, K., Byamukama, E., Byrne, A., Collins, A., Cowger, C., Cummings, J., & Chapara, V. (2018). Efficacy of Miravis Ace for FHB and DON management across environments and grain market classes: A progress report. *Natl. Fusarium Head Blight Forum Proc*, 40–42.

Sarinelli, J. M. (2019). *Sarinelli 2019 Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel _ Enhanced Reader.pdf*.

Scherer, A., & Christensen, G. B. (2016). Concepts and relevance of genome-wide association studies. *Science Progress*, *99*(1), 59–67. https://doi.org/10.3184/003685016X14558068452913

Schulthess, A. W., Zhao, Y., Longin, C. F. H., & Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *131*(3), 685–701. https://doi.org/10.1007/s00122-017-3029-7

Semagn, K., Skinnes, H., Bjørnstad, Å., Marøy, A. G., & Tarkegne, Y. (2007). Quantitative trait loci controlling Fusarium head blight resistance and low deoxynivalenol content in hexaploid wheat population from 'Arina'and NK93604. *Crop Science*, *47*(1), 294–303.

Silva, C. L. Da. (2018). Genomic approaches for mapping and predicting disease resistance in wheat (Triticum aestivum L .) by Cristiano Lemes Da Silva *PHD DISSERTATION*.

Snijders, C. H. A. (2004). Resistance in wheat to Fusarium infection and trichothecene formation. *Toxicology Letters*, *153*(1), 37–46.

Srinivasachary, Gosman, N., Steed, A., Hollins, T. W., Bayles, R., Jennings, P., & Nicholson, P. (2009). Semi-dwarfing Rht-B1 and Rht-D1 loci of wheat differ significantly in their influence on resistance to Fusarium head blight. *Theoretical and Applied Genetics*, *118*(4), 695–702. https://doi.org/10.1007/s00122-008-0930-0

Srinivasachary, Gosman, N., Steed, A., Simmonds, J., Leverington-Waite, M., Wang, Y., Snape, J., & Nicholson, P. (2008). Susceptibility to Fusarium head blight is associated with the Rht-D1b semi-dwarfing allele in wheat. *Theoretical and Applied Genetics*, *116*(8), 1145–1153. https://doi.org/10.1007/s00122-008-0742-2

Stack, R. W., & McMullen, M. P. (2011). A visual scale to estimate severity of Fusarium head blight in wheat. *North Dakota State University Extension Service. Extension Publication PP-1095.*, *PP-1095*(November), 2. https://library.ndsu.edu/ir/bitstream/handle/10365/9187/PP1095_1998.pdf?sequence=1&isAllowed=y

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., & Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Tropical Plant Pathology*, *42*(3), 165–174. https://doi.org/10.1007/s40858-017-0127-7

Steiner, B., Lemmens, M., Griesser, M., Scholz, U., Schondelmaier, J., & Buerstmayr, H. (2004). Molecular mapping of resistance to Fusarium head blight in the spring wheat cultivar Frontana. *Theoretical and Applied Genetics*, *109*(1), 215–224.

Su, Z., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H., Cai, S., Liu, D., Zhang, D., Li, T., Trick, H., St. Amand, P., Yu, J., Zhang, Z., & Bai, G. (2019). A deletion mutation in TaHRC confers Fhb1 resistance to Fusarium head blight in wheat. *Nature Genetics*, *51*(7), 1099–1105. https://doi.org/10.1038/s41588-019-0425-8

Su, Z., Jin, S., Zhang, D., & Bai, G. (2018). Development and validation of diagnostic markers for Fhb1 region, a major QTL for Fusarium head blight resistance in wheat. *Theoretical and Applied Genetics*, *131*(11), 2371–2380. https://doi.org/10.1007/s00122-018-3159-6

Thambugala, D., Brûlé-Babel, A. L., Blackwell, B. A., Fedak, G., Foster, A. J., MacEachern, D., Gilbert, J., Henriquez, M. A., Martin, R. A., McCallum, B. D.,

Spaner, D., Iqbal, M., Pozniak, C. J., N'Diaye, A., & McCartney, C. A. (2020). Genetic analyses of native Fusarium head blight resistance in two spring wheat populations identifies QTL near the B1, Ppd-D1, Rht-1, Vrn-1, Fhb1, Fhb2, and Fhb5 loci. *Theoretical and Applied Genetics*, *133*(10), 2775–2796. https://doi.org/10.1007/s00122-020-03631-y

Tundo, S., Paccanaro, M. C., Elmaghraby, I., Moscetti, I., Ovidio, R. D., Favaron, F., & Sella, L. (2020). *The Xylanase Inhibitor TAXI-I Increases Plant*.

Van Ooijen, G., Mayr, G., Kasiem, M. M. A., Albrecht, M., Cornelissen, B. J. C., & Takken, F. L. W. (2008). Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany*, *59*(6), 1383–1397. https://doi.org/10.1093/jxb/ern045

Venske, E., dos Santos, R. S., Farias, D. da R., Rother, V., da Maia, L. C., Pegoraro, C., & Costa de Oliveira, A. (2019). Meta-analysis of the QTLome of fusarium head blight resistance in bread wheat: Refining the current puzzle. *Frontiers in Plant Science*, *10*(June), 1–19. https://doi.org/10.3389/fpls.2019.00727

Verges, V. L., Lyerly, J., Dong, Y., & Van Sanford, D. A. (2020). Training Population Design With the Use of Regional Fusarium Head Blight Nurseries to Predict Independent Breeding Lines for FHB Traits. *Frontiers in Plant Science*, *11*(July), 1–13. https://doi.org/10.3389/fpls.2020.01083

Vogelgsang, S., Hecker, A., Musa, T., Dorn, B., & Forrer, H. R. (2011). On-farm experiments over 5 years in a grain maize/winter wheat rotation: Effect of maize residue treatments on Fusarium graminearum infection and deoxynivalenol

contamination in wheat. *Mycotoxin Research*, *27*(2), 81–96. https://doi.org/10.1007/s12550-010-0079-y

Waldron, B. L., Moreno-Sevilla, B., Anderson, J. A., Stack, R. W., & Frohberg, R. C. (1999). RFLP mapping of QTL for Fusarium head blight resistance in wheat. *Crop Science*, *39*(3), 805–811.

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., Lyu, Z., Chen, L., Xu, S., Guo, J., Li, M., Su, P., Li, X., Wang, G., Bo, C., Fang, X., Zhuang, W., Cheng, X., Wu, J., … Kong, L. (2020). Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. *Science*, *368*(6493). https://doi.org/10.1126/science.aba5435

Wang, R., Chen, J., Anderson, J. A., Zhang, J., Zhao, W., Wheeler, J., Klassen, N., See, D. R., & Dong, Y. (2017). Genome-wide association mapping of Fusarium head blight resistance in spring wheat lines developed in the Pacific Northwest and CIMMYT. *Phytopathology*, *107*(12), 1486–1495. https://doi.org/10.1094/PHYTO-02-17-0073-R

Wang, S.-B., Feng, J.-Y., Ren, W.-L., Huang, B., Zhou, L., Wen, Y.-J., Zhang, J., Dunwell, J. M., Xu, S., & Zhang, Y.-M. (2016). Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Scientific Reports*, *6*(1), 1–10.

Willian W Wilson. (2017). Опыт аудита обеспечения качества и безопасности медицинской деятельности в медицинской организации по разделу

«Эпидемиологическая безопасностьNo Title. In *Вестник Росздравнадзора* (Vol. 4, Issue 774).

Willyerd, K. T., Li, C., Madden, L. V, Bradley, C. A., Bergstrom, G. C., Sweets, L. E., McMullen, M., Ransom, J. K., Grybauskas, A., & Osborne, L. (2012). Efficacy and stability of integrating fungicide and cultivar resistance to manage Fusarium head blight and deoxynivalenol in wheat. *Plant Disease*, *96*(7), 957–967.

Wu, J., Wang, Y., Kim, S. T., Kim, S. G., & Kang, K. Y. (2013). Characterization of a newly identified rice chitinase-like protein (OsCLP) homologous to xylanase inhibitor. *BMC Biotechnology*, *13*. https://doi.org/10.1186/1472-6750-13-4

Xie, C., Gessler, D. D. G., & Xu, S. (1998). Combining different line crosses for mapping quantitative trait loci using the identical by descent-based variance component method. *Genetics*, *149*(2), 1139–1146. https://doi.org/10.1093/genetics/149.2.1139

Xu, S. (2013). Genetic mapping and genomic selection using recombination breakpoint data. *Genetics*, *195*(3), 1103–1115. https://doi.org/10.1534/genetics.113.155309

Xue, S., Li, G., Jia, H., Xu, F., Lin, F., Tang, M., Wang, Y., An, X., Xu, H., Zhang, L., Kong, Z., & Ma, Z. (2010). Fine mapping Fhb4, a major QTL conditioning resistance to Fusarium infection in bread wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *121*(1), 147–156. https://doi.org/10.1007/s00122-010-1298-5

Xue, S., Xu, F., Tang, M., Zhou, Y., Li, G., An, X., Lin, F., Xu, H., Jia, H., & Zhang, L. (2011). Precise mapping Fhb5, a major QTL conditioning resistance to Fusarium

infection in bread wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *123*(6), 1055–1063.

Yu, J., & Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, *17*(2), 155–160. https://doi.org/10.1016/j.copbio.2006.02.003

Zeng, Z.-B. (1994). A composite interval mapping method for locating multiple QTLs. *Proceedings, 5th World Congress on Genetics Applied to Livestock Production, University of Guelph, Guelph, Ontario, Canada*, *7*.

Zhou, J., Loh, Y. T., Bressan, R. A., & Martin, G. B. (1995). The tomato gene Pti1 encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. *Cell*, *83*(6), 925–935. https://doi.org/10.1016/0092-8674(95)90208-2

Zhu, Z., Chen, L., Zhang, W., Yang, L., Zhu, W., Li, J., Liu, Y., Tong, H., Fu, L., Liu, J., Rasheed, A., Xia, X., He, Z., Hao, Y., & Gao, C. (2020). Genome-Wide Association Analysis of Fusarium Head Blight Resistance in Chinese Elite Wheat Lines. *Frontiers in Plant Science*, *11*(February), 1–11. https://doi.org/10.3389/fpls.2020.00206

Chapter 3 Multi-locus genome-wide association studies to characterize Fusarium head blight (FHB) resistance in hard winter wheat

3.1 Abstract

Fusarium head blight (FHB), caused by the fungus *Fusarium graminearum* Schwabe is one of the most devastating diseases of wheat that can cause severe yield losses along with quality and food/feed safety concerns due to the accumulation of mycotoxins in the grains. Incorporating the resistant alleles from wild relatives, landraces or exotic materials remain challenging and have limited success. Therefore, a better understanding of the genetic basis of native FHB resistance in hard winter wheat and combining it with major QTLs can facilitate development of FHB resistant cultivars. In this study, we evaluated a panel of 257 elite breeding lines from the South Dakota State University (SDSU) breeding program to uncover the genetic basis of native FHB resistance in the US hard winter wheat. We conducted a multi-locus genome-wide association study (ML-GWAS) with 9,321 high-quality single nucleotide polymorphisms (SNPs) covering all 21 wheat chromosomes. Marker-trait associations (MTAs) were identified with eight different ML-GWAS models, the most appropriate being Fixed and random model Circulating Probability Unification (FarmCPU) for FHB disease index (DIS) and Fusarium damaged kernels (FDK). A total of six distinct QTNs were identified for DIS on five different chromosomes such as 2A, 2B, 3B, 4B, and 7A, where five were considered ‘reliable QTNs’ as those were identified by multiple models. For FDK, a total of eight unique QTNs were identified on six different chromosomes 3B, 5A, 6B, 6D, 7A, and 7B, where four QTNs were considered reliable. The additive effect of favorable alleles of reliable QTNs was found to be significant as the mean DIS and FDK score decreased with the accumulation of resistant alleles. This current study sheds light on the

genetic basis of native FHB resistance in hard winter wheat germplasm from the US Great Plains region and the QTNs identified in this study would be useful resources for FHB resistance breeding via marker-assisted selection.

Keywords: FHB resistance; GBS; multi-locus GWAS; hard winter wheat; winter wheat breeding for scab

3.2 Introduction

Fusarium head blight (FHB), also known as wheat scab or scab, is one of the most devastating diseases of wheat primarily caused by the fungus *Fusarium graminearum* Schwabe. FHB can cause severe losses in yield due to the shriveled grains and quality concerns due to lower test weight (Gilbert and Tekauz, 2000; Bai and Shaner, 2004). Further, *Fusarium* sp. produces harmful mycotoxins such as deoxynivalenol (DON) that can accumulate in the infected grains and poses a serious threat to food and feed safety, and negatively impact the wheat trade (Pestka, 2010; Ferrigo et al., 2016). In the USA, FHB was first reported by Arthur (1891) in Indiana and since then, FHB has expanded its horizons to all major wheat-growing states in the US. This expansion of FHB is likely due to a suitable climate, increased acreage under no-till cultivation, and adoption of maize-wheat rotations over the last several decades, causing huge economic losses (Nganje et al., 2004; McMullen et al., 2012; Wilson et al., 2017). For instance, wheat producers suffered revenue losses worth $850 million due to FHB outbreaks in the US (Wilson et al., 2017).

Although fungicides are used for FHB prevention and control, the development of FHB resistant varieties is still the most effective and economical approach to minimize the losses caused by this disease (Bai and Shaner, 2004; Gilbert and Haber, 2013). Genetic resistance to FHB is complex and controlled by multiple quantitative trait loci (QTLs) with small to medium effects. Further resistance expression is also significantly influenced by environmental conditions (Miedaner et al., 2001; Buerstmayr et al., 2012). Several types of resistance mechanisms have been proposed and commonly evaluated, including resistance to the initial infection (Type I), resistance to the spread of infection

within the spike (Type II), resistance to accumulation of mycotoxins such as deoxynivalenol (DON) (Type III), and resistance against damaged kernels (Type IV) (Bai and Shaner, 2004; Gilbert and Haber, 2013), with type II resistance being more stable and utilized in many wheat breeding programs. Nevertheless, type III and type IV resistance have also received attention in wheat breeding because of being associated with end-use quality, which is the biggest concern of the growers (Mesterhazy, 2020; Verges et al., 2020). The genetic analysis of FHB resistance in wheat has been extensively reported, and a large number of QTLs were identified on 21 wheat chromosomes (Liu et al., 2009; Venske et al., 2019), including seven cataloged FHB genes, *Fhb1* to *Fhb7* (Liu et al., 2009; Su et al., 2019; Venske et al., 2019; Ma et al., 2020), with most of these originating from Asian germplasm such as the Chinese wheat variety 'Sumai-3' and 'Wahgnshuibai' or landraces (Bai et al., 1999; Anderson et al., 2001; Buerstmayr et al., 2009; Xue et al., 2010; Steiner et al., 2017) and wild relatives (*Fhb3*, *Fhb6*) (REF). Nevertheless, the transfer of resistance from wild relatives, landraces or exotic materials is challenging and leads to linkage drag and adaptability issues. Thus, only a few QTLs, in particular *Fhb1*, with a major effect on FHB resistance are successfully employed by wheat breeding programs mostly by spring wheat programs (Bai et al., 2018). Contrarily, majority of the germplasm from hard winter wheat region of the US relies upon the variation in FHB resistance from native sources including cultivars like 'Everest', 'Overland', 'Lyman', and 'Expedition' (Clinesmith et al., 2019; Zhang et al., 2022). However, identification of genomic regions underlying native resistance and the development of reliable markers is needed to pyramid an effective level of resistance in required backgrounds. Thus, it is

important to determine the genetic basis of native FHB resistance from this region and exploit it in the regional breeding programs.

Though numerous QTLs for FHB resistance have been identified using traditional linkage mapping, this approach can encompass limited diversity. Genome-wide association studies (GWAS) provide a good alternative by providing a much higher resolution to capture insights into the genetic architecture of complex traits because of historically accumulated mutations or recombination events (Francioli et al., 2016; Scherer and Christensen, 2016). GWAS has been successfully used to dissect several traits of economic importance in wheat (Sukumaran et al., 2014; Sidhu et al., 2020; AlTameemi et al., 2021), including a few studies for FHB resistance analysis (Kollers et al., 2013; Arruda et al., 2016; Wang et al., 2017; Zhu et al., 2020). However, none of these studies have been performed in the US hard winter wheat. Furthermore, recent developments in multi-locus GWAS (ML-GWAS) models have improved the power and reliability of this approach to identify causal loci for complex traits. For instance, more powerful methods like FarmCPU and BLINK have improved the ability of GWAS to detect smaller effects loci (Liu et al., 2016; Huang et al., 2019). Apart from these models, several important ML-GWAS models that have been reported to outperform conventional GWAS models include the multi-locus random-SNP-effect mixed linear model (mrMLM), fast multi-locus random-SNP-effect mixed linear model (FASTmrMLM), fast multi-locus random-SNP-effect efficient mixed-model analysis (FASTmrEMMA), iterative modified-sure independence screening Expectation-Maximization-Bayesian least absolute shrinkage and selection operator (ISIS EM-BLASSO), polygenic-background-control based least angle regression plus empirical Bayes (pLARmEB), and

pKWmEB (Wang et al., 2016; Tamba et al., 2017; Zhang et al., 2017; Ren et al., 2018; Wen et al., 2018). The ML-GWAS models are not merely more reliable and efficient, they also overcome the requirement of multiple testing corrections that always results in false negatives (Zhang et al., 2019).

Majority of the GWAS studies make use of assembled diversity panels or landraces in various crop species (Ward et al., 2019). However, several studies effectively used the panels consisting of elite breeding lines to dissect the genetic basis of various traits of economic importance (Sukumaran et al., 2014, 2018), however, none of these studies have explored FHB resistance in the US hard winter wheat breeding germplasm. This approach permits identification, mapping and the direct transfer and pyramiding of identified QTLs to the new backgrounds in the breeding programs without any linkage drag. For this study, we envisaged using a panel of elite lines from the South Dakota State University (SDSU) breeding program to uncover the genetic basis of native FHB resistance in our program and exploit that for the development of improved FHB resistance varieties. The panel of elite lines was phenotyped for FHB resistance in a controlled FHB field nursery over two years and genotyped using genotyping-by-sequencing (GBS) approach. The specific objectives of this study were to (i) evaluate the genetic basis of FHB resistance in hard winter wheat elite breeding material and identify associated markers to facilitate marker-assisted selection; (ii) identify candidate genes in the regions significantly associated with FHB traits.

3.3 Materials and method

3.3.1 Plant materials and FHB screening

A set of 257 breeding lines from the SDSU winter wheat breeding program was used in this study. The SDSU winter wheat breeding program evaluates a set of advanced and elite breeding lines for FHB resistance in a mist irrigated field nursery each growing season. Most of the breeding lines are evaluated at either $F_{4:7}$ or $F_{4:8}$ filial generation. The 257 lines used in the current study were evaluated in the growing season of 2019 and 2020. Among these lines, 169 were screened in the 2019 nursery and 154 in the 2020 nursery, with an overlap of 58 lines between the two seasons. Owing to the missing genotype data or inconsistent replications, a total of eight and one lines were removed from 2019 and 2020, respectively, leaving 257 unique lines for downstream analysis.

The FHB nurseries were planted at Brookings, South Dakota (44.3114°N, 96.7984°W) during the growing seasons of 2019 and 2020. A randomized complete block design with 2 or 3 replicates for different sets of lines was used to design the experiment. The resistant and susceptible checks genotypes for FHB resistance were included in each nursery, where cultivars ‘Lyman’ and ‘Emerson’ were used as resistant checks while ‘Flourish’ was the susceptible checks. Each experimental unit consisted of a 1-meter-long row plot with an inter-row spacing of 20 cm. The experiments were managed using the regional standard cultural practices for the proper growth and development of wheat plants. Days to heading (DTH) were recorded by calculating the Julian date when 50% of the plot had completely emerged heads. Plant height (PH) was measured from the soil surface to the top of spikes excluding awns at maturity.

The FHB nurseries were inoculated using corn spawn and inoculum spraying with *F. graminearum* isolates (SD-FG1) as described in (Halder et al., 2019). Briefly, the Fusarium-infected corn kernels (scabby corn inoculum) were scattered in the field first boot stage (Feekes 10), and followed by another round at the heading (Feekes 10.1) stages to ensure maximum infection in the FHB nursery. In addition, direct spray inoculation was used at 50% anthesis using a conidial suspension containing 100,000 spores/ml to avoid any disease escape. The sprinkler head irrigation system was used to mist the nurseries every night (19:00–7:00) for 2 minutes (every 15 minutes) to promote a humid micro-environment for disease development. FHB disease incidence (INC) and disease severity (SEV) were scored after 21 days of inoculation following the scale described by Stack and McMullen, 2011. These traits were scored using 20 spikes per replicate/genotype. The FHB disease index (DIS) was used for resistance evaluation and calculated as (INC × SEV)/100. The percentage of Fusarium damaged kernels (FDK) was evaluated after the rows were harvested utilizing low air-speeds to prevent the loss of shriveled kernels. The grain samples were visually scored for FDK by using a set of known FDK standards (https://agcrops.osu.edu/newsletter/corn-newsletter/2015-21/rating-fusarium-damaged-kernels-fdk-scabby-wheat) two replications per sample.

3.3.2 Statistical analysis

The phenotypic data from two seasons was analyzed to obtain the best linear unbiased estimates (BLUEs) for FHB traits using the following model:

$$y_{ijk} = \mu + E_i + R_{j(i)} + G_k + GE_{ik} + e_{ijk}$$

where y_{ijk} is the trait of interest, μ is the overall mean, E_i is the effect of the i^{th} environment, $R_{j(i)}$ is the effect of the j^{th} replicate nested within the ith environment, G_k is the effect of the k^{th} genotype, GE_{ik} is the effect of the genotype × environment (G × E) interaction, and e_{ijk} is the residual error associated with the replication and genotype effects. The broad-sense heritability (H^2) of a trait of interest in a combined environment analysis was assessed based on the variance estimates from the linear mixed model as follows:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{ge}^2/nLoc + \sigma_e^2/(nLoc \times nRep)}$$

where σ_g^2 and σ_e^2 , are the genotype and error variance components, σ_g^2 is the G × E interaction variance component and *nLoc* is the number of environments in the analysis. The analysis was performed in META-R (Alvarado et al., 2019) which is based on LME4 (Bates et al., 2015) R package. The summary statistics, correlations, visualization, and comparison tests were performed in R (R Core Team, 2014).

3.3.3 Genotyping and quality control

The genotyping of the plant material was performed at USDA Central Small Grain Genotyping Lab, Manhattan, KS, using the Genotyping-by-sequencing GBS procedure (Poland et al., 2012). Briefly, the genomic DNA was extracted from young leaf tissue for each line using a Cetyl Trimethylammonium Bromide (CTAB) method (Bai et al., 1999). GBS libraries were prepared by double restriction digestion with HF-*PstI* and *MspI* enzymes (Poland et al., 2012) and sequenced using an Ion Proton sequencer (Thermo Fisher Scientific, Waltham, MA, USA). The Chinese Spring wheat genome reference RefSeq v2.0 (IWGSC, 2018; Zhu et al., 2021) was used to align the GBS reads using the

default settings of Burrows-Wheeler Aligner v0.6.1 and the single nucleotide polymorphisms (SNPs) were called using the GBS v2.0 discovery pipeline in TASSEL v5.0 (Bradbury et al., 2007). For quality control, SNPs with more than 30% missing calls, minor allele frequency (MAF) of less than 5%, and unmapped on any chromosome were filtered out, leaving 9,321 high-quality SNPs for downstream analysis. The high-quality SNPs were imputed using BEAGLE v4.1 (Browning and Browning, 2007) for further analysis.

3.3.4 Population Structure and linkage disequilibrium analysis

The principal component analysis (PCA) of the filtered and imputed genotypic data was conducted to analyze population structure in 257 breeding lines in R (R Core Team 2014) and visualized using 'ggplot2' (Wickham, 2016). Further, we analyzed the population stratification using a Bayesian model-based clustering program, STRUCTURE v2.3.4 assuming an Admixture model (Pritchard et al., 2000). STRUCTURE analysis was performed by assuming ten subgroups (*K = 1 - 10*) with ten independent runs for each subgroup using a burn-in period of 10,000 iterations followed by 20,000 Monte-Carlo iterations. The analysis was implemented in parallel using StrAuto v1.0 on the SDSU high-performance computing (HPC) cluster (Chhatre and Emerson, 2017; Tange, 2018). An ad-hoc statistic (DeltaK) was used to infer the most likely number of subgroups using STRUCTURE HARVESTER (Evanno et al., 2005; Earl and vonHoldt, 2012). The Linkage disequilibrium (LD) parameter r^2 for the whole genome as well as each sub-genome was estimated separately in TASSEL v5.0 (Bradbury et al., 2007) by computing r^2 values for all pairwise markers using a sliding window size of 50 markers. LD decay

over genetic distance was visualized by fitting a non-linear model using the modified Hill and Weir method (Hill and Weir, 1988) in R (Team, 2014).

3.3.5 Multi-locus genome-wide association analysis

We used ML-GWAS to identify marker-trait associations (MTAs) using BLUEs for FHB traits obtained using the mixed effect model and 9,321 high-quality SNPs. For association analysis, we compared a total of eight ML-GWAS models. Two models, Fixed and random model

Circulating Probability Unification (FarmCPU) and the Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK), were implemented in Genomic Association and Prediction Integrated Tool (GAPIT) v3.0 (Wang and Zhang, 2021) in the R environment. In addition, we used six recently developed ML-GWAS methods including mrMLM, FASTmrMLM, FASTmrEMMA, pLARmEB, ISIS EM-BLASSO, and pkWmEB. All these six models were implemented in the R package ‘mrMLM v4.0.2’ (Zhang et al., 2020) using default parameters. The ML-GWAS models included the estimated kinship (K) and the first two principal components from PCA as covariates to account for relatedness and the population structure. Based on the comparison using quantile-quantile (QQ) plots for all the models, we decided to report the results using the FarmCPU model as it showed better control of false positives and false negatives. Furthermore, we used a strict threshold based on False Discovery Rate (FDR, alpha = 0.05) correction for multiple testing. Though final results were reported from a single best model (FarmCPU), we used the results from the other seven ML-GWAS models to validate the FarmCPU MTAs as reliable, if they were also identified by

other models. The Manhattan plots and QQ plots were generated using the R package ‘CMPlot’ to visualize the results from the FarmCPU analysis.

We also used a pairwise t-test to compare differences in trait means for different alleles of significant MTAs. For each MTA, mean trait values for two groups of alleles (resistant v/s susceptible) were compared using a t-test and visualized using boxplots with R package ‘ggplot2’ (Wickham, 2016). Furthermore, the allelic frequencies of significant MTAs were analyzed to compare the effect of the combination of resistant alleles for DIS and FDK. The 257 accessions were grouped based on the resistant alleles carried for each trait. These groups were compared using an FDR-adjusted pairwise t-test.

3.3.6 Candidate gene analysis

Two highly significant QTNs for FDK were subjected to candidate gene analysis to identify genes with putative functions of interest. Linkage blocks harboring these two QTNs were deduced using the confidence interval method in Haploview (Barrett et al., 2005). These MTAs were physically mapped to Chinese Spring RefSeq v2.1 using marker sequences of significant SNPs (IWGSC, 2018; Zhu et al., 2021). The high confidence (HC) genes from IWGSC v2.1 RefSeq annotation were extracted from a flanking window around each MTA based on the LD decay in the respective region. The HC genes were annotated manually using Blast2GO (Conesa et al., 2005) for the identification of genes of interest. Further, a gene expression browser (http://www.wheat-expression.com/) and a thorough review of literature were used to exclude the unlikely candidates. For the gene expression browser, we used expression data from several

studies related to Fusarium infection to narrow down the genes of interest (Borrill et al., 2016).

3.4 Results

3.4.1 Observed variation for FHB traits

The BLUE values across two seasons exhibited a significant genotypic variation ($P < 0.001$) for DIS and FDK in the panel of 257 breeding lines. The distribution for DIS and FDK is presented in Table 1. The variation for DIS ranged from 12.6 to 90.3, while the FDK ranged from 13.6 to 97.6 (Table 1). We observed a high broad-sense heritability (H^2) for DIS ($H^2 = 0.85$), whereas a moderate heritability for FDK ($H^2 = 0.76$). Based on the BLUE values across seasons, the disease indices (DIS) for two moderately resistant checks namely ‘Emerson’ and ‘Lyman were 36.3 and 31.6, respectively, whereas the susceptible check ‘Flourish’ was 77.1. Similarly, the FDK percentage for ‘Emerson’ and ‘Lyman was 48.9% and 31.5%, whereas ‘Flourish’ was rated 84.5%. Pearson correlation coefficients estimated using the phenotypic BLUEs for DIS and FDK were significant ($P < 0.001$), showing a positive correlation value of 0.44 (Figure 1). Furthermore, we also estimated Person correlation among the FHB traits, PH, and DTH. A significant negative correlation was observed between DIS and DTH ($r = 0.24$, $P < 0.001$) as well as FDK and DTH $r = 0.17$, $P < 0.01$). Intriguingly, FHB traits (DIS and FDK) were not significantly correlated with PH (Figure 1).

3.4.2 Genotyping, population structure, and linkage disequilibrium

The genotyping using the GBS approach yielded a total of 9,321 high-quality SNPs which were used for downstream analysis. Among 9,321 SNPs, the numbers of SNPs

from the A and B sub-genomes were comparatively higher than the D sub-genome, with the B sub-genome (4,202; 45.1%) having the highest and the D sub-genome (1,418; 15.2%) having lowest SNP density (Supplementary Table S1). The highest SNPs were found on chromosome 7A (796) whereas chromosome 4D had the lowest number (36 SNPs). The LD analysis revealed a different pattern of LD decay among the three subgenomes, with the whole genome LD decay being around 3.5 Mbp (Supplementary Figure S1). Further, sub-genomes A and B showed a smaller LD decay distance compared to sub-genome D (Supplementary Figure S1). The population structure among 257 accessions was inferred using PCA and STRUCTURE analysis (Figure 2). The DeltaK statistic was used to estimate optimal subgroups based on STRUCTURE analysis and it showed a major peak at $K = 2$, suggesting only two major groups in the panel (Figure 1a). The principal component analysis also showed considerable admixture in the population, indicating the presence of two subgroups within 257 accessions (Figures 1b and 1c), with the first two principal components explaining around only 6.5% and 3.4% of the total variance, respectively.

3.4.3 Genomic loci associated with FHB traits

Association analysis was initially performed using eight different ML-GWAS methods. Overall, these eight ML-GWAS models identified 52 quantitative trait nucleotides (QTNs) for DIS and 53 QTNs for FDK (Supplementary Table S2). Nevertheless, we compared all these models based on QQ plots and FarmCPU was found to fit best for both the traits suggesting better control of false positives and false negatives. Thus, we used this model to report final QTNs for DIS and FDK (Figure 3).

A total of six distinct QTNs were identified for DIS using FarmCPU on five different chromosomes based on FDR corrected threshold (Table 2, Figure 3A). The most significant QTN was identified on the short arm of chromosome 4B (*S4B_40315424*) physically mapped at 40 Mbp. The second most significant QTN (*S3B_773516625*) was found on chromosome 3B located at 773 Mbp. Of the six QTNs for DIS, five were considered 'reliable QTNs' as these were identified by at least one another ML-GWAS model except one QTN on chromosome 2B (*S2B_725552556*) (Table 2). For FDK, a total of eight unique QTNs were identified using FarmCPU located on six different chromosomes (Table 2, Figure 3B), with one QTN each on chromosomes 3B, 5A, 6D, and 7B, two QTNs on 6B and three QTNs on 7A. Among the eight QTNs for FDK, four were considered reliable as these were identified by another ML-GWAS model. The four reliable QTNs for FDK were located on chromosomes 3B (*S3B_768314878*), 5A (*S5A_619020400*), 6B (*S6B_718194425*), and 7B (*S7B_707550430*) (Table 2). Out of the 14 QTNs identified for DIS and FDK, two QTNs (*S3B_768314878* and *S4B_647586119*) were found to be pleiotropic for both the traits (Supplementary Table S2).

Further, we compared trait values for DIS and FDK among the two alleles of the identified reliable QTNs (Figure 4). Of the five reliable QTNs for DIS, four exhibited a statistically significant difference in mean DIS score (Figure 4a). The mean DIS score of the lines with resistant allele (43.4) of QTN represented by *S3B_773516625* was 25.6% lower than those with susceptible allele (58.3). Similarly, QTN on 4B (*S4B_40315424*) showed a decrease of 25.8% in the mean DIS score from susceptible (47.3) to resistant (35.1) alleles. For FDK, all four reliable QTNs showed statistically significant differences

in mean FDK percentage among the two alleles (Figure 4b). Intriguingly, favorable alleles for two QTNs (*S6B_718194425* and *S7B_707550430*) exhibited a decrease of around 13% in mean FDK percentage over the unfavorable allele.

3.4.4 Additive effect of identified QTNs

We investigated the effect of accumulating favorable alleles for reliable QTNs on DIS and FDK. The panel of 257 accessions was categorized into groups based on the number of favorable or resistant alleles carried by accessions. For DIS, five groups were identified carrying one, two, three, four, or five resistant alleles. We observed a significant decrease in the mean DIS score as the number of resistant alleles increased (Figure 5a). The mean DIS for the group of accession having only '1' resistant allele was 62.1, while the mean DIS for the group with '5' resistant alleles was 31.5 (Figure 5a). Similarly, five groups were observed based on the resistant alleles for FDK (Figure 5b). A significant reduction was observed in mean FDK with an increase in resistant alleles (Figure 5b).

3.4.5 Relationship between FHB and height genes

The most significant QTN for DIS (*S4B_40315424*) was identified on the short arm of chromosome 4B physically mapped at 40 Mbp, the region which harbors wheat dwarfing gene *Rht1*. We used the PH data for 257 accessions to identify marker-trait association for this trait. Interestingly, we identified the most significant QTN for plant height at 40 Mbp on chromosome 4B (Supplementary Table S3), suggesting the same loci linked to both traits.

3.4.6 Candidate gene analysis for important QTNs

The candidate gene analysis was performed for two QTNs (*S6B_718194425* and *S7B_707550430*) associated with FDK, as we were able to define a narrow region based on LD decay for these QTNs. For the QTN located on chromosome 6B (*S6B_718194425*), a 1.7 Mbp linkage block was identified harboring significant SNP (Supplementary Figure S3a). Similarly, we identified a 2.3 Mbp long linkage block harboring QTN *S7B_707550430* (Supplementary Figure S3b). Based on Chinese Spring RefSeq v2.1, 28 and 20 high-confidence genes were retrieved for the 6B and 7B QTNs, respectively. Further analysis using the wheat expression browser (http://www.wheat-expression.com) with *Fusarium* specific studies, and comparison with literature removed 32 genes (Supplementary Table S4), leaving only 16 genes with putative functions of interest (Table 3). Among these 16 genes, five genes from 6B region (*TraesCS6B02G448800*, *TraesCS6B02G448900*, *TraesCS6B02G450000*, *TraesCS6B02G450200*, and *TraesCS6B02G450500*) and three genes from 7B region (*TraesCS7B02G417000*, *TraesCS7B02G429800* and *TraesCS7B02G430000*) were

of specific interest as they exhibited a differential expression between mock and *Fusarium* inoculated spikes in Chinese Spring (Supplementary Figures S4 and S5).

3.5 Discussion

The utilization of host resistance to develop FHB resistant wheat cultivars is the most economical and sustainable approach to manage FHB. This necessitates the continuous identification and validation of novel sources of FHB resistance and their utilization in breeding programs using marker-assisted selection. Thus, research efforts have resulted

in the identification of several major and minor genes for FHB resistance including *Fhb1*, and their pyramiding across various breeding programs, particularly in spring wheat (Steiner et al., 2017; Bai et al., 2018; Ghimire et al., 2020). Nevertheless, using wild introgression (*Fhb3*, *Fhb6*) or exotic resistance sources, such as Sumai3, leads to the linkage drag or undesirable agronomic traits, making it difficult to incorporate these genes into regional breeding programs. Thus, breeders rely on identifying and utilizing native FHB resistance for improving the FHB resistance.

Majority of the hard winter wheat (HWW) cultivars from the Great Plain region of the US, including the SDSU winter wheat program, do not carry *Fhb1* likely due to yield drag with only one HWW variety (TAM 205) carrying *Fhb1* has been released to date (Zhang et al., 2022). Fortunately, there are several cultivars including 'Everest', 'Overland', 'Lyman', 'Heyne', 'Century', and 'Hondo' that exhibit moderate resistance to FHB but do not carry *Fhb1* have been released in the US hard winter wheat region (Jin et al., 2013; Bai et al., 2018; Clinesmith et al., 2019; Zhang et al., 2022), showing the importance of the native FHB resistance in the regional programs. Further, various studies have successfully identified QTLs for native resistance using cultivars like 'Art', 'Everest', and 'Lyman' (Clinesmith et al., 2019; Hashimi, 2019). Thus, we used a set of advanced and elite breeding lines from the SDSU program to identify genomic regions associated with FHB resistance, which could be readily employed in developing improved cultivars.

Two FHB traits, DIS and FDK percentage were used to evaluate FHB resistance in a panel of advanced breeding lines. We observed a significant variation for both the traits, with DIS scores ranging from 12.6 to 90.3% and FDK from 13.6 to 97.6%. As the

majority of our material did not have *Fhb1* based on the parentage, the significant genotypic variation observed for both the traits suggests a presence of underlying sources for native resistance. Furthermore, we observed a moderate to high heritability for DIS and FDK, which was in corroboration with several previous studies (Larkin et al., 2020; Xu et al., 2020; Zhu et al., 2020).

ML-GWAS using the FarmCPU algorithm identified a total of six and eight QTNs associated with DIS and FDK, respectively (Table 2). As previous studies used various types of marker systems to map FHB resistance and are currently mapped to Chinese Spring RefSeq 1.0, it is difficult to precisely compare the QTNs from the current study (mapped using Chinese Spring RefSeq 2.0) with previously identified regions. Therefore, we identified the approximate physical locations of previous QTLs and QTNs from the current study on Chinese Spring RefSeq 1.0 to facilitate the comparison and validation of the genomic regions (Supplementary Table S5) (IWGSC, 2018; Zhao et al., 2019).

Out of the six QTNs identified for DIS (Table 2), four were found to be located in genomic regions previously reported to have QTLs for various FHB traits. We identified a QTN on the long arm of chromosome 2A (718 Mbp), which corresponds to a stable QTL for FHB resistance (~709 Mbp) identified in US Soft Red Winter Wheat Breeding Line 'VA00W-38' (Liu et al., 2012). Another QTN (*S3B_773516625*) was identified for DIS on chromosome 3B which mapped to 753 Mbp on RefSeq v1.0. This QTN was found in similar location (753 Mbp) to that of a recently identified genomic region for type-III FHB resistance from Canadian spring wheat cultivar 'AAC Tenacious', validating the importance of this region (Dhariwal et al., 2020). Further, a QTN identified on 7AS (*S7A_48708273*) was present in the genomic region (~28 – 68 Mbp) reported to

harbor QTLs for FHB resistance (Zhang et al., 2010; Jiang et al., 2020; Thambugala et al., 2020).

Several studies have reported a co-localization of QTLs for FHB resistance with dwarfing genes such as *Rht-B1* and *Rht-D1*, with dwarfing alleles at these loci related to FHB susceptibility (Miedaner and Voss, 2008; Srinivasachary et al., 2008; Liu et al., 2012; Dhariwal et al., 2020; Thambugala et al., 2020; Goddard et al., 2021). In the current study, we identified a strong QTN for DIS on the short arm of chromosome 4B (37 Mbp), which co-localized with the *Rht-B1* region. Intriguingly, we did not observe a strong association based on plant height and any of the FHB traits based on Pearson's correlation (Figure 1). Further, we conducted a GWAS for PH using 257 accessions and identified the most significant QTN for plant height at the same location as DIS QTN (37 Mbp) (Supplementary Table S3). In corroboration with previous studies (Srinivasachary et al., 2008; Buerstmayr and Buerstmayr, 2016), our results suggest that the susceptibility associated with dwarfing allele of *Rht-B1* might be caused by a potential linkage of susceptible genes with the dwarfing genes. Thus, it will be beneficial to either evaluate other reduced height genes, such as *Rht24*, or identification of recombinants where the linkage between *Rht* genes and susceptible gene(s) has been broken. Except for these four QTNs for DIS, we did not find any previously reported QTN in the proximity of QTNs on chromosome 2BL (*S2B_725552556*) and 4BL (*S4B_647586119*). Thus, these two QTNs could likely represent novel genomic regions for native FHB resistance in hard winter wheat.

Apart from DIS, we identified a total of eight QTNs for FDK (Table 2). Comparison with previous studies revealed that six QTNs were found in the genomic regions associated

with FHB resistance. The QTN identified for FDK on chromosome 3B (*S3B_768314878*) co-localized with the QTN for DIS (*S3B_773516625*) in the same region, showing a pleiotropic effect on both traits. As discussed earlier, Dhariwal et al. (2020) also identified a stable QTL for DON at similar position, suggesting it is an important region for FHB resistance. Similarly, the QTN identified on 5A (*S5A_619020400*) at 617 Mbp for FDK was found in the proximity of two previously reported QTLs, one mapped at ~596 Mbp for FDK in soft red winter wheat varieties AGS 2060 and AGS 2035 (Castro Aviles et al., 2020), whereas the second QTL for DON mapped at 621 Mbp in Canadian spring wheat cultivar 'AAC Tenacious' (Dhariwal et al., 2020). Further, we identified two QTNs (*S6B_320696398* and *S6B_718194425*) for FDK on chromosome 6B at 314 Mbp and 708 Mbp, respectively (Table 2). QTN at 708 Mbp was found to align within the confidence interval of a meta-QTL for FHB resistance on chromosome 6BL (Venske et al., 2019). However, we did not find any previous QTL around 314 Mbp. Another QTN (*S6D_110313864*) from this study was mapped on the short arm of 6D. Though few studies reported QTLs for FHB resistance in this region, we were unable to compare the exact location due to different marker systems.

Two QTNs (*S7A_713432647* and *S7A_738859192*) for FDK identified in this study correspond to 707 Mbp and 731 Mbp on CS RefSeq v1.0 (Table 2). Several studies have identified QTLs for different FHB traits in this genomic region. An interval of 611 to 724 Mbp was delimited for a Fusarium-damaged kernels QTL in 'Nanda 2419 × Wangshuibai' population (Li et al., 2008). Apart from this, QTLs were reported for different FHB traits in the 7A region but could not be compared precisely based on physical position (Li et al., 2012; Lu et al., 2013; Guo et al., 2020). Further, we identified

a significant QTN on chromosome 7B mapped at 698 Mbp. Castro Aviles et al. (2020) also mapped a QTL for DON at ~ 718 Mbp in soft red winter wheat. Similarly, a QTL for FHB resistance was mapped at ~683 Mbp from a cross between Ningmai-9 and Yangmai-158 (Jiang et al., 2020). Overall, we identified 14 QTNs associated with DIS and FDK in the current study. Out of these, ten QTNs co-localized with previously reported genomic regions for different FHB traits. Our study validated the previous QTLs in hard winter wheat and with higher marker coverage compared to the majority of previous studies, we identified more tightly associated SNPs with these QTLs. Further, identified SNPs associated with these QTLs could be used to develop KASP markers, which can be effective in tracking and pyramiding these reliable QTLs for native FHB resistance in required backgrounds using marker-assisted selection (Gill et al., 2019). Apart from the validated regions, we also identified four likely novel genomic regions associated with FHB resistance that could be subjected to further investigation.

Further, we performed candidate gene analysis for two important FDK QTNs (S6B_718194425 and S7B_707550430) to identify candidate genes with putative functions of interest. In wheat, the majority of disease genes that have been cloned are reported to encode intracellular immune receptors of the nucleotide binding-site–leucine-rich repeat (NBS-LRR) family, receptor-like kinases (RLKs), or wall-associated kinases (WAKs) as the protein product (Keller et al., 2018). However, two recently characterized FHB resistance genes, *Fhb1* and *Fhb7*, were reported to have different mechanisms. Several candidates for *Fhb1* have been reported including a pore-forming toxin-like (PFT) gene encoding a chimeric lectin with two agglutinin domains; a mutation of a histidine-rich calcium-binding protein gene (Rawat et al., 2016; Bai et al., 2018; Li et al.,

2019; Su et al., 2019). Nevertheless, none of these genes did share any conserved domains related to the disease-resistance gene cloned in plants (Li et al., 2019; Su et al., 2019). Recently, a gene encoding a glutathione S-transferase (GST) was determined as the *Fhb7*, which can detoxify pathogen-produced toxins by conjugating a glutathione (GSH) unit onto the epoxide moieties of the pathogenic molecule (Wang et al., 2020). Based on this information, we were able to identify several putative candidate genes for two QTNs (Table 3). We found several genes encoding putative disease resistance proteins including LRR receptor-like serine/ threonine-protein kinase, or nicotinate N-methyltransferase 1-like proteins (Table 3), that may play a role in the process of intracellular detection of pathogen-derived molecules and signal transduction (Zhou et al., 1995) in the candidate region. Further, we used a wheat expression browser (Borrill et al., 2016) using expression data from *Fusarium* infected spikes. We identified a few genes in 6B and 7B QTN regions that showed differential expression among mock-inoculated, and *Fusarium* inoculated spikes (Supplementary Figures S4 and S5). It could be interesting to further investigate the role of these genes in FHB resistance.

In summary, the current study provides new insights into the genetic basis of native FHB resistance in hard winter wheat germplasm from the Great Plains region of the US. The study validates the role of ten QTLs in regulation of FHB resistance in wheat including hard winter wheat. Further, four potential novel QTN and the associated markers for 14 QTN can facilitate the deployment of these QTLs through marker-assisted selection. Further, the information on genomic regions associated with FHB resistance could be useful for the breeders to improve the genomic selection models to select breeding lines with improved FHB resistance.

3.6 References

AlTameemi, R., Gill, H. S., Ali, S., Ayana, G., Halder, J., Sidhu, J. S., et al. (2021). Genome-wide association analysis permits characterization of Stagonospora nodorum blotch (SNB) resistance in hard winter wheat. *Sci. Rep.* 11, 12570. doi:10.1038/s41598-021-91515-6.

Alvarado, G., López, M., Vargas, M., Pacheco, Á., Rodríguez, F., Burgueño, J., et al. (2019). META-R (Multi Environment Trail Analysis with R for Windows) Version 6.04. Available at: https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10201 [Accessed December 30, 2020].

Anderson, J. A., Stack, R. W., Liu, S., Waldron, B. L., Fjeld, A. D., Coyne, C., et al. (2001). DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theor. Appl. Genet.* 102, 1164–1168. doi:10.1007/s001220000509.

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., et al. (2016). Genome-Wide Association Mapping of Fusarium Head Blight Resistance in Wheat using Genotyping-by-Sequencing. *Plant Genome* 9. doi:10.3835/plantgenome2015.04.0028.

Bai, G., Kolb, F. L., Shaner, G., and Domier, L. L. (1999). Amplified fragment length polymorphism markers linked to a major quantitative trait locus controlling scab resistance in wheat. *Phytopathology* 89, 343–348. doi:10.1094/PHYTO.1999.89.4.343.

Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to fusarium head blight. *Annu. Rev. Phytopathol.* 42, 135–161.

doi:10.1146/annurev.phyto.42.040803.140340.

Bai, G., Su, Z., and Cai, J. (2018). Wheat resistance to Fusarium head blight. *Can. J. Plant Pathol.* 40, 336–346. doi:10.1080/07060661.2018.1476411.

Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265. doi:10.1093/bioinformatics/bth457.

Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67. doi:10.18637/jss.v067.i01.

Borrill, P., Ramirez-Gonzalez, R., and Uauy, C. (2016). expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform. *Plant Physiol.* 170, 2172–2186. doi:10.1104/pp.15.01667.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi:10.1093/bioinformatics/btm308.

Browning, S. R., and Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097. doi:10.1086/521987.

Buerstmayr, H., Ban, T., and Anderson, J. A. (2009). QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: A review. *Plant Breed.* 128, 1–26. doi:10.1111/j.1439-0523.2008.01550.x.

Buerstmayr, M., and Buerstmayr, H. (2016). The semidwarfing alleles Rht-D1b and Rht-B1b show marked differences in their associations with anther-retention in wheat

heads and with fusarium head blight susceptibility. *Phytopathology* 106, 1544–1552. doi:10.1094/PHYTO-05-16-0200-R.

Buerstmayr, M., Huber, K., Heckmann, J., Steiner, B., Nelson, J. C., and Buerstmayr, H. (2012). Mapping of QTL for Fusarium head blight resistance and morphological and developmental traits in three backcross populations derived from Triticum dicoccum × Triticum durum. *Theor. Appl. Genet.* 125, 1751–1765. doi:10.1007/s00122-012-1951-2.

Castro Aviles, A., Alan Harrison, S., Joseph Arceneaux, K., Brown-Guidera, G., Esten Mason, R., and Baisakh, N. (2020). Identification of QTLs for Resistance to Fusarium Head Blight Using a Doubled Haploid Population Derived from Southeastern United States Soft Red Winter Wheat Varieties AGS 2060 and AGS 2035. *Genes (Basel).* 11, 699. doi:10.3390/genes11060699.

Chhatre, V. E., and Emerson, K. J. (2017). StrAuto: Automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics* 18, 192. doi:10.1186/s12859-017-1593-0.

Clinesmith, M. A., Fritz, A. K., da Silva, C. L., Bockus, W. W., Poland, J. A., Dowell, F. E., et al. (2019). QTL mapping of fusarium head blight resistance in winter wheat cultivars ‘art’ and ‘everest.’ *Crop Sci.* 59, 911–924. doi:10.2135/cropsci2018.04.0276.

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi:10.1093/bioinformatics/bti610.

Dhariwal, R., Henriquez, M. A., Hiebert, C., McCartney, C. A., and Randhawa, H. S. (2020). Mapping of major fusarium head blight resistance from canadian wheat cv. Aac tenacious. *Int. J. Mol. Sci.* 21, 1–24. doi:10.3390/ijms21124497.

Earl, D. A., and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi:10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14, 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x.

Ferrigo, D., Raiola, A., and Causin, R. (2016). Fusarium toxins in cereals: Occurrence, legislation, factors promoting the appearance and their management. *Molecules* 21. doi:10.3390/molecules21050627.

Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., et al. (2016). HHS Public Access. 47, 822–826. doi:10.1038/ng.3292.Genome-wide.

Ghimire, B., Sapkota, S., Bahri, B. A., Martinez-Espinoza, A. D., Buck, J. W., and Mergoum, M. (2020). Fusarium Head Blight and Rust Diseases in Soft Red Winter Wheat in the Southeast United States: State of the Art, Challenges and Future Perspective for Breeding. *Front. Plant Sci.* 11, 1080. doi:10.3389/fpls.2020.01080.

Gilbert, J., and Haber, S. (2013). Overview of some recent research developments in fusarium head blight of wheat. *Can. J. Plant Pathol.* 35, 149–174. doi:10.1080/07060661.2013.772921.

Gilbert, J., and Tekauz, A. (2000). Review: Recent developments in research on fusarium head blight of wheat in Canada. *Can. J. Plant Pathol.* 22, 1–8.

doi:10.1080/07060660009501155.

Gill, H. S., Li, C., Sidhu, J. S., Liu, W., Wilson, D., Bai, G., et al. (2019). Fine Mapping of the Wheat Leaf Rust Resistance Gene Lr42. *Int. J. Mol. Sci.* 20. doi:10.3390/ijms20102445.

Goddard, R., Steed, A., Scheeren, P. L., Maciel, J. L. N., Caierão, E., Torres, G. A. M., et al. (2021). Identification of Fusarium head blight resistance loci in two Brazilian wheat mapping populations. *PLoS One* 16, e0248184. doi:10.1371/journal.pone.0248184.

Guo, J., Khan, J., Pradhan, S., Shahi, D., Khan, N., Avci, M., et al. (2020). Multi-Trait Genomic Prediction of Yield-Related Traits in US Soft Wheat under Variable Water Regimes. *Genes (Basel).* 11, 1270. doi:10.3390/genes11111270.

Halder, J., Zhang, J., Ali, S., Sidhu, J. S., Gill, H. S., Talukder, S. K., et al. (2019). Mining and genomic characterization of resistance to tan spot, Stagonospora nodorum blotch (SNB), and Fusarium head blight in Watkins core collection of wheat landraces. *BMC Plant Biol.* 19, 1–15. doi:10.1186/s12870-019-2093-3.

Hashimi, A. (2019). Mapping Quantitative Trait Loci for Fusarium Head Blight Resistance in US Winter Wheat.

Hill, W. G., and Weir, B. S. (1988). Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* 33, 54–78. doi:10.1016/0040-5809(88)90004-4.

Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8, 1–12. doi:10.1093/gigascience/giy154.

IWGSC (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361, eaar7191. doi:10.1126/science.aar7191.

Jiang, P., Zhang, X., Wu, L., He, Y., Zhuang, W., Cheng, X., et al. (2020). A novel QTL on chromosome 5AL of Yangmai 158 increases resistance to Fusarium head blight in wheat. *Plant Pathol.* 69, 249–258. doi:10.1111/ppa.13130.

Jin, F., Zhang, D., Bockus, W., Baenziger, P. S., Carver, B., and Bai, G. (2013). Fusarium head blight resistance in U.S. Winter wheat cultivars and elite breeding lines. *Crop Sci.* 53, 2006–2013. doi:10.2135/cropsci2012.09.0531.

Keller, B., Wicker, T., and Krattinger, S. G. (2018). Advances in Wheat and Pathogen Genomics: Implications for Disease Control. *Annu. Rev. Phytopathol.* 56, 67–87. doi:10.1146/annurev-phyto-080516-035419.

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., et al. (2013). Whole Genome Association Mapping of Fusarium Head Blight Resistance in European Winter Wheat (Triticum aestivum L.). *PLoS One* 8. doi:10.1371/journal.pone.0057500.

Larkin, D. L., Holder, A. L., Mason, R. E., Moon, D. E., Brown-Guedira, G., Price, P. P., et al. (2020). Genome-wide analysis and prediction of Fusarium head blight resistance in soft red winter wheat. *Crop Sci.* 60, 2882–2900. doi:10.1002/csc2.20273.

Li, C., Zhu, H., Zhang, C., Lin, F., Xue, S., Cao, Y., et al. (2008). Mapping QTLs associated with Fusarium-damaged kernels in the Nanda 2419 × Wangshuibai population. *Euphytica* 163, 185–191. doi:10.1007/s10681-007-9626-9.

Li, G., Zhou, J., Jia, H., Gao, Z., Fan, M., Luo, Y., et al. (2019). Mutation of a histidine-

rich calcium-binding-protein gene in wheat confers resistance to Fusarium head blight. *Nat. Genet.* 51, 1106–1112.

Li, T., Bai, G., Wu, S., and Gu, S. (2012). Quantitative trait loci for resistance to Fusarium head blight in the Chinese wheat landrace Huangfangzhu. *Euphytica* 185, 93–102. doi:10.1007/s10681-012-0631-2.

Liu, S., Christopher, M. D., Griffey, C. A., Hall, M. D., Gundrum, P. G., and Brooks, W. S. (2012). Molecular characterization of resistance to fusarium head blight in U.S. soft red winter wheat breeding line VA00W-38. *Crop Sci.* 52, 2283–2292. doi:10.2135/cropsci2012.03.0144.

Liu, S., Hall, M. D., Griffey, C. A., and McKendry, A. L. (2009). Meta-Analysis of QTL associated with fusarium head blight resistance in wheat. *Crop Sci.* 49, 1955–1968. doi:10.2135/cropsci2009.03.0115.

Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLOS Genet.* 12, e1005767. doi:10.1371/journal.pgen.1005767.

Lu, Q., Lillemo, M., Skinnes, H., He, X., Shi, J., Ji, F., et al. (2013). Anther extrusion and plant height are associated with Type I resistance to Fusarium head blight in bread wheat line “Shanghai-3/Catbird.” *Theor. Appl. Genet.* 126, 317–334. doi:10.1007/s00122-012-1981-9.

Ma, Z., Xie, Q., Li, G., Jia, H., Zhou, J., Kong, Z., et al. (2020). Germplasms, genetics and genomics for better control of disastrous wheat Fusarium head blight. *Theor. Appl. Genet.* 133, 1541–1568. doi:10.1007/s00122-019-03525-8.

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G.,

et al. (2012). Fusarium Head Blight Disease Cycle , Symptoms , and Impact on Grain Yield and Quality Frequency and Magnitude of Epidemics Since 1997. *Plant Dis.* 96, 1712–1728.

Miedaner, Reinbrecht, Lauber, Schollenberger, and Geiger (2001). Effects of genotype and genotype-environment interaction on deoxynivalenol accumulation and resistance to Fusarium head blight in rye, triticale, and wheat. *Plant Breed.* 120, 97–105. doi:10.1046/j.1439-0523.2001.00580.x.

Miedaner, T., and Voss, H.-H. (2008). Effect of Dwarfing *Rht* Genes on Fusarium Head Blight Resistance in Two Sets of Near-Isogenic Lines of Wheat and Check Cultivars. *Crop Sci.* 48, 2115–2122. doi:10.2135/cropsci2008.02.0107.

Nganje, W. E., Kaitibie, S., Wilson, W. W., Leistritz, L. F., and Bangsund, D. A. (2004). *Economic impacts of fusarium head blight in wheat and barley: 1993-2001*.

Pestka, J. J. (2010). Toxicological mechanisms and potential health effects of deoxynivalenol and nivalenol. *World Mycotoxin J.* 3, 323–347. doi:10.3920/WMJ2010.1247.

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome* 5, plantgenome2012.06.0006. doi:10.3835/plantgenome2012.06.0006.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155, 945 LP – 959. Available at: http://www.genetics.org/content/155/2/945.abstract.

Rawat, N., Pumphrey, M. O., Liu, S., Zhang, X., Tiwari, V. K., Ando, K., et al. (2016). Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming

toxin-like domain conferring resistance to Fusarium head blight. *Nat. Genet.* 48, 1576–1580. doi:10.1038/ng.3706.

Ren, W. L., Wen, Y. J., Dunwell, J. M., and Zhang, Y. M. (2018). PKWmEB: Integration of Kruskal-Wallis test with empirical Bayes under polygenic background control for multi-locus genome-wide association study. *Heredity (Edinb).* 120, 208–218. doi:10.1038/s41437-017-0007-4.

Robert W. Stack, and Marcia P. McMullen (2011). A Visual Scale to Estimate Severity of Fusarium Head Blight in Wheat. Fargo Available at: www.ag.ndsu.edu/agcomm/creative-commons. [Accessed April 23, 2021].

Scherer, A., and Christensen, G. B. (2016). Concepts and relevance of genome-wide association studies. *Sci. Prog.* 99, 59–67. doi:10.3184/003685016X14558068452913.

Sidhu, J. S., Singh, D., Gill, H. S., Brar, N. K., Qiu, Y., Halder, J., et al. (2020). Genome-Wide Association Study Uncovers Novel Genomic Regions Associated With Coleoptile Length in Hard Winter Wheat. *Front. Genet.* 10, 1345. doi:10.3389/fgene.2019.01345.

Srinivasachary, Gosman, N., Steed, A., Simmonds, J., Leverington-Waite, M., Wang, Y., et al. (2008). Susceptibility to Fusarium head blight is associated with the Rht-D1b semi-dwarfing allele in wheat. *Theor. Appl. Genet.* 116, 1145–1153. doi:10.1007/s00122-008-0742-2.

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., and Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Trop. Plant Pathol.* 42, 165–174. doi:10.1007/s40858-

017-0127-7.

Su, Z., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H., et al. (2019). A deletion mutation in TaHRC confers Fhb1 resistance to Fusarium head blight in wheat. *Nat. Genet.* 51, 1099–1105. doi:10.1038/s41588-019-0425-8.

Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P., and Reynolds, M. P. (2014). Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor. Appl. Genet.* 128, 353–363. doi:10.1007/s00122-014-2435-3.

Sukumaran, S., Lopes, M., Dreisigacker, S., and Reynolds, M. (2018). Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theor. Appl. Genet.* 131, 985–998. doi:10.1007/s00122-017-3037-7.

Tamba, C. L., Ni, Y. L., and Zhang, Y. M. (2017). Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies. *PLoS Comput. Biol.* 13. doi:10.1371/journal.pcbi.1005357.

Tange, O. (2018). GNU Parallel 2018. doi:10.5281/ZENODO.1146014.

Team, R. C. (2014). R: A language and environment for statistical computing. Available at: http://www.r-project.org/.

Thambugala, D., Brûlé-Babel, A. L., Blackwell, B. A., Fedak, G., Foster, A. J., MacEachern, D., et al. (2020). Genetic analyses of native Fusarium head blight resistance in two spring wheat populations identifies QTL near the B1, Ppd-D1, Rht-1, Vrn-1, Fhb1, Fhb2, and Fhb5 loci. *Theor. Appl. Genet.* 133, 2775–2796. doi:10.1007/s00122-020-03631-y.

Venske, E., dos Santos, R. S., Farias, D. da R., Rother, V., da Maia, L. C., Pegoraro, C., et al. (2019). Meta-analysis of the QTLome of fusarium head blight resistance in bread wheat: Refining the current puzzle. *Front. Plant Sci.* 10, 727. doi:10.3389/fpls.2019.00727.

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., et al. (2020). Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. *Science (80-.).* 368. doi:10.1126/science.aba5435.

Wang, J., and Zhang, Z. (2021). GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. *Genomics. Proteomics Bioinformatics*. doi:10.1016/j.gpb.2021.08.005.

Wang, R., Chen, J., Anderson, J. A., Zhang, J., Zhao, W., Wheeler, J., et al. (2017). Genome-wide association mapping of Fusarium head blight resistance in spring wheat lines developed in the Pacific Northwest and CIMMYT. *Phytopathology* 107, 1486–1495. doi:10.1094/PHYTO-02-17-0073-R.

Wang, S. B., Feng, J. Y., Ren, W. L., Huang, B., Zhou, L., Wen, Y. J., et al. (2016). Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* 6, 1–10. doi:10.1038/srep19444.

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS One* 14, e0208217. doi:10.1371/journal.pone.0208217.

Wen, Y.-J., Zhang, H., Ni, Y.-L., Huang, B., Zhang, J., Feng, J.-Y., et al. (2018). Methodological implementation of mixed linear models in multi-locus genome-wide

association studies. *Brief. Bioinform.* 19, 700–712. doi:10.1093/bib/bbw145.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. Available at: https://cran.r-project.org/web/packages/ggplot2/citation.html [Accessed February 27, 2022].

Wilson, W. W., McKee, G., Nganje, W., Dahl, B., and Bangsund, D. (2017). Economic Impact of USWBSI's Scab Initiative to Reduce FHB. Fargo doi:10.22004/AG.ECON.264672.

Xu, Q., Xu, F., Qin, D., Li, M., Fedak, G., Cao, W., et al. (2020). Molecular mapping of qtls conferring fusarium head blight resistance in chinese wheat cultivar jingzhou 66. *Plants* 9, 1–15. doi:10.3390/plants9081021.

Xue, S., Li, G., Jia, H., Xu, F., Lin, F., Tang, M., et al. (2010). Fine mapping Fhb4, a major QTL conditioning resistance to Fusarium infection in bread wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 121, 147–156. doi:10.1007/s00122-010-1298-5.

Zhang, J., Feng, J. Y., Ni, Y. L., Wen, Y. J., Niu, Y., Tamba, C. L., et al. (2017). PLARmEB: Integration of least angle regression with empirical Bayes for multilocus genome-wide association studies. *Heredity (Edinb).* 118, 517–524. doi:10.1038/hdy.2017.8.

Zhang, J., Gill, H. S., Brar, N. K., Halder, J., Ali., S., Liu, X., et al. (2022). Genomic prediction of Fusarium head blight resistance in early stages using advanced breeding lines in hard winter wheat. *Crop J.* doi:10.1016/j.cj.2022.03.010.

Zhang, M., Zhang, R., Yang, J., and Luo, P. (2010). Identification of a new QTL for Fusarium head blight resistance in the wheat genotype "'Wang shui-bai.'" *Mol. Biol. Rep.* 37, 1031–1035. doi:10.1007/s11033-009-9809-7.

Zhang, Y.-M., Jia, Z., and Dunwell, J. M. (2019). Editorial: The Applications of New Multi-Locus GWAS Methodologies in the Genetic Dissection of Complex Traits. *Front. Plant Sci.* 10, 100. doi:10.3389/fpls.2019.00100.

Zhang, Y. W., Tamba, C. L., Wen, Y. J., Li, P., Ren, W. L., Ni, Y. L., et al. (2020). mrMLM v4.0.2: An R Platform for Multi-locus Genome-wide Association Studies. *Genomics, Proteomics Bioinforma.* 18, 481–487. doi:10.1016/j.gpb.2020.06.006.

Zhao, C., Sun, H., Guan, C., Cui, J., Zhang, Q., Liu, M., et al. (2019). Physical information of 2705 PCR-based molecular markers and the evaluation of their potential use in wheat. *J. Genet.* 98, 69. doi:10.1007/s12041-019-1114-1.

Zhou, J., Loh, Y. T., Bressan, R. A., and Martin, G. B. (1995). The tomato gene Pti1 encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. *Cell* 83, 925–935. doi:10.1016/0092-8674(95)90208-2.

Zhu, T., Wang, L., Rimbert, H., Rodriguez, J. C., Deal, K. R., De Oliveira, R., et al. (2021). Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *Plant J.* 107, 303–314. doi:10.1111/tpj.15289.

Zhu, Z., Chen, L., Zhang, W., Yang, L., Zhu, W., Li, J., et al. (2020). Genome-Wide Association Analysis of Fusarium Head Blight Resistance in Chinese Elite Wheat Lines. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.00206.

3.7 Appendix

List of tables and figures:

Table 3.1. Phenotypic variation, variance estimates, and broad-sense heritability for studied FHB traits. DIS, FHB disease index; FDK, fusarium damaged kernel percentage; CV, coefficient of variation; H^2, broad-sense heritability.

Trait	**Genotypic variance**	**Mean**	**Min**	**Max**	**CV**	**H^2**
DIS	109.1***	45.4	12.6	90.3	20.7	0.85
FDK	156.2***	70.6	13.6	97.6	16.3	0.76

[a]Statistically significant differences are denoted by an asterisk (*) where * $P \leq 0.05$, ** $P \leq 0.01$, and *** $P \leq 0.001$.

Table 3.2. Significant marker-trait associations (MTAs) identified by genome-wide association studies (GWAS) using the FarmCPU model for DIS and FDK. The MTAs were declared significant based on FDR corrected *P*-value threshold of 0.05.

Trait	SNP	Chr	Pos[a]	MAF[b]	SNP effect	P value	FDR-adj P value	Another model[c]
DIS	S2A_722857568	2A	722,857,568	0.17	3.70	2.72E-05	0.046	1,2,3,6
	S2B_725552556	2B	725,552,556	0.18	3.98	2.94E-05	0.046	-
	S3B_773516625	3B	773,516,625	0.14	-5.47	2.29E-07	0.001	1,2,3,6,7
	S4B_40315424	4B	40,315,424	0.16	-7.00	5.46E-14	0.000	1,2,3,4,5,6,7
	S4B_647586119	4B	647,586,119	0.26	3.51	2.37E-06	0.007	1,4,5
	S7A_48708273	7A	48,708,273	0.07	-4.74	1.96E-05	0.047	7
FDK	S3B_768314878	3B	768,314,878	0.40	3.61	8.23E-08	0.000	1,2,3,4
	S5A_619020400	5A	619,020,400	0.13	4.63	7.83E-07	0.000	1,2,7
	S6B_320696398	6B	320,696,398	0.12	4.05	7.41E-06	0.012	-
	S6B_718194425	6B	718,194,425	0.23	3.64	5.98E-07	0.001	1,2,3,4,5,6,7
	S6D_110313864	6D	110,313,864	0.06	-7.50	4.09E-08	0.000	-
	S7A_713432647	7A	713,432,647	0.12	-4.44	1.13E-05	0.015	-
	S7A_738859192	7A	738,859,192	0.34	2.87	3.51E-07	0.001	-
	S7B_707550430	7B	707,550,430	0.30	-2.86	3.7E-05	0.043	1,2,6

[a]The physical position is based on IWGSC RefSeq v2.0 (IWGSC, 2018)

[b]MAF refers to minimum allele frequency for the corresponding MTA

[c]This column enlists ML-GWAS model(s), which identified the corresponding MTA in addition to FarmCPU. The QTN was referred to as 'reliable' if identified by at least two ML-GWAS models. Various ML-GWAS models are: 1, mrMLM; 2, FastmrMLM; 3, FastmrEMMA, 4, pLARmEB; 5, pKWmEB; 6, ISIS EM-BLASSO; and 7, BLINK.

Table 3.3. Candidate genes identified for two QTNs for FDK, *S6B_718194425* and *S7B_707550430*, with putative functions of interest and their functional annotation.

Gene ID[a]	Previous ID[b]	Start position[c]	Annotation
TraesCS6B03G1248600	TraesCS6B02G449300	718084520	FHA domain-containing protein DDL-like
TraesCS6B03G1248900	TraesCS6B02G449400	718124729	nicotinate N-methyltransferase 1-like
TraesCS6B03G1249200	TraesCS6B02G449500	718134558	disease resistance protein RGA5-like
TraesCS6B03G1249300	TraesCS6B02G449600	718142603	disease resistance protein RGA5-like isoform X1
TraesCS6B03G1249800	TraesCS6B02G450000	718403799	aquaporin PIP1-5-like
TraesCS6B03G1250200	TraesCS6B02G450200	718437357	aquaporin PIP1-5-like
TraesCS6B03G1250700	TraesCS6B02G450500	718634434	50S ribosomal protein L9, chloroplastic
TraesCS6B03G1251200	TraesCS6B02G450700	718948307	acyl-CoA-binding domain-containing protein 4-like
TraesCS6B03G1252500	TraesCS6B02G451300	719516962	NAC domain-containing protein 78-like
TraesCS7B03G1160200	TraesCS7B02G417000	706703917	hypothetical protein CFC21_105377
TraesCS7B03G1160400	TraesCS7B02G417100	706707637	NBS-LRR disease resistance protein
TraesCS7B03G1161500	TraesCS7B02G417300	706844055	putative disease resistance protein RGA3
TraesCS7B03G1162000	TraesCS7B02G417400	706905895	probable LRR receptor-like serine/threonine-protein kinase
TraesCS7B03G1166000	TraesCS7B02G429700	708194926	hydroquinone glucosyltransferase-like
TraesCS7B03G1167100	TraesCS7B02G429800	708341181	uncharacterized methyltransferase At2g41040,
TraesCS7B03G1167600	TraesCS7B02G430200	708568667	putative disease resistance RPP13-like protein 1 isoform X1

[a]Gene ID based on the IWGSC RefSeq Annotation v2.1 (IWGSC 2018; Zhu et al. 2021)

[b]Previous IDs for respective genes to the IDs used in IWGSC RefSeq Annotation v1.1 (IWGSC 2018)

[c]Physical position of start points for respective genes are based on IWGSC RefSeq v2.0 (IWGSC 2018)

Figures

Figure 3.1 Correlation coefficients among investigated FHB traits, PH, and DTH calculated by using the best linear unbiased estimates (BLUEs) obtained from a combined analysis of two years. DIS, FHB disease index; FDK, Fusarium damaged kernel percentage; PH, plant height; and DTH, days to heading. The diagonal of the pair plot elucidates the frequency distribution for four traits. Statistically significant differences are denoted by an asterisk (*) where * P ≤ 0.05, ** P ≤ 0.01, and *** P ≤ 0.001.

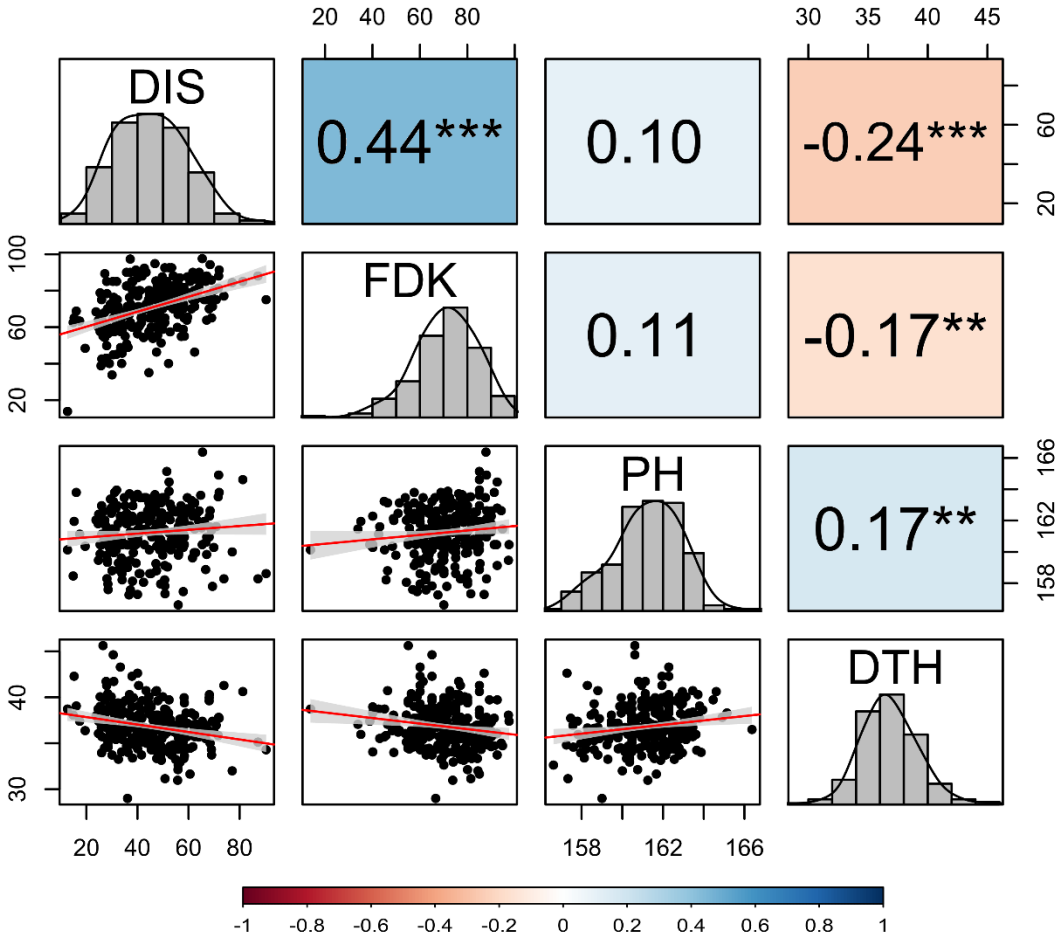


Figure 3.2 Population structure analysis in 257 accessions based on the 9,321 SNPs, (a) Evanno plot of Delta-K statistic from the STRUCTURE analysis. (b) Scree plot for first ten components obtained from principal component analysis (PCA). (c) Scatterplot for the first two components (PC1 and PC2) from PCA.

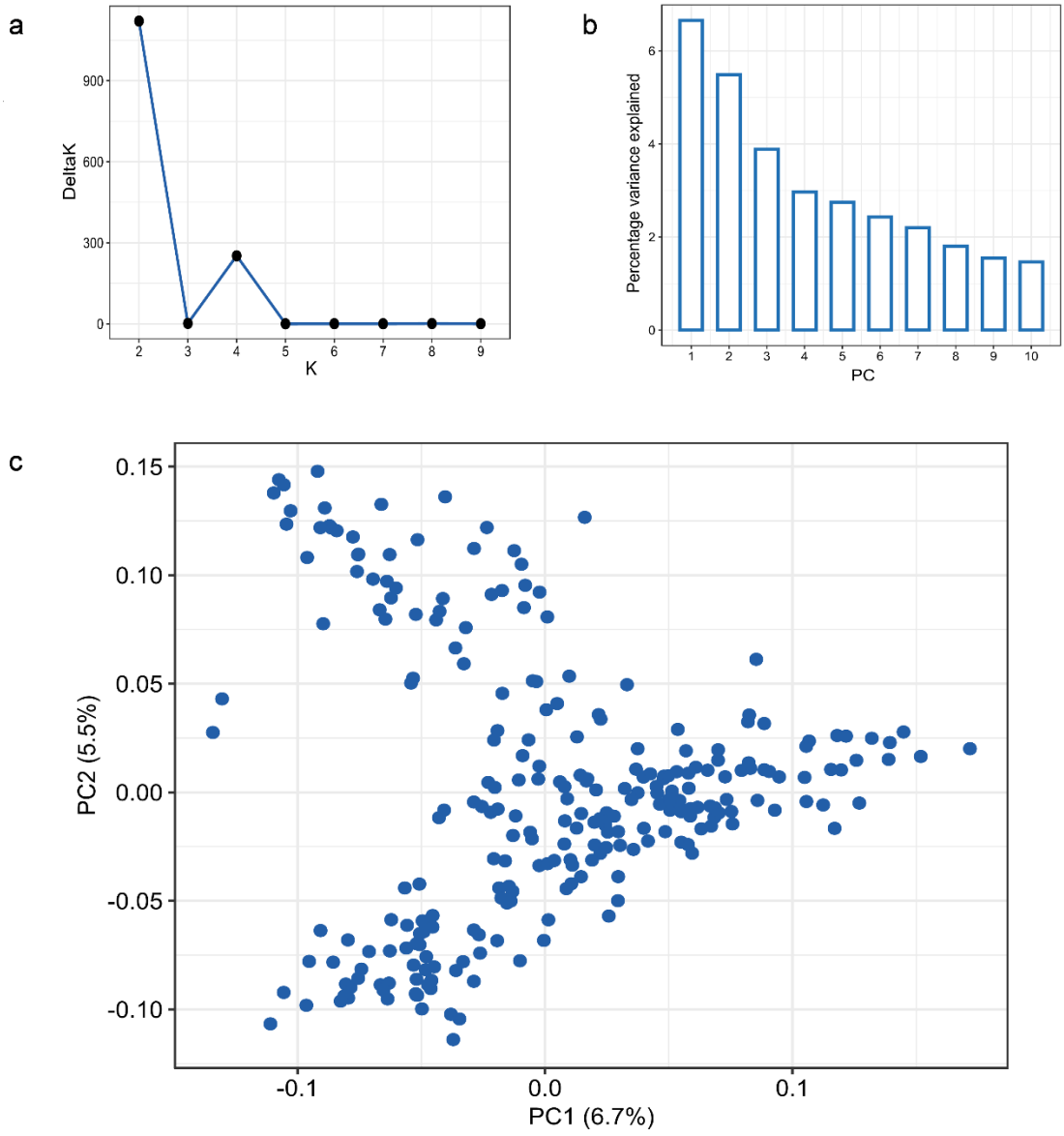


Figure 3.3 Multi-locus marker trait association for DIS and FDK using the FarmCPU model. Manhattan and QQ plot for (a) DIS and (b) FDK showing the distinct peaks for identified QTNs. The significant associations (FDR $P < 0.05$) are elucidated using solid pink vertical lines.

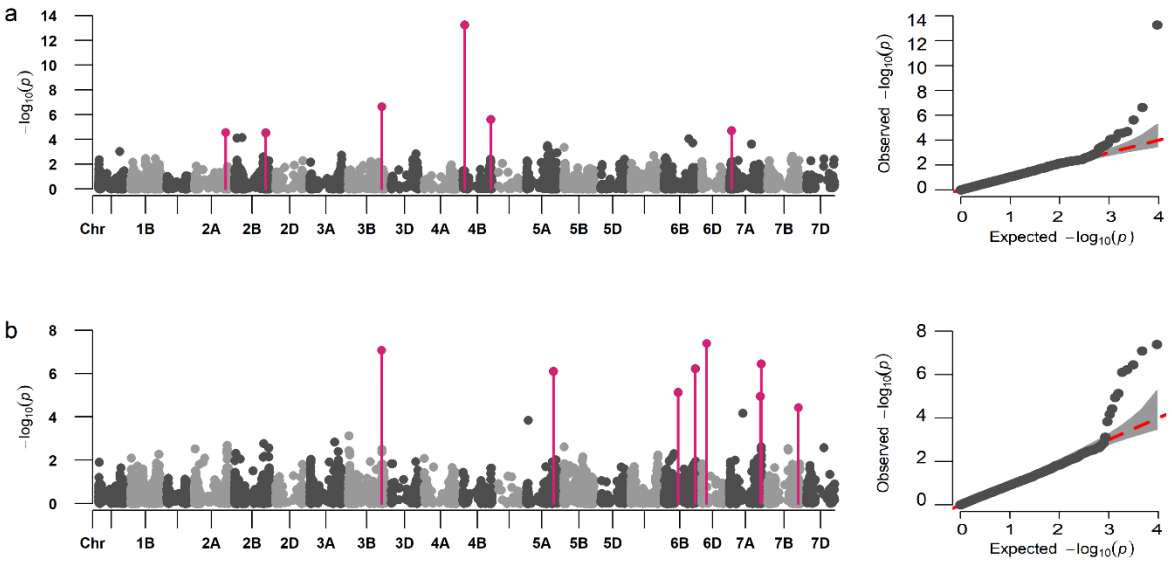


Figure 3.4 Boxplots showing the effect of two alleles (favorable v/s unfavorable) of the reliable QTNs (enlisted in Table 2) on the trait means for (a) DIS (b) FDK. Statistically significant differences are denoted by an asterisk (*) where * $P \leq 0.05$, ** $P \leq 0.01$, and *** $P \leq 0.001$.

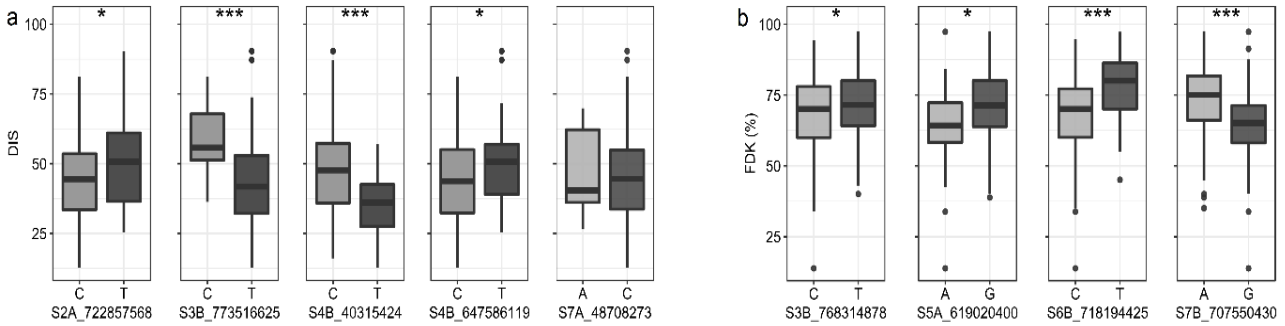
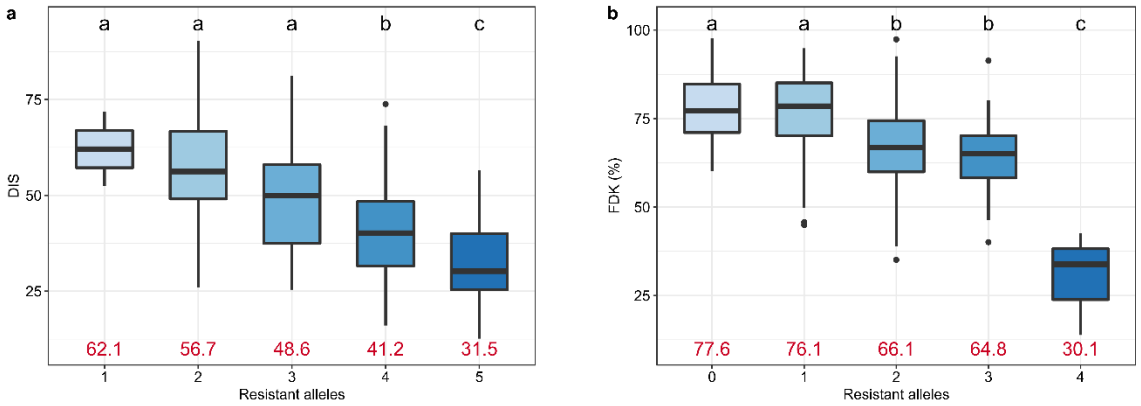


Figure 3.5 Boxplots comparing the trait performance of the lines carrying different numbers of resistant alleles for (a) DIS and (b) FDK, compared using an FDR adjusted Least Significance Difference (LSD) test. Different letters on top of the boxplots denote statistically different groups. The mean trait value for DIS and FDK for the corresponding group is given using red text.



Supplementary Table S1. The distribution of 9,321 SNPs across 21 wheat chromosomes in the panel of 257 accessions.

Sub-genome	Chromosome	Number of SNPs	% SNPs
A	1	532	
	2	447	
	3	518	
	4	391	
	5	508	
	6	509	
	7	796	
Subtotal A		3,701	39.71
B	1	588	
	2	689	
	3	749	
	4	221	
	5	617	
	6	743	
	7	595	
Subtotal B		4,202	45.08
D	1	202	
	2	357	
	3	276	
	4	36	
	5	191	
	6	162	
	7	194	
Subtotal D		1,418	15.21
Total (A, B, and D)		9,321	100

Supplementary Table S2. Summary of all the QTNs for DIS and FDK identified using eight ML-GWAS models. For FarmCPU and BLINK, the threshold to declare an association as significant was FDR adjusted $P < 0.05$. For other six models, the associations were declared significant based on LOD > 3.

Trait	Model	SNP	Chromosome	Position[a]
DIS	BLINK	S2B_173523267	2B	173523267
	BLINK	S3B_773516625	3B	773516625
	BLINK	S4B_40315424	4B	40315424
	BLINK	S5A_621373206	5A	621373206
	BLINK	S6B_658140124	6B	658140124
	BLINK	S7A_42614676	7A	42614676
	BLINK	S7A_510433772	7A	510433772
	FASTmrEMMA	S2A_722857568	2A	722857568
	FASTmrEMMA	S3B_773516625	3B	773516625
	FASTmrEMMA	S4B_40315424	4B	40315424
	FASTmrEMMA	S7A_27378888	7A	27378888
	FASTmrMLM	S2A_722857568	2A	722857568
	FASTmrMLM	S2B_662259821	2B	662259821
	FASTmrMLM	S3B_773516625	3B	773516625
	FASTmrMLM	S4B_40315424	4B	40315424
	FASTmrMLM	S7A_27378888	7A	27378888
	FarmCPU	S2A_722857568	2A	722857568
	FarmCPU	S2B_725552556	2B	725552556
	FarmCPU	S3B_773516625	3B	773516625
	FarmCPU	S4B_40315424	4B	40315424
	FarmCPU	S4B_647586119	4B	647586119
	FarmCPU	S7A_48708273	7A	48708273
	ISIS EM-BLASSO	S2A_722857568	2A	722857568
	ISIS EM-BLASSO	S2B_662259821	2B	662259821
	ISIS EM-BLASSO	S3B_773516625	3B	773516625
	ISIS EM-BLASSO	S4B_40315424	4B	40315424
	ISIS EM-BLASSO	S5D_551464583	5D	551464583

	ISIS EM-BLASSO	S6B_96579342	6B	96579342
	ISIS EM-BLASSO	S7A_27378888	7A	27378888
	mrMLM	S2A_722857568	2A	722857568
	mrMLM	S2B_662259821	2B	662259821
	mrMLM	S3B_773516625	3B	773516625
	mrMLM	S4B_40315424	4B	40315424
	mrMLM	S6B_96579342	6B	96579342
	mrMLM	S7A_27378888	7A	27378888
	mrMLM	S7A_3972877	7A	3972877
	pKWmEB	S2B_789008561	2B	789008561
	pKWmEB	S4B_40315424	4B	40315424
	pKWmEB	S4B_642604572	4B	642604572
	pKWmEB	S6B_663525758	6B	663525758
	pKWmEB	S6D_70342	6D	70342
	pKWmEB	S7A_27378888	7A	27378888
	pKWmEB	S7A_510433772	7A	510433772
	pLARmEB	S2B_662259821	2B	662259821
	pLARmEB	S2B_789008561	2B	789008561
	pLARmEB	S4B_40315424	4B	40315424
	pLARmEB	S4B_642604572	4B	642604572
	pLARmEB	S5D_551464583	5D	551464583
	pLARmEB	S6B_634018818	6B	634018818
	pLARmEB	S6B_663525758	6B	663525758
	pLARmEB	S7A_27378888	7A	27378888
	pLARmEB	S7A_510433772	7A	510433772
FDK	BLINK	S5A_39854554	5A	39854554
	BLINK	S5A_619020400	5A	619020400
	BLINK	S6B_718194425	6B	718194425
	FASTmrEMMA	S2A_748396092	2A	748396092
	FASTmrEMMA	S3A_568391268	3A	568391268
	FASTmrEMMA	S3B_771956508	3B	771956508
	FASTmrEMMA	S6B_716336898	6B	716336898

FASTmrMLM	S2A_748396092	2A	748396092
FASTmrMLM	S3A_12668487	3A	12668487
FASTmrMLM	S3A_568391268	3A	568391268
FASTmrMLM	S3B_771956508	3B	771956508
FASTmrMLM	S5A_618308582	5A	618308582
FASTmrMLM	S6B_716336898	6B	716336898
FASTmrMLM	S7B_707550430	7B	707550430
FarmCPU	S3B_768314878	3B	768314878
FarmCPU	S5A_619020400	5A	619020400
FarmCPU	S6B_320696398	6B	320696398
FarmCPU	S6B_718194425	6B	718194425
FarmCPU	S6D_110313864	6D	110313864
FarmCPU	S7A_713432647	7A	713432647
FarmCPU	S7A_738859192	7A	738859192
FarmCPU	S7B_707550430	7B	707550430
ISIS EM-BLASSO	S2A_748396092	2A	748396092
ISIS EM-BLASSO	S3A_530215377	3A	530215377
ISIS EM-BLASSO	S5A_39854554	5A	39854554
ISIS EM-BLASSO	S6B_716336898	6B	716336898
ISIS EM-BLASSO	S7B_472484704	7B	472484704
ISIS EM-BLASSO	S7B_707550430	7B	707550430
mrMLM	S2A_748396092	2A	748396092
mrMLM	S3A_12668487	3A	12668487
mrMLM	S3A_568391268	3A	568391268
mrMLM	S3B_771956508	3B	771956508
mrMLM	S5A_618308582	5A	618308582
mrMLM	S6B_716336898	6B	716336898
mrMLM	S7B_707550430	7B	707550430
pKWmEB	S2A_748396092	2A	748396092
pKWmEB	S3A_13811558	3A	13811558
pKWmEB	S3A_528989206	3A	528989206
pKWmEB	S5A_39854554	5A	39854554

pKWmEB	S6B_716171099	6B	716171099
pKWmEB	S7B_472484704	7B	472484704
pLARmEB	S1A_10439185	1A	10439185
pLARmEB	S1B_595397803	1B	595397803
pLARmEB	S2A_748396092	2A	748396092
pLARmEB	S2A_781672337	2A	781672337
pLARmEB	S2B_662259821	2B	662259821
pLARmEB	S3A_568391268	3A	568391268
pLARmEB	S3B_621303861	3B	621303861
pLARmEB	S3B_771956508	3B	771956508
pLARmEB	S4B_642901115	4B	642901115
pLARmEB	S5A_11637897	5A	11637897
pLARmEB	S6B_716336898	6B	716336898
pLARmEB	S7A_50507497	7A	50507497

[a]Physical position is based on IWGSC RefSeq v2.0 (IWGSC, 2018)

Supplementary Table S3. Table summarizing the GWAS results for plant height (PH) using the FarmCPU model. The threshold used to declare an association as significant was FDR adjusted $P < 0.05$.

SNP	Chromosome	Position[a]	*P* value	MAF	FDR adjusted *P* value
S2A_614445772	2A	614,445,772	4.8E-06	0.15	1E-02
S4B_40019966	4B	40,019,966	8.76E-13	0.18	8E-09
S5B_295150730	5B	295,150,730	1.23E-06	0.46	6E-03
S7B_9558646	7B	9,558,646	2.21E-06	0.38	7E-03

[a]Physical position is based on IWGSC RefSeq v2.0 (IWGSC, 2018)

Supplementary Table S4. Table enlisting all high-confidence genes identified in the flanking region of two QTNs, *S6B_718194425* and *S7B_707550430* along with their annotation.

Gene ID[a]	Previous Gene ID[b]	Start[c]	Annotation
TraesCS6B03G1247100	*TraesCS6B02G448800*	717691663	cytochrome P450 714C2-like
TraesCS6B03G1247900	*TraesCS6B02G448900*	717933372	zinc finger protein ZAT5-like
TraesCS6B03G1248300	*TraesCS6B02G449000*	718038855	serine/threonine-protein phosphatase PP1-like
TraesCS6B03G1248400	*TraesCS6B02G449100*	718072499	hypothetical protein CFC21_090362
TraesCS6B03G1248500	*TraesCS6B02G449200*	718081314	rop guanine nucleotide exchange factor 3-like
TraesCS6B03G1248600	*TraesCS6B02G449300*	718084520	FHA domain-containing protein DDL-like
TraesCS6B03G1248900	*TraesCS6B02G449400*	718124729	nicotinate N-methyltransferase 1-like
TraesCS6B03G1249200	*TraesCS6B02G449500*	718134558	disease resistance protein RGA5-like
TraesCS6B03G1249300	*TraesCS6B02G449600*	718142603	disease resistance protein RGA5-like isoform X1
TraesCS6B03G1249400	*TraesCS6B02G449700*	718193470	NEP1-interacting protein-like 1
TraesCS6B03G1249500	*TraesCS6B02G449800*	718194677	predicted protein
TraesCS6B03G1249600	*TraesCS6B02G449900*	718351170	lysine-specific histone demethylase 1 homolog 3-like
TraesCS6B03G1249800	*TraesCS6B02G450000*	718403799	aquaporin PIP1-5-like
TraesCS6B03G1249900	*TraesCS6B02G450100*	718408744	laccase-15-like
TraesCS6B03G1250200	*TraesCS6B02G450200*	718437357	aquaporin PIP1-5-like
TraesCS6B03G1250500	*TraesCS6B02G450300*	718537206	peptidyl-prolyl cis-trans isomerase-like
TraesCS6B03G1250600	*TraesCS6B02G450400*	718550171	peptidyl-prolyl cis-trans isomerase-like
TraesCS6B03G1250700	*TraesCS6B02G450500*	718634434	50S ribosomal protein L9, chloroplastic
TraesCS6B03G1251100	*TraesCS6B02G450600*	718946263	4-coumarate--CoA ligase-like 3
TraesCS6B03G1251200	*TraesCS6B02G450700*	718948307	acyl-CoA-binding domain-containing protein 4-like
TraesCS6B03G1251400	*TraesCS6B02G450800*	718975754	hypothetical protein CFC21_090380
TraesCS6B03G1251500	*TraesCS6B02G450900*	718995041	hypothetical protein CFC21_090381
TraesCS6B03G1251800	*TraesCS6B02G451000*	719407030	csAtPR5, putative, expressed
TraesCS6B03G1251900	*TraesCS6B02G451100*	719416172	signal peptide peptidase-like 4
TraesCS6B03G1252300	*TraesCS6B02G451200*	719509971	phosphatidylinositol 4-phosphate 5-kinase 9-like
TraesCS6B03G1252500	*TraesCS6B02G451300*	719516962	NAC domain-containing protein 78-like

TraesCS6B03G1252600	*TraesCS6B02G451400*	719562411	probable mitochondrial saccharopine dehydrogenase-like oxidoreductase At5g39410
TraesCS6B03G1252700	*TraesCS6B02G451500*	719564065	predicted protein
TraesCS7B03G1159300	*TraesCS7B02G416700*	706566583	uncharacterized protein LOC123161760
TraesCS7B03G1159800	*TraesCS7B02G416800*	706617114	hydroquinone glucosyltransferase-like
TraesCS7B03G1160000	*TraesCS7B02G416900*	706684670	hypothetical protein CFC21_105376
TraesCS7B03G1160200	*TraesCS7B02G417000*	706703917	hypothetical protein CFC21_105377
TraesCS7B03G1160400	*TraesCS7B02G417100*	706707637	NBS-LRR disease resistance protein
TraesCS7B03G1161400	*TraesCS7B02G417200*	706838899	hypothetical protein CFC21_105379
TraesCS7B03G1161500	*TraesCS7B02G417300*	706844055	putative disease resistance protein RGA3
TraesCS7B03G1162000	*TraesCS7B02G417400*	706905895	probable LRR receptor-like serine/threonine-protein kinase At3g47570
TraesCS7B03G1163500	*TraesCS7B02G429100*	707818550	NADH-ubiquinone oxidoreductase chain 1
TraesCS7B03G1163700	*TraesCS7B02G429200*	707819633	ribosomal protein S13
TraesCS7B03G1164100	*TraesCS7B02G429300*	707825040	ribosomal protein L16
TraesCS7B03G1164800	*TraesCS7B02G429400*	707834318	cytochrome c biogenesis protein ccmFC
TraesCS7B03G1165100	*TraesCS7B02G429500*	707836552	Cytochrome c biogenesis Fc
TraesCS7B03G1165200	*TraesCS7B02G429600*	707836902	39 kDa protein in mitochondrial S-1 and S-2 DNA
TraesCS7B03G1166000	*TraesCS7B02G429700*	708194926	hydroquinone glucosyltransferase-like
TraesCS7B03G1167100	*TraesCS7B02G429800*	708341181	uncharacterized methyltransferase At2g41040,
TraesCS7B03G1167200	*TraesCS7B02G429900*	708370570	AAA-ATPase ASD, mitochondrial-like
TraesCS7B03G1167300	*TraesCS7B02G430000*	708543954	uncharacterized protein LOC119341039
TraesCS7B03G1167400	*TraesCS7B02G430100*	708558030	uncharacterized protein LOC123162851
TraesCS7B03G1167600	*TraesCS7B02G430200*	708568667	putative disease resistance RPP13-like protein 1 isoform X1

[a]Gene ID based on the IWGSC RefSeq Annotation v2.1 (IWGSC 2018; Zhu et al. 2021)

[b]Previous IDs for respective genes to the IDs used in IWGSC RefSeq Annotation v1.1 (IWGSC 2018)

[c]Physical position of start points for respective genes are based on IWGSC RefSeq v2.0 (IWGSC 2018)

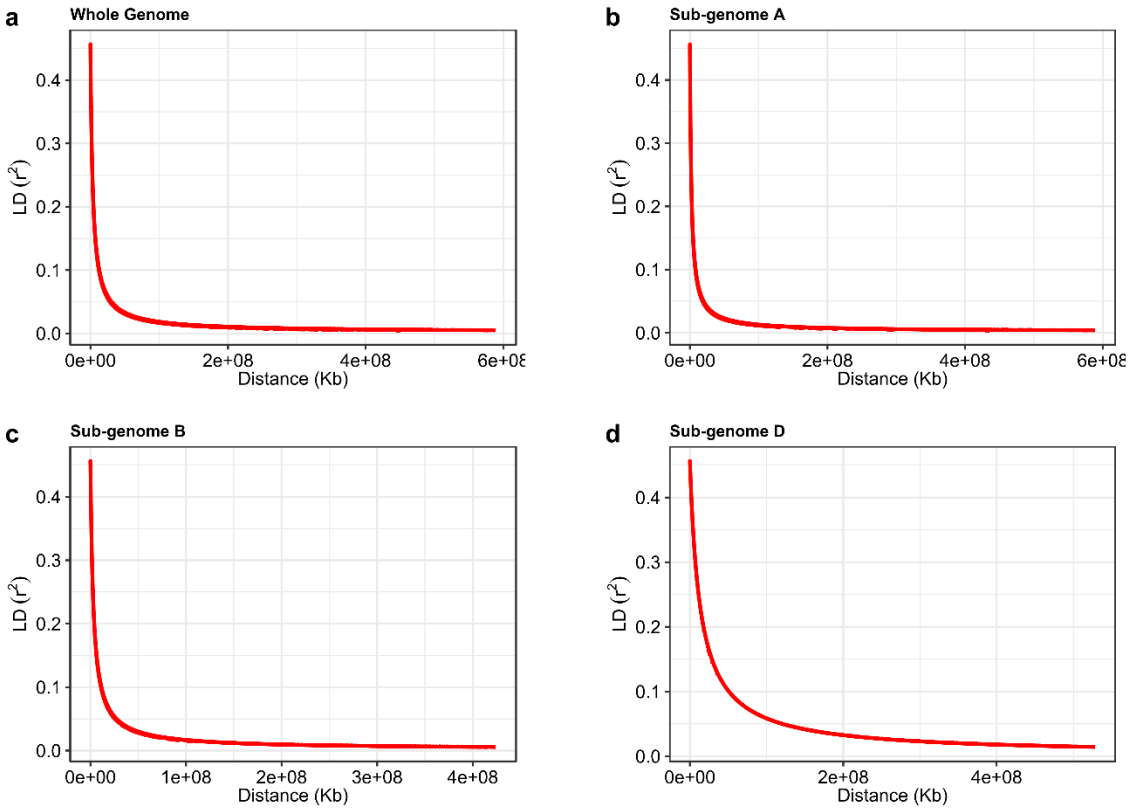
Supplementary Table S5. Marker sequences for the significant SNPs associated with DIS and FDK along with the physical position corresponding to Chinese Spring RefSeq v2.0 and RefSeq v1.0 (IWGSC, 2018).

SNP	Sequence	RefSeq v2.0	RefSeq v1.0
S2A_722857568	GTTCTAACCTCTAGTGTAAGGTTACATCCTTACAGGGCAGAGCACGTTGTTCATCACAAGGTT TATAACTCACATACTCATTAAACCTACAAGGTTTTGAGCGCGCAACACTGTGTGGCAATTTGG AAACATCATGAAAAACTGAAGAATCTTGGCAGCTGCAGGGAAACAAACACACATTCATGCA AATGTAAGGATACGCCCCCGCGTCGCCCCCGCTCGGGCGACTCGGGCGGGATCTAAAACCCT AGCCGCCAGGGGTCCCCATCCGTCCTCCCCTCCCTCCGCCGCTGCCGAAGGACGCCGCCGGCT ATAGGCCAGGGCCAGGGGTGGATCCAGAATTGAGCCTCGCCCCGGGCCCCAAGCTTACTACA GTGTCAGCTCAGTGCTAAACAGTATC	722857568	718979885
S2B_725552556	ACTGAAGTGACTCCGTGAGGCACAGCCGAATCGTTGGGCCGTTCGAATTTGGAACCGTGTGT GGCGTCTACCAACACGAGTATGTTCGCGCTTGACACGCGCGTACGTATGTCCACGGCCGGTA ACTGAGATGTTGTCCCGATCGACGCGTCCACGCACGCACGTACCAAGCTAGCTGTTGAATTA ATGATCGGGTGCCGCTGCGCTGCTTTATATGTAGAGTATTCGTGTACGTACAGGACAACACA ATGGCTGCAGATTCAGTTATGCAAAAAAAAAGGGTGAACTTTTCCTCTGTTGCGGTGGTATCCG CGAGATGGAGGTGGGTACGATCTGGTGCACGACGCGAAAGCGACCGTTGGCTTCACTAATCA TGTGATGATGAAAACGTGTGCAGTACA	725552556	717127511
S3B_773516625	AGGGTGATCTAGATTACGTCCGGTACTATGACGAGATGTCAAGTGAAATAAAGGTAGGCATG ATGTGATCAGCTTTTGAATTAAACTGCTAATCATGTACTAGCATATTGGTATAACAGTGTTCT TATCATGTGACAGTGGATTGAAGATAAGGTTGGACTATACAAGAAGGAGGTAAGCATCCCCT TTTGGTTAACTATTAAACAATATCTGCAGGTTCTGTAATGTCTAACATCACACATGACAATTA ACTATGTATTTTCATCATATAGTGACGGAAATACGAGAGCAGGGCTTTTCTTCAAGCAATGAA GATTGGATCAGGGTTTCATAATCTATTCAGACATTCAGTTACGAGAGCCTACAACGTATTGCG CCATGCCCCTCCCTCGCCCATAAA	773516625	758202918
S4B_40315424	TAGTAGTATCATCTTGATTCTTGAGCCTAGAGCAGCAAAGTTTGAAGTTGAAAGTCTGCCTGG AAACTACGCGTGTCCCCTCTGTCCGGCCGTGTCTGGCATGACAAACATGCATGCAGCACATGT CGGCGAATGGTTTTGTAGTACACCGTATCCTCTATACTACTCGTAACACGTGAACCATCTTGC GAAGAAGATATACTACCACTGCAGTCTACAGTAGTATAGGAGTATCAAGGATTTGCTGAGCT GTAAAAGTGCTGCATCTCCTAGTAGCTTGCAAAGTAGCATATATACAATGGGCAGCTGGTGG TCGGGGCCAGAGCACACAAGTTTCTTCATGTTCACACTGCATGCATGCAGAGTTGGTTCGTCA CCCCTTTCATGCATATACATATAC	40315424	37575130
S4B_647586119	TACTCGACTACTGCGCGGCCAATCAAGTTCCCCTGTGGAAGCACATAATGTTTATGTTTAGTA GAAGAACAACACATTTCATAGATTGCGTTATATTTTTTTCACAGGTTGTGTTGTAGAAAAATG CCAACAGCATATATATACTTATATACCTGCGACAGACCAACAATGTTATAGCCGCCGCTTAGC TCTTTCATCTCCTTTACCTGCAGAAACAAATTAGCAAGCAGGCATGGATGATGAATATCTATC	647586119	648502365

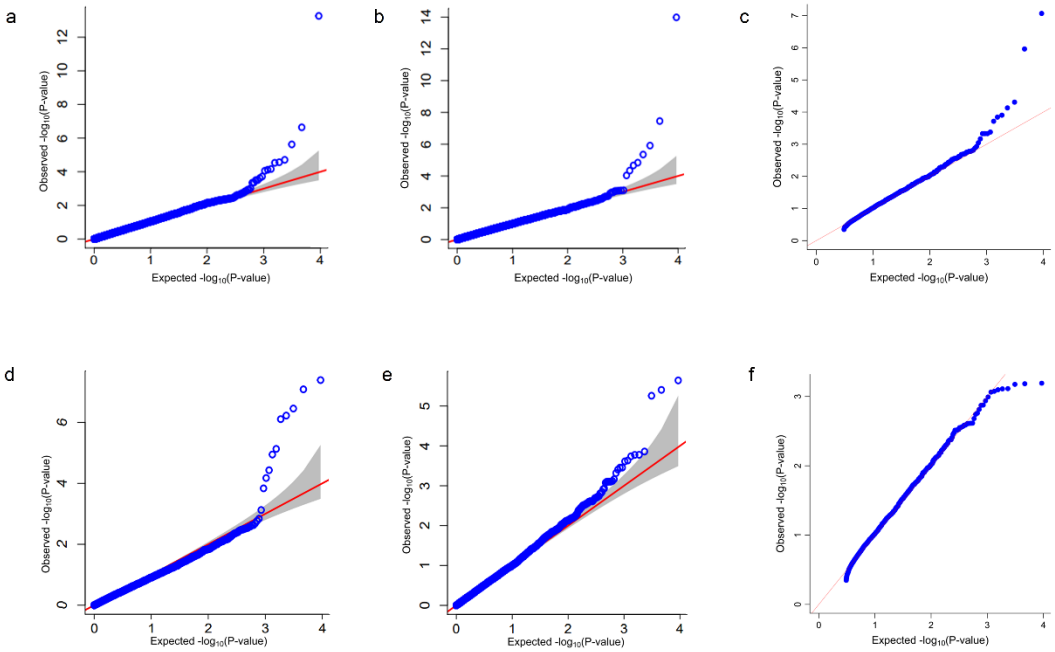
	TATCTTCTCAGCGTACAGCAGAGGAAAAGGCATCATCATCGCCCAAAGCGCCGAACCGATCA CCTTGTTGCACATGATATCAGCCTGTTCCTGTAGAGGCATGAGCCAGGAGTCCCGTGCCCCGT TTCCAATTTCTCTGAACACAAGC		
S7A_48708273	GGATTTTTTAAATGAATTGCTTAGTCCAGGATCTCTGGCCTCAAATAATTAGCGATTTTATGC AGAGGCCGGGTGTGTGCTCATTGTGTTTGTATTCCTTTGATGCTCCATTCTGAGCAAATAAAA TCCACCATTTATCTAAGGAAAAAACTAATTAGCGATTTTAGCAAGTCCAACATGTCCAACAAT GTTAAAATTTCACTTATTCTCTGCCTCGATAGACAGTGGCTGCTGCAGCATCAGGACACGCAG AAGAAGCATCGTCCATATTCAGGTTCAAACATGCAAAAACACCTTTCTACTTTGGTCATAATA GTCTTCGAAAAGAACAGAACCATCCCAGCGATCTAGCATGTGGTTTTCACTTTTCACAGAGA GGCACCAACGAAATCCTTGCTCC	48708273	46257001
S3B_768314878	TAATATGATCATGAACATGGGTGACCAAAATTCAAGAAATCCAAATAGTACACTGAATCAAT AGTAGCACGTCAGTTGTGTACATGTTACGATGAAGATTTATCATCGAAAACAAAAAAAAAAT CTCATTAGGAGAAAGCTTAGCTGCAGAACTCACATAGGCCTTACCAAAGGTTAGTGGACTTC ACAACCCTTGGACCAACCTCATGTCACAAGGCGCTGCATTTGGACATTATGTAGTTAGTGTGC GGCGAGTTAGTACTGGACACATGTGAAGAACCGTGAGACGTGAATCAATAGCACACTGAATC AATGTATCAAACCTGAAGCAGCACGTCATTGTGTGAAGAAAGACCTCCGGCCTCGCGCATGC GTGCGTGAAGGTGGCCTGTCCCCTGAC	768314878	752885972
S5A_619020400	CCGGAGACGAAGCAGAAGATGCGTGCGGTTCGACATCGCGCGGGGCTCTCAGACCTCCCTCT CGTGTGCCGCGGACCACCTGCTCAGGAGACCTGCATCTGTTTTATAGGAAGCAATCCATTCCA CGACCACGACGTCCGAAGAAAACAAAGGAGAAACAACCCAGGAGTGCAACCCAAGGAAAAT GTACATGACATCAGAGCTGCAGAGGAAAACGCCCTGAACTTCCCAGCGCCCAATGCTTCTGC ATCAGCCTTGATCTGATCAATGACACAATTATCTTGACCGATTGGAGCAATCAACAGGTGTTA ATGCAACCAGTCAATCAACCAGGTTTTGATGGGGTTTTGGAAATGTAATCGGAGGGAGACTT GGAAAGAGTCGGGTCTTCGACGGGGAT	619020400	617251477
S6B_320696398	CTCCAGAGGGGAGTAATCAATCATATTAATTACAAGAATATTTCCAAGTTTGGTCTTGCCTGT TTATCTGGTGAACCTACTAACTTCCGTGAAGCCCTTGGTGATGCTCGCTGGAAACAAGCAATG AGGAGGAATATTCTGCCCTGCAGAAAAATAAGACCTGGCATTTGGTTCCTCGTCAGGGAGGT AAAAACTTGATCGATTGCAAATGGGTCTATAGGACTAAGAGAAAATCAGATGACACGGTTGA TCGTTATAAAGCAAGGCTCGTCACAAAGGGATTTAAGCAAAGATACGGGATTGGTTATGAGG ACAATTTTAGTCCGGTTGTCAAAGCAGCATCCATTCGTCTTATGTTATCCATTGTTGTTTCCAG GGGATGGAGTCTCAGACAGCTAGA	320696398	314551477

S6B_718194425	TGTCGCTCCTCTTCTGCTGGAGCCACCGCGTTCCTCAAGAACCTTACCCTAAGCAATCCCTGC AAAACCAGACAATTTTCAGACCGGCACCCTCGACATTTCTGATAGCTCATTCCTGTAATGCAA CACGGTCAGCTTCACATCTCACCTGAATGCGGCTGCGGCTGCGCTGGGGGAACTGCGACACA ATGTGGTACCCCGATAGCCCCTGCAGCTCGCGCTGACGCTCCCGGGACATCCGAGTGATGAC ATCGGTGCGCGCCTGCCGGCCACGGAGCCGGGGAGGCTCCCCTCGCCTCTCCCTCTCCGCGG CGGTCGGCTCCTCCCTGCGGGAGGCCGCGCGGGCGTCGCGCGGCTGGCTGGCCATCTGCACC CACTCCCTGACGAGCCTGACCCTCTG	718194425	708786003
S6D_110313864	AACATATTTTTTTATGCGTACTTTTTTTGAAAAAACAATTATCCCACAGCGAGCTAATTAGCT CCACCCACAGTGGGCCGACAGGGGCCAGTCGGCCAGCTGTAGGCCGACTGGAGTCCAATCGG CCTGTTGTGGGCCGACTAGGCCCAGTCGGCCTGCAGCGGGCCGACGGTGTATTATTTTTGAAA AAAATTATCCAGCGTAATATTTATGAAAATTAATAAAAAAAGCATTATTTAAAAAAAATTAG CCCCGCCTCGCGACTTACAGCCAACTGACATAATTTGGACTGACAGCCCCCGGTTTGACCGA ATTAATTACTAGTAGAGTATTATTTTTCTTTTGCGAATGAGAGAATAAATTGCCTTTAACCAC TAATCTACTATTATTTCACGTTCCA	110313864	88831444
S7A_713432647	GCGGTCCTTGCACAGTTAATGAGGTCACGTCCATGGCATACGCTACCCCTCTTTGCTTGCCAG CCTTGTCTGCGCGAGTGCAGCAGAGGTGCAGTCCGTCCATGAGAGCCCGGTATTGCGCCGAT GGCTGGACCAAAGACGAAACCACTGTGTTGGCCTGGTGTTCCCAACGTCTTTCCCCGCACCA AGTTTGGATTCCGTTGTGGATTTCATTACTCTACTCGCTTATGTTGTTTGCTGCTGCTGCTGCA GCAGCAGAGGGAGGGTGAGAGAGACCCAGCCAAGCCAAGGCGCCGGGTTCCCACGTTACCA AGAACCTGTAAATATTTGCCATCACCCATCACTATACATGCCATTGCTTTGTTTGTCCTCGTAT ATTCAGTCTAACAAACAACAAATT	713432647	707834631
S7A_738859192	TTTGAGAAATCCTAACTCCGCCTCTGGTGATGAGGACCGGCACGAACGGTGGCGTCCTCGCG CCTAGGGAGGCAGCGCCACGCATACCACATAAAAGTCGACGCGTCGGGACGGCGGAGTGGA CCACGCGCCGCGACGCTACGGCCTGAAGGGGCGACGCAGCGGCAGCGCGCGGGTCGGGGCG ACGGCGAGGAAAACCACTGGGCGGCTGCAGGGACCATGTGCTACGCTCGCTAGCGCCCGAC GAGGCGACGCAGCGGCAGCGTGAGGGTCGGTGCAGCCTCGGGGACGGCCGGAAGCGGACGG CCTCGGGGCGGTGGTTGGCCGCGGGCTGCGGCACTACAGCCCGAAGGAGCATGCAGTGCCAG CATGCGGGTTGGTGCAACCACGGGAACGGCCT	738859192	731888363
S7B_707550430	CACCTTGCCCTAGCCGCCGCCCCCCTTCTTCTCCCTCGCGACGGCGCCAACACGCCGGAGCCG CCGCCCTCGCCCCCAACCCTCGCTGTTCTGCCCGTATAACCTACGGAGTAGAGCCGGTAGGA ACCCTAGTCCTACCATAGCCAGACAGAGGCAAGTCGCCATAATGAAAAGAGACGTGAAAAA CTGAAGTAATACTGAATTGGTACCTTTGTGCTGCAGGAGTCCTTATACTAGTAGTATATATAG CAGTATAACAAAATCCTAACTTAATAAGAAGAAAGAAACTCAAAGGTATTGGTAGTGAGAA CCAACATACTAGCTGCTTTCTTCCATCCTAGTACAAAAACGATGCACAAAAAAGGAAGGGGG GAAAACTGGGTGCCCCCTCAACTTCATC	707550430	698229993

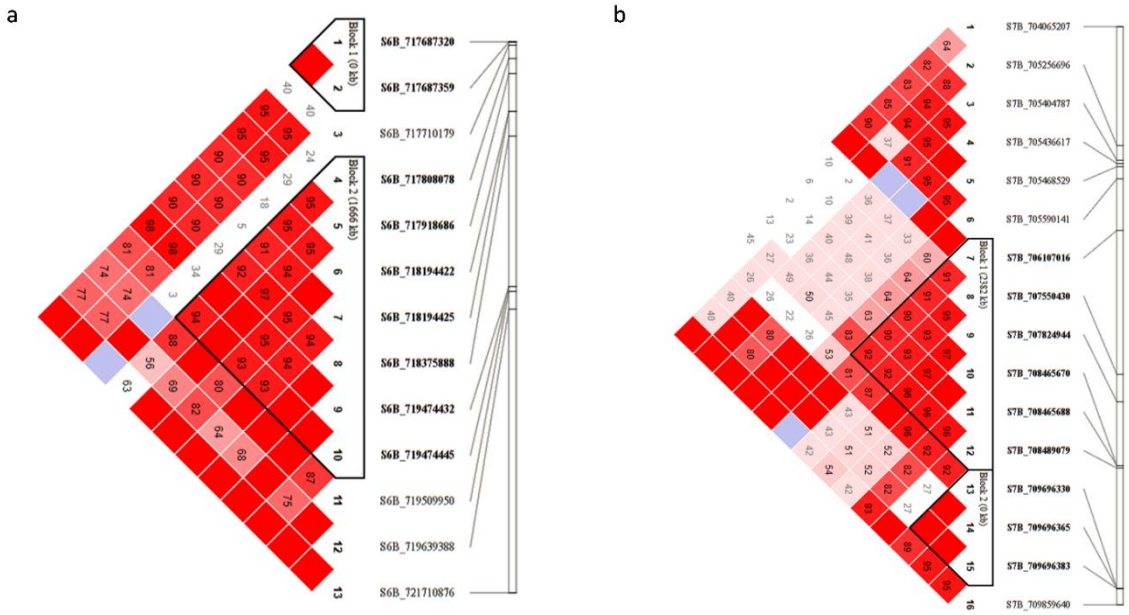
Supplementary Figure S1. Intra-chromosomal linkage disequilibrium (LD) in the panel of 257 lines.



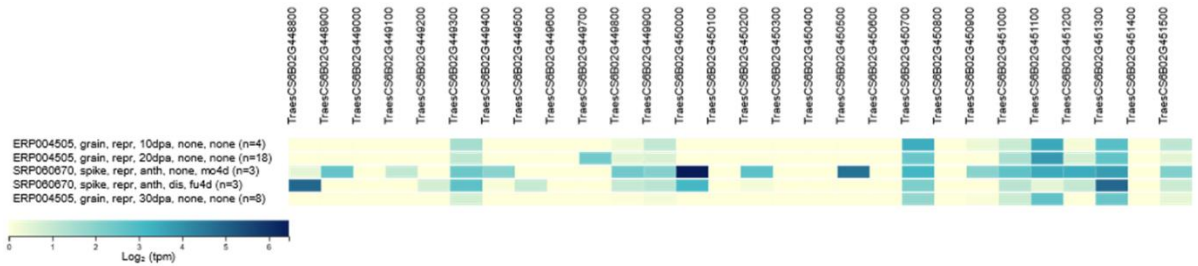
Supplementary Figure S2. QQ plots comparing the model performance of various ML-GWAS traits. The upper panel compares different ML-GWAS models for DIS, where (a) QQ plot using FarmCPU for DIS, (b) QQ plot using BLINK for DIS, and (c) a combined QQ plot for DIS obtained from six ML-GWAS models including mrMLM, FastmrMLM, FastmrEMMA, pLARmEB, pKWmEB, and ISIS EM-BLASSO. The lower panel shows the comparison of QQ plots for FDK, where (d) QQ plot using FarmCPU for FDK, (e) QQ plot using BLINK for FDK, and (f) a combined QQ plot for FDK obtained from six ML-GWAS models including mrMLM, FastmrMLM, FastmrEMMA, pLARmEB, pKWmEB, and ISIS EM-BLASSO.



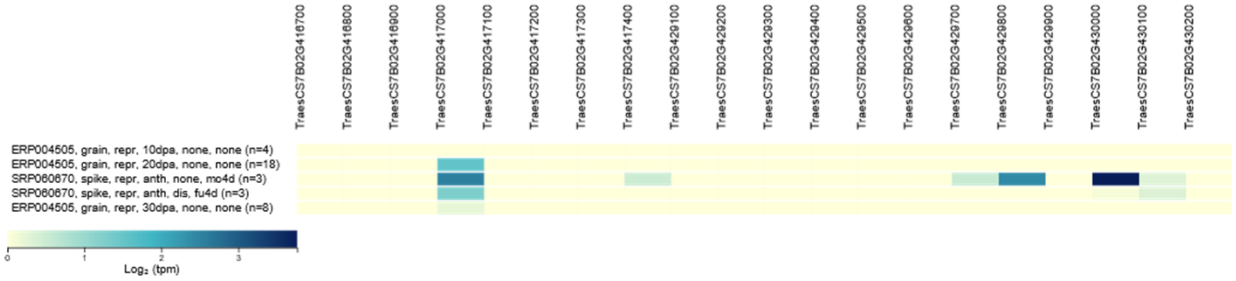
Supplementary Figure S3. Local linkage disequilibrium (LD) block for region harboring QTNs for FDK, (a) *S6B_718194425* and (b) *S7B_707550430*.



Supplementary Figure S4. Gene expression analysis for high confidence genes in the flanking region of QTN S6B_718194425 across several studies for time courses of Fusarium infection. Gene expression is presented as a heatmap with Gene IDs based on IWGSC RefSeq v1.1 are listed on the top and the stages/tissues of expression on the side.



Supplementary Figure S5. Gene expression analysis for high confidence genes in the flanking region of QTN S7B_707550430 across several studies for time courses of Fusarium infection. Gene expression is presented as a heatmap with Gene IDs based on IWGSC RefSeq v1.1 are listed on the top and the stages/tissues of expression on the side.



Chapter 4. Genomic prediction of Fusarium head blight resistance in early stages using advanced breeding lines in hard winter wheat

4.1 Abstract

Fusarium head blight (FHB), also known as scab, is a devastating fungal disease of wheat that causes significant losses in grain yield and quality. Quantitative inheritance and cumbersome phenotyping make FHB resistance a challenging trait for direct selection in wheat breeding. Genomic selection (GS) to predict FHB resistance traits has shown promise in several studies. Here, we used univariate and multivariate genomic prediction models to evaluate the prediction accuracy (PA) for different FHB traits using 476 elite and advanced breeding lines developed by South Dakota State University hard winter wheat breeding program. These breeding lines were assessed for FHB disease index (DIS), and percentage of Fusarium damaged kernels (FDK) in three FHB nurseries in 2018, 2019, and 2020 seasons (TP18, TP19, and TP20) and were evaluated as training populations (TP) for genomic prediction (GP) of FHB traits. We observed moderate PA using univariate models for DIS (0.39 and 0.35) and FDK (0.35 and 0.37) using TP19 and TP20, respectively, and slightly higher PA (0.41 for DIS and 0.38 for FDK) when TP19 and TP20 (TP19+20) were combined to leverage the advantage of a large training population. Although GP with multivariate approach including plant height and days to heading as covariates did not significantly improve PA for DIS and FDK over ST models, PA for DON increased by 20% using DIS, FDK, DTH as covariates using MT model in 2020. Finally, we used TP19, TP20, and TP19+20 in forward prediction to calculate genomic-estimated breeding values (GEBVs) for DIS and FDK in preliminary breeding lines at an early stage of the breeding program. We observed moderate PA of up

to 0.59 for DIS and 0.54 for FDK, demonstrating the promise in genomic prediction for FHB resistance in earlier stages using advanced lines. Our results suggest GP of expensive FHB traits like DON and FDK can facilitate the rejection of highly susceptible materials at an early stage in a breeding program.

Keywords: FHB, GBS, genomic selection, multi-trait models, winter wheat, wheat scab

4.2 Introduction

Several fungal pathogens continuously constrain global wheat production and food security. Fusarium head blight (FHB), also known as scab, is a devastating fungal disease of wheat that causes significant losses in grain yield and quality (McMullen et al., 2012). FHB is expanding its horizons throughout major wheat-producing areas due to climate change, an increased wheat-growing area under no-till cultivation, and adoption of maize-wheat rotations (Zhang et al., 2014; Singh et al., 2016). Though several *Fusarium* species can cause FHB, *Fusarium graminearum* is the prominent pathogen for FHB in the United States, Canada, China, and some European countries (Trail, 2009). In addition to grain yield losses, FHB results in reduced quality and contamination by mycotoxin, primarily deoxynivalenol (DON) that poses serious health consequences to humans and animals if ingested in certain quantities (Trail, 2009; Sobrova et al., 2010). In 2014, the revenue losses for hard wheat were estimated to be around $600 million in the Great Plains region of the US (Wilson et al., 2017).

Even though fungicides are frequently used to reduce FHB damage, the utilization of resistant varieties is considered the most effective and economical approach to combat diseases like FHB (Bai and Shaner, 2004; Gilbert and Haber, 2013). Being quantitative in nature, resistance to FHB is governed by multiple quantitative trait loci (QTLs) and highly influenced by changing environments. Resistance to FHB has been categorized as type I (resistance to initial pathogen penetration), type II (resistance to spread of FHB symptoms within a spike), or type III (low mycotoxin accumulation) (Bai et al., 2018). However, type II resistance is more stable and utilized in breeding programmes as compared to type I and III FHB resistance (Bai and Shaner, 2004; Bai et al., 2018).

Conventional QTL mapping and genome-wide association study (GWAS) have been used to dissect the genetic basis of FHB resistance and a large number of QTLs have been identified across all the wheat chromosomes, including seven named QTLs, *Fhb1* to *Fhb7* (Liu et al., 2009; Halder et al., 2019; Su et al., 2019; Venske et al., 2019). However, only a handful of QTLs have a major effect on type II resistance (resistance to FHB symptom spread in a spike) and have been effectively utilized in wheat breeding globally, in particular *Fhb1* (Steiner et al., 2017; Bai et al., 2018). In the US hard winter wheat region, most of the variation in FHB resistance are from native sources including cultivars 'Everest', 'Overland', 'Lyman', and 'Expedition'. Furthermore, phenotypic selection for FHB resistance in both field and greenhouse is very complicated and some of the measurements such as DON content can only be obtained after harvest and are costly. Thus, genomic selection (GS) could be a promising approach to improve FHB resistance in wheat with reduced phenotyping efforts and costs.

GS is an approach that employs linkage disequilibrium (LD) and estimates the genetic worth of an individual using genome-wide markers (Meuwissen et al., 2001; Heffner et al., 2009). GS addresses the primary limitation of QTL mapping and marker-assisted selection by using a joint estimate of all marker effects (Bassi et al., 2015). Thus, GS is useful for predicting and selecting complex traits controlled by several minor QTLs that are difficult to map using QTL mapping (Lorenz et al., 2011). The genomic prediction (GP) models are developed by using genotypic and phenotypic data in a training population (TP) to predict the genomic-estimated breeding value (GEBVs) of individuals in the breeding population (BP) (Meuwissen et al., 2001). GS has shown immense potential in plant breeding, and several studies have reported successful implementation

of these strategies in different crops in recent years (Poland et al., 2012; Bhat et al., 2016; Juliana et al., 2017). GS is particularly useful for these traits where phenotyping is cumbersome or costly (Battenfield et al., 2016; Gill et al., 2021).

Predictive ability (PA) of the GS model refers to the correlation between estimated GEBVs and the actual phenotypic values of the individuals in the validation set. The PA mainly depends on heritability of the traits, TP nature and size, and choice and optimization of the statistical models (Ali et al., 2020; Gill et al., 2021). Several studies have evaluated the GP models for predicting different FHB traits including disease index (DIS), fusarium damaged kernels (FDK), and DON in wheat. Most of these studies employed a cross-validation approach to evaluate the PA of GP in spring and soft winter wheat using various prediction models and strategies (Rutkoski et al., 2012; Arruda et al., 2015; Mirdita et al., 2015; Hoffstetter et al., 2016a; Dong et al., 2018; Larkin et al., 2020). However, the focus of these studies was limited to improving the PA within the TP being evaluated, rather than validating the improved models in forward prediction. Unlike other traits, such as yield, only a few studies reported the implementation of GP to select preliminary breeding lines for FHB resistance (Schulthess et al., 2018; Verges et al., 2020). So far, the potential of GS in improving the FHB resistance in early generations of a hard winter wheat breeding program has not been examined.

Most of the previous studies compared univariate GP approaches, including ridge-regression best linear unbiased prediction (rrBLUP), and genomic best linear unbiased prediction (GBLUP), LASSO, Random Forest (RF), and several Bayesian approaches (Rutkoski et al., 2012; Arruda et al., 2015; Dong et al., 2018). In most cases, the performance of these GP models varied with FHB traits and cross-validation schemes

used in the analyses. On the other hand, multi-trait (MT) approaches are used to improve the prediction ability for a primary trait, when secondary traits genetically correlated to the primary trait are available (Jia and Jannink, 2012). MT models are of particular importance when the primary trait is difficult or expensive to phenotype and has low heritability. Plant height (PH) and days to heading (DTH) are often associated with FHB resistance in bread wheat and durum wheat (Schulthess et al., 2018; Larkin et al., 2020). Thus, several studies have used PH and DTH as covariates to predict FHB resistance (Steiner et al., 2017; Schulthess et al., 2018; Larkin et al., 2020; Moreno-Amores et al., 2020), and only two studies suggested improvement in PA using PH or DTH as covariates (Schulthess et al., 2018; Larkin et al., 2020). Furthermore, different FHB traits, including DIS, FDK, and DON are known to have moderate positive genetic correlations (Gaire et al., 2021). Recent studies have evaluated the use of different combinations of FHB traits along with DTH and PH as secondary traits to utilize this genetic correlation for improving prediction ability of MT models (Larkin et al., 2020). For instance, FHB traits like DIS and FDK can be used as secondary traits for improving the prediction of DON, which is cumbersome and costly to phenotype. Thus, there is a need to evaluate the usefulness of these covariates or to figure out the best combinations of secondary traits in MT models to predict FHB traits, especially in hard winter wheat.

Another aspect of the successful application of GS in a breeding program is establishing a training population. Previous studies have evaluated strategies to optimize training populations using advanced breeding lines to predict GEBVs of preliminary breeding lines for various traits in wheat, including FHB (Hoffstetter et al., 2016a; Belamkar et al., 2018; Adeyemo et al., 2020; Verges et al., 2020). The application of GS in this scenario

can be handy as it is challenging to phenotype a large set of preliminary lines in expensive FHB nurseries.

The primary objective of this study was to use different sets of advanced breeding lines as training populations to predict FHB traits in preliminary breeding lines of our hard winter wheat breeding program. Secondly, we wanted to evaluate the performance of MT models to predict FHB traits, including DIS, FDK, and DON, using different combinations of secondary traits. For this, we evaluated FHB nurseries comprising advanced breeding lines from three years for their usability as training populations to predict these traits. Thus, our specific objectives for this study were to (a) evaluate the usability of FHB nurseries comprising advanced breeding lines for predicting FHB traits and assessing the improvement in predictive ability if the best-performing TPs from individual years were combined based on the lines shared among the FHB nurseries over years, (b) compare predictive abilities of MT models when different combinations of secondary traits are used to predict DIS, FDK, and DON, and (c) to validate the selected models and TPs in a forward prediction scheme to calculate the GEBVs for FHB traits in an independent breeding population comprising the preliminary breeding lines.

4.3 Materials and Methods

4.3.1 Plant materials

A panel of 476 wheat breeding lines from the winter wheat breeding program at South Dakota State University was used in this study. A set of breeding lines that is tested every year in the advanced yield trial (AYT) and the elite yield trial (EYT) was evaluated for FHB resistance in a mist irrigated field nursery. Among the 476 breeding lines, 153 were

evaluated in the 2018 nursery, 169 in the 2019 nursery, and 154 in the 2020 nursery to optimize a training population for predicting preliminary breeding lines. Further, 65 lines were shared between the 2018 and 2019 nurseries; and 58 lines were shared between the 2019 and 2020 nurseries. The majority of the breeding lines were either $F_{4:7}$ or $F_{4:8}$ filial generation. Lines without genotypic data and lacking consistency between replications were excluded and the final analysis was conducted on 152, 161, and 153 lines from 2018, 2019, and 2020 nurseries, respectively. In addition, a set of 200 breeding lines from the preliminary yield trial (PYT) was used as a breeding population (BP) for the prediction of FHB indices (independent validation) using the optimized training population. A random set of 60 lines was selected from the BP and evaluated in the 2020 nursery to validate the prediction accuracy (PA) in forward selection.

4.3.2 Experimental design and studied characters

Plant materials were planted in the FHB nurseries at Brookings, South Dakota (44.3114° N, 96.7984° W) during the 2018, 2019, and 2020 growing seasons (Supplementary Table S1) using a randomized complete block design with 2 or 3 replicates for different sets of lines with corresponding checks. We used cultivars ‘Lyman’ and ‘Emerson’ as resistant checks while ‘Overley’ and ‘Flourish’ were used as susceptible checks. The experimental unit was a single-row plot (40 plants/1 meter row) for each line. FHB resistance and related traits were evaluated in 2018, 2019, and 2020. Heading date (DTH) was recorded using Julian date when 50% of the main tillers in the row had completely emerged heads. Plant height (PH) was measured from the soil surface to the top of main tiller spikes excluding awns when the plant materials matured.

All FHB nurseries were artificially inoculated using both corn-spawn and spraying a spore suspension of *F. graminearum* isolates (SD-FG1) as described in Halder et al. (Halder et al., 2019). Briefly, Fusarium-infested corn kernels were scattered on soil surface twice with one at boot (Feeks 10) and another at heading (Feekes 10.1) stages to enhance the chances of maximum spore production in the field. At anthesis, a conidial suspension containing 100,000 spores/ml was sprayed on the heads of each line at 50% anthesis to avoid any escape. The nursery was misted using sprinklers every night (7 pm-7 am) to maintain the humidity for disease development. Disease incidence and severity was recorded by scoring FHB symptoms on 20 heads/replication/line 21 days post-flowering using a visual scale described by Stack and McMullen (Robert W. Stack and Marcia P. McMullen, 2011). The FHB disease index (DIS) was calculated as (Incidence (INC) × Severity (SEV))/100. Percentage of Fusarium damaged kernel (FDK) was evaluated using grain samples harvested by a low airspeed harvester. Sampled kernels were compared against a set of known FDK standards (https://agcrops.osu.edu/newsletter/corn-newsletter/2015-21/rating-fusarium-damaged-kernels-fdk-scabby-wheat) to estimate the FDK values in two replications per sample. We used DIS data from three seasons, whereas FDK from 2019 and 2020 seasons for evaluating GP. For DON estimation, the samples were analysed Department of Plant Science at North Dakota State University using gas chromatography–mass spectrometry method. As limited number of unreplicated samples were analysed for DON during 2018 and 2019, we used DON data from 2020 season only.

4.3.3 Genotyping-by-sequencing

Fresh leaf tissues were taken from each line for DNA isolation using the hexadecyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987). The seedlings for each line in advanced or elite trials in their respective season were grown in small pots/cones for tissue sampling and DNA extraction. A genotyping-by-sequencing (GBS) library was constructed using the *PstI* and *MspI* restriction enzymes (Poland et al., 2012) and sequenced in an IonTorrent Proton sequencer (Thermo Fisher Scientific, Waltham, MA, USA) at the USDA Central Small Grain Genotyping Lab, Manhattan, KS, USA. DNA sequence data was used to call single-nucleotide polymorphisms (SNP) using the previously described approach employing TASSEL v5.0 (Trait Analysis by aSSociation, Evolution and Linkage) (Bradbury et al., 2007). After removal of SNPs with more than 30% missing genotypes, minor allele frequency (MAF) of less than 0.05 and unmapped on any chromosome, 9,321 high-quality SNPs were imputed using BEAGLE v4.1 (Browning and Browning, 2007) for further analysis.

4.3.4 Phenotypic data analysis

The phenotypic data for FHB DIS and FDK was analyzed as best linear unbiased estimates (BLUEs) for individual nurseries. The following model was used to estimate the BLUEs:

$$y_{ij} = \mu + R_i + G_j + e_{ij}$$

where y_{ij} is the trait of interest, μ is the overall mean, R_i is the effect of the i^{th} replicate, G_j is the effect of the j^{th} genotype, and e_{ij} is the residual error effect associated with the i^{th}

replication and j^{th} genotype. The lines shared between 2019 and 2020 seasons were used to combine the DIS and FDK data, BLUEs was estimated across environments using the following statistical model:

$$y_{ijk} = \mu + E_i + R_{j(i)} + G_k + GE_{ik} + e_{ijk}$$

where y_{ijk} is the trait of interest, μ is the overall mean, E_i is the effect of the i^{th} environment, $R_{j(i)}$ is the effect of the j^{th} replicate nested within the ith environment, G_k is the effect of the j^{th} genotype, GE_{ik} is the effect of the genotype x environment (G x E) interaction, and e_{ijk} is the residual error associated with the i^{th} replication and j^{th} genotype.

The broad-sense heritability (H^2) for DIS and FDK was estimated for independent nurseries as follow:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2 / nRep}$$

where σ_g^2 and σ_e^2 , are the genotype and error variance components, respectively. We used META-R (Alvarado et al., 2020) based on the LME4 R-package (Bates et al., 2015) for the linear mixed model analysis and heritability estimation. The Pearson correlations among traits and environments were estimated based on the BLUEs for each trait using R environment (R Core Team, 2018).

The principal component analysis (PCA) was conducted using the genotypic data from 457 lines (257 lines from 2019 and 2020 nurseries and 200 lines from the breeding population) to study the relationship between the training and breeding populations. The '*prcomp'* function in R was selected to perform PCA using 9,321 SNP markers, and the first two principal components were used for the scatterplot.

4.3.5 Genomic prediction models

The univariate genomic prediction for DIS and FDK was performed using four different algorithms. The ridge regression (rrBLUP) model (Endelman, 2011) is the widely used GS model in plant breeding. In rrBLUP, we assume a normal distribution of marker effects with equal variance. The GEBVs for DIS and FDK were estimated for each training population using the trait BLUEs. A linear mixed model was implemented using the following model:

$$y = 1\mu + Zu + e$$

where y is the vector ($n \times 1$) of adjusted means (BLUEs) from n genotypes for a given trait; μ is the overall mean; Z is the design matrix $(n \times p)$ with known values of p markers for n genotypes; u is a genotypic predictor with $u \sim N(0, G_{nxn}\sigma_g^2)$, where G is positive semidefinite matrix, obtained from markers using *'A.mat'* which is an additive relation matrix function and σ_g^2 is the additive genetic variance; ε is the residual error with $e \sim N(0, \sigma_e^2)$. The model was implemented in the 'rrBLUP' R package (Endelman, 2011) for one trait at a time.

The rrBLUP model assumes common variance across the marker effects, which causes an underestimation of the large-effect QTLs. However, the Bayesian method assumes unequal variances across marker effects and uses different priors to estimate these variances to overcome the limitations of rrBLUP (Meuwissen et al., 2001). We used two Bayesian algorithms, BayesA (BA) and BayesB (BB), to estimate GEBVs for given traits. BA assumes that all markers have a non-zero effect by treating the proportion of markers with no effect (π) as zero. The markers are included in the model after shrinking their estimates to a normal distribution. BB is an extension of BA, which employs an inverse chi-square distribution for marker effects and assumes that some markers have no effect, which are excluded from the model. Thus, BB considers the presence of some large effect QTLs controlling the given trait (Meuwissen et al., 2001; Habier et al., 2011). A detailed description of Bayesian models can be found in Pérez and de los Campos (Pérez and De Los Campos, 2014). The Bayesian models were implemented in the 'BGLR' package using a Gibbs sampler with 5,000 burn-in and 15,000 iterations for each run (Pérez and De Los Campos, 2014). Random Forest (RF) is an ensemble method that uses a collection of classification or regression trees to conduct prediction. The idea is that by combining a large number of smaller decision trees, RF can reduce the variance of prediction. The bias of RF models converges to a limiting value in the limits. RF methods have been applied for both genetic association studies and phenotype predictions (Bureau et al., 2005). In this study, we extended the method used in Grinberg et al. (Grinberg et al., 2020) with 1000 iterations and 64 random states. The algorithm was implemented in Python using Sklearn library.

A multivariate model was used to predict DIS and FDK by including days to heading (DTH) and plant height (PH) as secondary traits in the model. Furthermore, we evaluated the performance of multivariate model to predict DON content using different combinations of secondary traits including DIS, FDK, DTH, and PH. A Bayesian Multivariate Gaussian model with an unstructured variance-covariance matrix was used for the multi-trait (MT) model (Lado et al., 2018). The MT model predicts FDK/DIS/DON using the secondary traits as described in the following:

$$y = 1\mu + Zu + \varepsilon$$

where y is the vector with a length of $n \times t$ (n genotypes and t traits); μ is the means vector; Z represents the incidence matrix of order $[(n \times t)p]$, ; $u_{[(n \times t)p]}$ is genotypic predictor for all individuals and traits with $u \sim N(0, \sum \otimes G)$. The matrix G represents the positive semidefinite matrix obtained from markers. The residuals of the MT model are represented by the vector ε, with $\varepsilon \sim N(0, R \otimes I)$. The matrices $\sum$ and R are the variance-covariance matrices for depicting the genetic and residual effects for each individual in all traits, respectively. $\sum$ was estimated as an unstructured matrix and R as a diagonal matrix following Lado et. al., 2018. The variance-covariance matrices were estimated using a Gibbs sampler with 15,000 iterations, where the first 5,000 iterations were used for burn-in. The MT model was implemented in R package 'MTM' (de los Campos and Grüneberg, 2016).

4.3.6 Training population design and cross-validation

We evaluated the breeding lines from 2018, 2019, and 2020 nurseries as training populations (TP) to predict the performance of preliminary breeding lines. Furthermore,

we combined and evaluated the lines from 2019 and 2020 as one training population. Though the data for these two nurseries come from different years, a set of 58 lines that were common between the two nurseries was used to obtain the trait BLUEs. The resulting TPs were referred to as TP18, TP19, TP20, and TP19+20 for further genomic prediction.

We first used each of these TPs for predicting single-year DIS and FDK. Briefly, each TP was randomly divided into five sets of equal size. Four of the five sets (80%) were used as a training set (phenotyped and genotyped) to train the model, and the remaining set (20%) was used as a testing set (genotyped only) for prediction (Supplementary Figure S1). Predictive ability was estimated as Pearson's correlation between the GEBVs and observed phenotypes for the testing set. The predictions were assessed using five different models, as discussed earlier. The cross-validation process was repeated 500 times for rrBLUP and 100 times for Bayesian models, where each iteration included different lines in the training and testing sets. Cross-validation was used to evaluate the ability of four TPs using different models to estimate GEBVs for DIS and FDK.

For the MT model, the lines were randomly split into a training set (80%) and a testing set (20%). To train the model, we used phenotypic data of secondary traits (PH and DTH) from both the training and testing sets, but the phenotypic data of the target trait (DIS or FDK) from the training set only (Supplementary Figure S1).

As mentioned earlier, we used phenotypic data for DON from 2020 season to evaluate the performance of MT models and compare it with standard ST model (rrBLUP). The lines were randomly split into a training set (80%) and a testing set (20%) as in earlier

case (Supplementary Figure S1). In this case, we evaluated the inclusion of different secondary traits (DIS, FDK, DTH, and PF) in different combinations to evaluate the predictive ability of MT model for DON. For instance, we used a MT model which included all four secondary traits to predict DON, then only three traits as secondary traits, and overall evaluating eight such combinations.

4.3.7 Forward prediction of breeding lines for DIS and FDK

Based on cross-validation analysis, we selected TP19, TP20, and TP19+20 for independent prediction of DIS and FDK in the preliminary lines from the breeding population (BP). TP18 was not used for forward predictions as we did not observe a good PA for DIS and FDK in cross-validation. As all the prediction models yielded comparable results for DIS and FDK using cross-validation, we selected rrBLUP over other models for independent predictions based on its easy and less-intensive implementation. The model was trained using genotypic and phenotypic data from TP19, TP20, and TP19+20 in the 'rrBLUP' package to predict the GEBVs of 200 individuals in the breeding population. To assess the predictive ability, we randomly selected 60 lines of the BP, phenotyped these lines for DIS and FDK in the 2020 nursery, and used the observed phenotypic values to compare to the GEBVs from the TPs.

4.4 Results

4.4.1 Phenotypic analysis

A significant genotypic variation in DIS and FDK ($p < 0.001$) was observed in FHB nurseries from three seasons (Supplementary Table 2). The largest variation in DIS was observed in 2019 ranging from 16.0 to 91.2 with a mean DIS of 49.1 (Table 1), while the

smallest variation was observed in 2018, ranging from 21.0 to 59.7 (Table 1, Figure 1). The mean percent FDK were also different between 2019 and 2020, with 80% in 2019 and 58.6% in 2020. Despite high mean FDK values, sufficient phenotypic variation was observed among evaluated lines in both years (Table 1). Overall, 2019 had the highest disease occurrence and 2018 the lowest.

Significant phenotypic correlations (0.45 and 0.46) were observed between DIS and FDK in both 2019 and 2020. DIS and FDK exhibited negative correlations with plant height in all three nurseries (Supplementary Figures S1, S2 and S3). A negative correlation was observed between DIS and days to heading in 2018. However, both DIS and FDK showed weak but positive correlations with days to heading in 2019 and 2020 (Supplementary Figure S2). As DON was also estimated in 2020, we observed significant positive phenotypic correlations for DON with DIS (0.34), FDK (0.33), and DTH (0.33) (Supplementary Figure S4). Broad-sense heritability for both FHB traits were moderate (Table 1) with the highest heritability for DIS (0.77) in 2020, and for FDK (0.75) in 2019.

4.4.2 Relatedness between TP and BP

The principal component analysis (PCA) was conducted using 9,321 SNP markers and 457 lines. The first two principal components (PCs) explained 6.9% (PC1) and 5.0% (PC2) of the genetic variance (Figure 1). The PCA revealed two primary clusters for the 457 lines, including 257 unique lines from the training populations (TP19 and TP20) and 200 lines of the breeding population (BP). The mixed distribution of the lines from both TPs and BP in the two clusters suggested a close relationship between the two

populations, therefore it will be useful to use the TP for forward prediction in a breeding program.

4.4.3 Cross-validation within TPs for DIS and FDK

Various GP models were used to evaluate the prediction accuracy within each TP using a cross-validation approach. We masked the observed phenotype in 20% of the lines in each TP, and treated them as untested new lines. The PA of DIS and FDK was measured as the Pearsons correlations between the predicted and the observed phenotypes of the masked lines. A general comparison of DIS and FDK prediction using different models is presented using boxplots in Figures 2 and 3. TP18 provided a low mean predictive ability of 0.20 for DIS, ranging from 0.15 to 0.23 using different GP models (Table 2). The TP20 provided a moderate mean predictive accuracy of 0.35 for DIS, and the TP19 yielded the highest mean prediction accuracy (0.39) for DIS, ranging from 0.37 to 0.41 using various models. Furthermore, for FDK, the mean predictive ability was 0.35 ranging from 0.32 to 0.37 in TP19, whereas, in TP20 mean predictive ability was slightly higher ranging from 0.34 to 0.40 with a mean of 0.37 (Table 2). Among four different ST models, rrBLUP outperformed other models in all the TPs for predicting DIS and FDK (Figures 2 and 3).

4.4.4 Combining two TPs into a large TP

To improve prediction accuracies for different traits, we combined two individual populations (TP19 and TP20) into one large population (TP19+20). The BLUE values based on the shared set of lines evaluated in both years were used to assess PA for DIS and FDK using a cross-validation procedure. TP18 had poor performance in the cross-

validation, therefore, it was excluded from the combined TP. The resulting population produced a slightly higher average PA of 0.41 for DIS and 0.38 for FDK using different GP models (Table 2; Figures 3 and 4) than for either TP19 or TP20. Besides a slight increase in the PA using larger TP, we observed that PA of all the models was quite similar, indicating a stable PA irrespective of the models.

4.4.5 Univariate v/s multivariate prediction models for DIS and FDK

PA also varied with types of GP models used for predicting DIS and FDK. Among four univariate models used for both DIS and FDK predictions, the rrBLUP model performed slightly better than BayesA (BA), BayesB (BB), and Random Forest (RF) models in all TP scenarios (Table 2). When the univariate GP models were compared with a multi-trait (MT) model with plant height and days to heading as covariates, the MT model did not show much improvement in DIS predictions than univariate models using TP18, TP19, and TP19+20. However, the prediction accuracy of MT was improved by 20% (0.35 to 0.42) for DIS in TP20 when two secondary traits (DTH and PH) were included in the model (Table 2). Improvement using the MT model in TP20 can be attributed to the moderate correlation between DIS and secondary traits in the model (Supplementary Figure S4). For FDK, the MT model did not improve prediction accuracy in any of the TPs (Table 2) as FDK seemed to be less correlated with DTH or PH in any of the growing seasons.

4.4.6 Evaluating MT model to predict DON using different combinations of secondary traits

The lines from 2020 nursery were also evaluated for DON along with other FHB traits. As DON is a costly trait to phenotype and a smaller number of samples are analyzed for DON each year, we were interested to see if we can use other FHB traits in MT models to predict DON. For this, we used the MT model with different combinations of DIS, FDK, DTH, and PH as covariates to predict DON. We used ST rrBLUP model as a benchmark to compare the performance of MT model (Figure 4). Using TP20, the rrBLUP model yielded a predictive accuracy of 0.49 for DON, which is higher as compared to DIS and FDK using the same model. The MT model showed an improvement of up to 20%, yielding PA ranging from 0.54 to 0.59 with different combinations of secondary traits (Supplementary Table S3). The MT model having DIS, FDK, DTH, and PH as covariates had PA of 0.56, whereas the MT model with DIS, FDK, and DTH had the highest PA (0.59) among all the combinations (Supplementary Table S3). We also evaluated the MT model if only single trait (DIS, FDK, or DTH) is used as covariate to predict DON. Interestingly, the MT model with only FDK as covariate had PA of 0.56, which is comparable to the combination having all traits as covariates (Supplementary Table S3).

4.4.7 Accuracy of independent predictions for DIS and FDK in forward breeding

To validate usefulness of different FHB nurseries as possible TPs, GEBVs for DIS and FDK of a random set of preliminary breeding lines from PYT were predicted using TP19, TP20, and TP19+20 with the rrBLUP model. Moderate prediction accuracies (Table 3) were observed for DIS and FDK using the three TPs. The TP19 provided the highest prediction accuracy (0.59) for DIS, following the trend observed in our cross-validation.

The TP19+20 produced the highest prediction accuracy of 0.54 for FDK, followed by TP19 and TP20 (0.50 and 0.49, Table 3). Overall, independent predictions provided better PA than the cross-validation. Furthermore, we used a scatterplot to compare the breeding lines which were rejected based on estimated GEBVs but retained based on observed data (Figure 5; Supplementary Figure S5). Interestingly, we observed there was low probability of rejecting lines with lower DIS or FDK ratings as most of lines rejected based on GEBVs were having an moderate DIS or FDK observed value in both 2019 and 2020 (refer to top-left quadrant of scatterplots in Figure 5 and Supplementary Figure S5). These results demonstrate that the genomic prediction can be implemented to improve FHB resistance in wheat breeding programs.

4.5 Discussion

Several studies have evaluated the inclusion of GS in wheat breeding programs to predict FHB resistance in recent years (Rutkoski et al., 2012; Arruda et al., 2015; Dong et al., 2018; Verges et al., 2020); however, most of these studies were done in soft winter wheat. The current study used a cross-validation strategy to evaluate the potential of hard winter wheat breeding lines as training populations for GS of FHB resistance. The multi-trait GP model was also evaluated for predicting different FHB traits. Finally, the current study demonstrates the use of different sets of advanced breeding lines as training sets to predict preliminary breeding lines in a forward prediction scheme.

4.5.1 Phenotypic response to FHB

We used an FHB disease index (DIS) estimated based on incidence and severity, and FDK percentage to evaluate FHB resistance in advanced breeding lines. We observed a

wide variation for FHB resistance in the advanced breeding lines from the South Dakota State University winter wheat breeding program in three disease nurseries from 2018-2020. For example, variations for DIS were from 16.0 to 91.2% and FDK from 38.1 to 99.0% in the 2019 nursery (Table 1). A similar trend was observed in two other nurseries (Table 1). The majority of the advanced and elite breeding material from our program does not carry *Fhb1* likely due to yield drag and negative agronomic potential, suggesting minor genes from native sources predominantly govern FHB resistance in our breeding program. This is similar to other hard winter wheat breeding programs in the region as only one variety (TAM 205) carrying *Fhb1* has been released till date. Therefore, GS seems to be a suitable approach to breed for the FHB resistance.

Although the FHB nursery had a controlled mist system, the prevailing environment in respective years is believed to play a significant role in varying disease pressure. The 2018 season was very dry in South Dakota, but it was reasonably wet in 2019. These environmental factors lead to fluctuations in the disease pressure affecting the spread of the data over the years, consistent with previous studies (Hoffstetter et al., 2016a; Adeyemo et al., 2020; Larkin et al., 2020; Verges et al., 2020). However, these fluctuations did not alter the ranking of different check genotypes over different nurseries. For instance, the DIS indices were 27.2, 29.8, and 32.6 in 2018, 2019, and 2020, respectively, for resistant check ‘Lyman’, and 84.6 and 72.4 in 2019 and 2020, respectively, for susceptible check ‘Overley’. The consistent performance of checks shows uniformity and reliability of phenotyping across nurseries, which is supported by higher heritability for DIS (0.54 – 0.77) and FDK (0.66 – 0.75) (Table 1). The moderate H^2 estimates were in similar range as that of related studies using different types of

population in spring- or winter wheat (Hoffstetter et al., 2016b; Adeyemo et al., 2020; Larkin et al., 2020; Xu et al., 2020; Zhu et al., 2020).

4.5.2 Within-population cross-validation for DIS and FDK

Several studies successfully used advanced or preliminary breeding lines from a breeding program (Hoffstetter et al., 2016a; Belamkar et al., 2018; Adeyemo et al., 2020) and from unrelated regional nurseries (Verges et al., 2020) as training populations to predict various traits in wheat. We evaluated advanced breeding lines in three FHB nurseries as three possible training sets for genomic prediction using cross-validation (Supplementary Table 1), and obtained moderate prediction accuracies when TP19 and TP20 were used to predict DIS (up to 0.41 using TP19 and 0.42 in TP20) and FDK (up to 0.37 in TP19 and 0.40 in TP20) (Table 2), consistent with several previous studies in soft winter or spring wheat (Hoffstetter et al., 2016a, Dong et al., 2018, Verges et al., 2020, Adeyemo et al., 2020). Similarly, prediction accuracies for FDK in our study were also similar to Rutkoski et al., (2012) and Adeyemo et al., (2020), but lower than those reported by Arruda et al. (2015) and Verges et al. (2020). The poor performance of TP18 for DIS prediction could result from low disease pressure and phenotypic variation observed in the 2018 nursery, hence, such a nursery is not recommended for a forward genomic prediction. Overall, based on mutliple years of data our study demonstrates the usefulness of advanced breeding lines as TPs for genomic prediction, given that the quality of phenotyping is robust.

Training population size is another crucial factor that affects the PA of GP models. Previous studies have reported an increase in PA with an increased TP size (Lorenz et al., 2012; Michel et al., 2017; Sarinelli et al., 2019). Lorenz et al. (2012) and Arruda et al. (2015) obtained higher PA for FHB traits in barley and wheat when the TP size contained 250 to 300 lines. In the current study, the three populations, each with around 150 lines, were used as independent TPs for FHB trait prediction. Also, the performance of a larger TP (TP19+20) developed by combining two TPs (TP19 and TP20) were evaluated. Although TP19 and TP20 were phenotyped in two independent nurseries, they shared a subset of 58 breeding lines. Thus the large TP (TP19+20) was formed with 265 unique lines. The higher PA were observed for DIS (0.41) and FDK (0.38) when TP19+20 was used as the TP. Apart from improvement in PA using TP19+20, we observed that all prediction models yielded similar results which was not the case when using smaller TPs (Table 2). Furthermore, a lower standard deviation for PA based on several repeats of cross-validation suggests consistency of prediction when using a large TP (Figures 3 and 4). The results of this study suggest that selection of the right TP is crucial to improve PA. TP19+20 showed the best PA due to the large TP size and correction of BLUEs for DIS and FDK across two different nurseries/years, thus it can be used as the TP in forward genomic prediction of preliminary breeding lines.

4.5.3 Model comparison for DIS and FDK

We compared the PA for DIS and FHB using four univariate and one multivariate GP model. Among the univariate models, rrBLUP outstripped the other three models for predicting DIS and FDK in all individual TPs (Table 2). Further, rrBLUP is preferred model for predicting FHB traits owing to its better performance and computational

advantage. Previous studies have also reported that rrBLUP has better PA than Bayesian models (Rutkoski et al., 2012; Arruda et al., 2015) is one of the most often used methods in GS for FHB resistance.

Using a multi-trait (MT) GP model is another approach to increase the PA for FHB traits. This model includes correlated secondary traits such as PH and DTH as covariates to predict DIS and FDK. Schulthess et al. (Schulthess et al., 2018) and Larkin et al. (Larkin et al., 2020) reported an increase in PA by including PH or DTH in the GP model. In the current study, PH and DTH were included as secondary traits in a MT-GBLUP model to predict DIS and FDK, however, improvement of PA were not considerable except for DIS in TP20 where the PA was improved by at least 20% over the other univariate models (Table 2). The better PA using MT model in the TP20 resulted from the moderate correlation of DIS with PH and DTH (Supplementary Figure S1). In other TPs, a lower correlation was observed among evaluated traits. The MT models have been used to improve the PA of low-heritability traits by using the information from correlated traits with high heritability (Calus and Veerkamp, 2011; Jia and Jannink, 2012; Gill et al., 2021). However, moderate to high heritability estimates for DIS and FDK in this study could be another reason for MT models showing no advantages over univariate models. Thus, our results suggest that the MT models could be useful and employed in forward prediction in early generation nurseries if the observed correlations between FHB traits and the covariates are moderate to high.

4.5.4 Optimization of MT model with different combinations of secondary traits to predict DON

DON is an important and primary FHB trait; however, most of the winter wheat breeders across the US are unable to make decisions based on DON estimates. For instance, the SDSU winter wheat program harvests FHB nursery in August and then plants the next cycle in September, making the turnaround cycle very short. Contrary to this, phenotyping for DON is mostly outsourced and it takes several months before the breeder gets the data. Thus, it will be of great importance if we can predict DON and utilize the predictions to forward resistant lines. Furthermore, previous studies have suggested the use of different pre-harvest traits, such as DIS, FDK, DTH or PH, in MT models to predict DON with better accuracy (Larkin et al., 2020). We used DON estimates available for TP20 to evaluate the predictive ability for DON using ST and different versions of MT model based on cross-validation. The ST (rrBLUP) model predicted DON with PA of 0.49, which is better than DIS and FDK using any of the models (Supplementary Table S3). Further, we compared MT model having different combinations of traits as covariates. We observed that using trait combinations of DIS, FDK, DTH (0.59) and FDK, DTH (0.58) had the highest PA (Fig. 4), similar to the results reported by Larkin et al. (Larkin et al., 2020). Interestingly, it was found that using FDK as ‘only’ covariate in the MT model had high PA (0.56), suggesting that using FDK can improve the performance of MT model over the ST model. Overall, the results suggest that we can use GP for predicting DON in earlier stages and the PA can be further improved by using secondary traits such as FDK.

4.5.5 Genomic predictions in forward breeding

Relatedness among the individuals in TP and BP is considered crucial for getting higher PA in genomic prediction. In this study, we aimed to evaluate the usefulness of advanced breeding lines as TPs to predict earlier generations lines (PYT) from our breeding program. The PCA showed a good association between the lines from TP and BP (Figure 2). Hence, we obtained moderate prediction accuracy for DIS and FDK in an independent BP (Table 3), comparable to the other reports (Jiang et al., 2017; Herter et al., 2019; Verges et al., 2020). Though a moderate PA for DIS and FDK were achieved in the current study, the phenotypic data of advanced lines evaluated from FHB nurseries could be used to predict GEBVs for the early generation breeding lines in wheat breeding programs (Table 3). The predicted GEBVs can be used to discard susceptible lines at the earlier stages of the breeding cycle (Figure 5, Supplementary Figure 3). For instance, we achieved a PA of 0.59 for DIS using TP19 and discarded 50% of the most susceptible lines based on the TP19 based GEBVs and selected the remaining 50% lines for further selection (Table 3). Among these discarded lines, 87% were highly susceptible based on their observed disease index in the mist irrigated inoculated FHB nursery, and only 13% susceptible genotypes were carried forward based on GEBVs (Figure 5). Similar results were observed for FDK (Supplementary Figure 3), which suggests that the estimated GEBVs from the TP can be used to discard the most susceptible lines at an early stage to reduce phenotyping costs. Our results also suggest that these lines discarded based on GEBVs had an extremely low chance to be highly resistant in terms of DIS and FDK.

In summary, our study demonstrated that advanced breeding lines evaluated in the FHB nurseries can serve as TPs and predict GEBVs for untested earlier generation lines in the breeding programs. Further advanced lines from several years can be combined to increase the PA in the forward breeding. However, we recommend evaluating the performance of individual advanced breeding nurseries through cross-validation before pooling multiple nurseries to develop larger TPs for forward prediction. Furthermore, the MT models using secondary traits as covariates could be useful in predicting cumbersome FHB traits (DON and FDK). Finally, our results suggest genomic prediction can be successfully applied in a wheat breeding program in discarding the most FHB susceptible lines at an early stage as opposed to laborious phenotypic selection in most years and especially in abnormal years when phenotypic evaluation is unreliable or unavailable due to environmental conditions.

Author Contributions

SKS, JZ and HSG conceptualized the experiment and designed the methodology; JZ, NK, SA, JH and HSG performed the investigation; HSG and SKS performed the data curation; JZ, HSG and SKS performed the data analysis and visualization; HSG, JZ performed the software implementation; AB, PSA, and GB carried out genotyping and SNP discovery; JZ, HSG and SKS wrote the original manuscript; SA, GB, UG, and BT contributed to the interpretation of results. All the authors revised the manuscript.

Funding

This project was collectively funded by the USDA hatch projects SD00H695-20, USDA-ARS agreement 59-0206-0-177 (USDA-USWBSI), the Agriculture and Food Research

Initiative Competitive Grants 2017-67007-25939 (Wheat-CAP) and South Dakota Wheat Commission grant 3X1340. The funders had no role in the study design, data collection and analysis, decision to publish, or manuscript preparation.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The authors would like to thank the South Dakota Agriculture Experimental Station (Brookings, SD, USA) for providing the resources to conduct the experiments.

4.6 References

Adeyemo, E., Bajgain, P., Conley, E., Sallam, A., and Anderson, J. (2020). Optimizing Training Population Size and Content to Improve Prediction Accuracy of FHB-Related Traits in Wheat. *Agronomy* 10, 543. doi:10.3390/agronomy10040543.

Ali, M., Zhang, Y., Rasheed, A., Wang, J., and Zhang, L. (2020). Genomic prediction for grain yield and yield-related traits in chinese winter wheat. *Int. J. Mol. Sci.* 21. doi:10.3390/ijms21041342.

Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., et al. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *Crop J.* 8, 745–756. doi:10.1016/j.cj.2020.03.010.

Arruda, M. P., Brown, P. J., Lipka, A. E., Krill, A. M., Thurber, C., and Kolb, F. L. (2015). Genomic Selection for Predicting *Fusarium* Head Blight Resistance in a Wheat Breeding Program. *Plant Genome* 8. doi:10.3835/plantgenome2015.01.0003.

Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to fusarium head blight. *Annu. Rev. Phytopathol.* 42, 135–161. doi:10.1146/annurev.phyto.42.040803.140340.

Bai, G., Su, Z., and Cai, J. (2018). Wheat resistance to Fusarium head blight. *Can. J. Plant Pathol.* 40, 336–346. doi:10.1080/07060661.2018.1476411.

Bassi, F. M., Bentley, A. R., Charmet, G., Ortiz, R., and Crossa, J. (2015). Breeding schemes for the implementation of genomic selection in wheat (Triticum spp.). *Plant Sci.* 242, 23–36. doi:10.1016/j.plantsci.2015.08.021.

Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67. doi:10.18637/jss.v067.i01.

Battenfield, S. D., Guzmán, C., Gaynor, R. C., Singh, R. P., Peña, R. J., Dreisigacker, S., et al. (2016). Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program. *Plant Genome* 9. doi:10.3835/plantgenome2016.01.0005.

Belamkar, V., Guttieri, M. J., Hussain, W., Jarquín, D., El-basyoni, I., Poland, J., et al. (2018). Genomic selection in preliminary yield trials in a winter wheat breeding program. *G3 Genes, Genomes, Genet.* 8, 2735–2747. doi:10.1534/g3.118.200415.

Bhat, J. A., Ali, S., Salgotra, R. K., Mir, Z. A., Dutta, S., Jadon, V., et al. (2016). Genomic selection in the era of next generation sequencing for complex traits in plant breeding. *Front. Genet.* 7. doi:10.3389/fgene.2016.00221.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi:10.1093/bioinformatics/btm308.

Browning, S. R., and Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097. doi:10.1086/521987.

Bureau, A., Dupuis, J., Falls, K., Lunetta, K. L., Hayward, B., Keith, T. P., et al. (2005). Identifying SNPs predictive of phenotype using random forests. *Genet. Epidemiol.* 28, 171–182. doi:10.1002/gepi.20041.

Calus, M. P. L., and Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genet. Sel. Evol.* 43, 26. doi:10.1186/1297-9686-43-26.

de los Campos, G., and Grüneberg, A. (2016). MTM Package. Available at:

https://github.com/QuantGen/MTM/.

Dong, H., Wang, R., Yuan, Y., Anderson, J., Pumphrey, M., Zhang, Z., et al. (2018). Evaluation of the Potential for Genomic Selection to Improve Spring Wheat Resistance to Fusarium Head Blight in the Pacific Northwest. *Front. Plant Sci.* 9, 911. doi:10.3389/fpls.2018.00911.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* Available at: https://worldveg.tind.io/record/33886 [Accessed January 20, 2021].

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome* 4, 250–255. doi:10.3835/plantgenome2011.08.0024.

Gaire, R., Brown-Guedira, G., Dong, Y., Ohm, H., and Mohammadi, M. (2021). Genome-Wide Association Studies for Fusarium Head Blight Resistance and Its Trade-Off With Grain Yield in Soft Red Winter Wheat. *Plant Dis.* 105, 2435–2444. doi:10.1094/PDIS-06-20-1361-RE.

Gilbert, J., and Haber, S. (2013). Overview of some recent research developments in fusarium head blight of wheat. *Can. J. Plant Pathol.* 35, 149–174. doi:10.1080/07060661.2013.772921.

Gill, H. S., Halder, J., Zhang, J., Brar, N. K., Rai, T. S., Hall, C., et al. (2021). Multi-Trait Multi-Environment Genomic Prediction of Agronomic Traits in Advanced Breeding Lines of Winter Wheat. *Front. Plant Sci.* 12. doi:10.3389/fpls.2021.709545.

Grinberg, N. F., Orhobor, O. I., and King, R. D. (2020). An evaluation of machine-learning for predicting phenotype: studies in yeast, rice, and wheat. *Mach. Learn.*

109, 251–277. doi:10.1007/s10994-019-05848-5.

Habier, D., Fernando, R. L., Kizilkaya, K., and Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics* 12. doi:10.1186/1471-2105-12-186.

Halder, J., Zhang, J., Ali, S., Sidhu, J. S., Gill, H. S., Talukder, S. K., et al. (2019). Mining and genomic characterization of resistance to tan spot, Stagonospora nodorum blotch (SNB), and Fusarium head blight in Watkins core collection of wheat landraces. *BMC Plant Biol.* 19, 1–15. doi:10.1186/s12870-019-2093-3.

Heffner, E. L., Sorrells, M. E., and Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12. doi:10.2135/cropsci2008.08.0512.

Herter, C. P., Ebmeyer, E., Kollers, S., Korzun, V., Würschum, T., and Miedaner, T. (2019). Accuracy of within- and among-family genomic prediction for Fusarium head blight and Septoria tritici blotch in winter wheat. *Theor. Appl. Genet.* 132, 1121–1135. doi:10.1007/s00122-018-3264-6.

Hoffstetter, A., Cabrera, A., Huang, M., and Sneller, C. (2016a). Optimizing training population data and validation of genomic selection for economic traits in soft winter wheat. *G3 Genes, Genomes, Genet.* 6, 2919–2928. doi:10.1534/g3.116.032532.

Hoffstetter, A., Cabrera, A., and Sneller, C. (2016b). Identifying Quantitative Trait Loci for Economic Traits in an Elite Soft Red Winter Wheat Population. *Crop Sci.* 56. doi:10.2135/cropsci2015.06.0332.

Jia, Y., and Jannink, J. L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* 192, 1513–1522.

doi:10.1534/genetics.112.144246.

Jiang, Y., Albert, ·, Schulthess, W., Rodemann, B., Ling, J., Plieske, J., et al. (2017). Validating the prediction accuracies of marker-assisted and genomic selection of Fusarium head blight resistance in wheat using an independent sample. *Theor. Appl. Genet.* 130, 471–482. doi:10.1007/s00122-016-2827-7.

Juliana, P., Singh, R. P., Singh, P. K., Crossa, J., Huerta-Espino, J., Lan, C., et al. (2017). Genomic and pedigree-based prediction for leaf, stem, and stripe rust resistance in wheat. *Theor. Appl. Genet.* 130, 1415–1430. doi:10.1007/s00122-017-2897-1.

Lado, B., Vázquez, D., Quincke, M., Silva, P., Aguilar, I., and Gutiérrez, L. (2018). Resource allocation optimization with multi-trait genomic prediction for bread wheat (Triticum aestivum L.) baking quality. *Theor. Appl. Genet.* 131, 2719–2731. doi:10.1007/s00122-018-3186-3.

Larkin, D. L., Holder, A. L., Mason, R. E., Moon, D. E., Brown-Guedira, G., Price, P. P., et al. (2020). Genome-wide analysis and prediction of Fusarium head blight resistance in soft red winter wheat. *Crop Sci.* 60, 2882–2900. doi:10.1002/csc2.20273.

Liu, S., Hall, M. D., Griffey, C. A., and McKendry, A. L. (2009). Meta-Analysis of QTL associated with fusarium head blight resistance in wheat. *Crop Sci.* 49, 1955–1968. doi:10.2135/cropsci2009.03.0115.

Lorenz, A. J., Chao, S., Asoro, F. G., Heffner, E. L., Hayashi, T., Iwata, H., et al. (2011). *Genomic Selection in Plant Breeding. Knowledge and Prospects.* Academic Press Inc. doi:10.1016/B978-0-12-385531-2.00002-5.

Lorenz, A. J., Smith, K. P., and Jannink, J. L. (2012). Potential and optimization of

genomic selection for Fusarium head blight resistance in six-row barley. *Crop Sci.* 52, 1609–1621. doi:10.2135/cropsci2011.09.0503.

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., et al. (2012). A unified effort to fight an enemy of wheat and barley: Fusarium head blight. *Plant Dis.* 96, 1712–1728. doi:10.1094/PDIS-03-12-0291-FE.

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. Available at: https://www.genetics.org/content/157/4/1819.short [Accessed January 12, 2021].

Michel, S., Ametz, C., Gungor, H., Akgöl, B., Epure, D., Grausgruber, H., et al. (2017). Genomic assisted selection for enhancing line breeding: merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theor. Appl. Genet.* 130, 363–376. doi:10.1007/s00122-016-2818-8.

Mirdita, V., He, S., Zhao, Y., Korzun, V., Bothe, R., Ebmeyer, E., et al. (2015). Potential and limits of whole genome prediction of resistance to Fusarium head blight and Septoria tritici blotch in a vast Central European elite winter wheat population. *Theor. Appl. Genet.* 128, 2471–2481. doi:10.1007/s00122-015-2602-1.

Moreno-Amores, J., Michel, S., Miedaner, T., Longin, C. F. H., and Buerstmayr, H. (2020). Genomic predictions for Fusarium head blight resistance in a diverse durum wheat panel: an effective incorporation of plant height and heading date as covariates. *Euphytica* 216, 1–19. doi:10.1007/s10681-019-2551-x.

Pérez, P., and De Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* 198, 483–495. doi:10.1534/genetics.114.164442.

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome* 5, plantgenome2012.06.0006. doi:10.3835/plantgenome2012.06.0006.

R Core Team (2018). R: A language and environment for statistical computing; 2015. Available at: https://scholar.google.com/scholar?cluster=9441913529578809097&hl=en&as_sdt=5,42&sciodt=0,42 [Accessed September 17, 2019].

Robert W. Stack, and Marcia P. McMullen (2011). A Visual Scale to Estimate Severity of Fusarium Head Blight in Wheat. Fargo Available at: www.ag.ndsu.edu/agcomm/creative-commons. [Accessed April 23, 2021].

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J.-L., and Sorrells, M. (2012). Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *Plant Genome* 5, 51–61. doi:10.3835/plantgenome2012.02.0001.

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., et al. (2019). Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel. *Theor. Appl. Genet.* 132, 1247–1261. doi:10.1007/s00122-019-03276-6.

Schulthess, A. W., Zhao, Y., Longin, C. F. H., and Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 131, 685–701. doi:10.1007/s00122-017-3029-7.

Singh, R. P., Singh, P. K., Rutkoski, J., Hodson, D. P., He, X., Jørgensen, L. N., et al.

(2016). Disease Impact on Wheat Yield Potential and Prospects of Genetic Control. *Annu. Rev. Phytopathol.* 54, 303–322. doi:10.1146/annurev-phyto-080615-095835.

Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., and Kizek, R. (2010). Deoxynivalenol and its toxicity. *Interdiscip. Toxicol.* 3, 94–99. doi:10.2478/v10102-010-0019-x.

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., and Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Trop. Plant Pathol.* 42, 165–174. doi:10.1007/s40858-017-0127-7.

Su, Z., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H., et al. (2019). A deletion mutation in TaHRC confers Fhb1 resistance to Fusarium head blight in wheat. *Nat. Genet.* 51, 1099–1105. doi:10.1038/s41588-019-0425-8.

Trail, F. (2009). For Blighted Waves of Grain: *Fusarium graminearum* in the Postgenomics Era: Figure 1. *Plant Physiol.* 149. doi:10.1104/pp.108.129684.

Venske, E., dos Santos, R. S., Farias, D. da R., Rother, V., da Maia, L. C., Pegoraro, C., et al. (2019). Meta-analysis of the QTLome of fusarium head blight resistance in bread wheat: Refining the current puzzle. *Front. Plant Sci.* 10, 727. doi:10.3389/fpls.2019.00727.

Verges, V. L., Lyerly, J., Dong, Y., and Van Sanford, D. A. (2020). Training Population Design With the Use of Regional Fusarium Head Blight Nurseries to Predict Independent Breeding Lines for FHB Traits. *Front. Plant Sci.* 11, 1083. doi:10.3389/fpls.2020.01083.

Wilson, W. W., McKee, G., Nganje, W., Dahl, B., and Bangsund, D. (2017). Economic

Impact of USWBSI's Scab Initiative to Reduce FHB. Fargo doi:10.22004/AG.ECON.264672.

Xu, Q., Xu, F., Qin, D., Li, M., Fedak, G., Cao, W., et al. (2020). Molecular Mapping of QTLs Conferring Fusarium Head Blight Resistance in Chinese Wheat Cultivar Jingzhou 66. *Plants* 9. doi:10.3390/plants9081021.

Zhang, X., Halder, J., White, R. P., Hughes, D. J., Ye, Z., Wang, C., et al. (2014). Climate change increases risk of fusarium ear blight on wheat in central China. *Ann. Appl. Biol.* 164, 384–395. doi:10.1111/aab.12107.

Zhu, Z., Chen, L., Zhang, W., Yang, L., Zhu, W., Li, J., et al. (2020). Genome-Wide Association Analysis of Fusarium Head Blight Resistance in Chinese Elite Wheat Lines. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.00206.

4.7 Appendix: Figures and tables

Figures

Figure 4.1 The scatterplot for principal component analysis (PCA) of 457 lines based on 9,321 SNP markers. The 457 breeding lines include 257 lines from training populations (TP19 and TP20) and 200 lines of the breeding population (BP) used in the forward prediction. The blue triangles represent the lines from TPs, and the green circles represents lines from BP. The first two PCs explained 6.9% and 5.0% of the total variation, respectively.

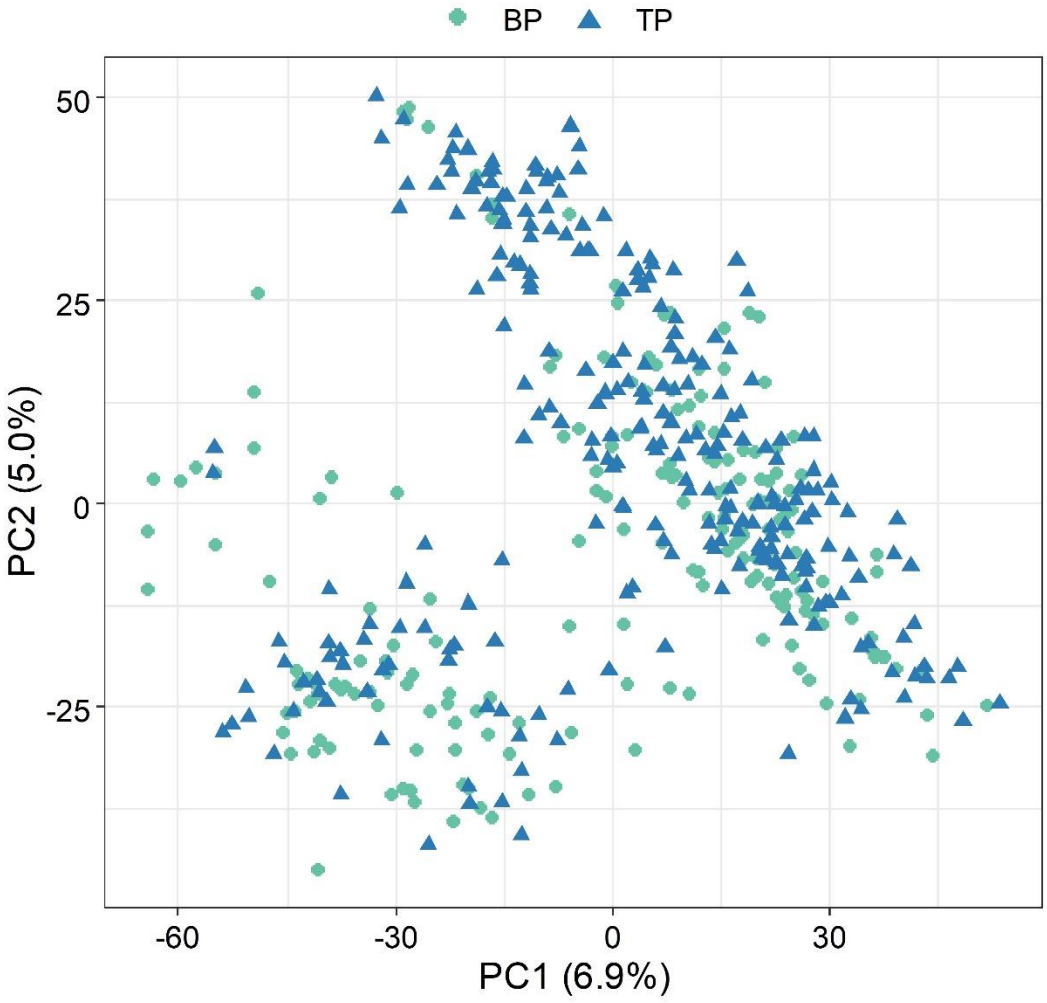


Figure 4.2 The predictive ability (PA) for Fusarium head blight (FHB) resistance disease index (DIS) in different sets of training populations (TPs) used in the study. Boxplots compare the PA using five genomic prediction models: rrBLUP, ridge-regression best linear unbiased prediction; BayesA, BA; BayesB, BB; Random Forest, RF; and Multi-trait model, MT. Training population based on 2018 FHB nursery, TP18; Training population based on 2019 FHB nursery, TP19; Training population based on 2020 FHB nursery, TP20; Training population combining 2019 and 2020 FHB nurseries, TP19+20.

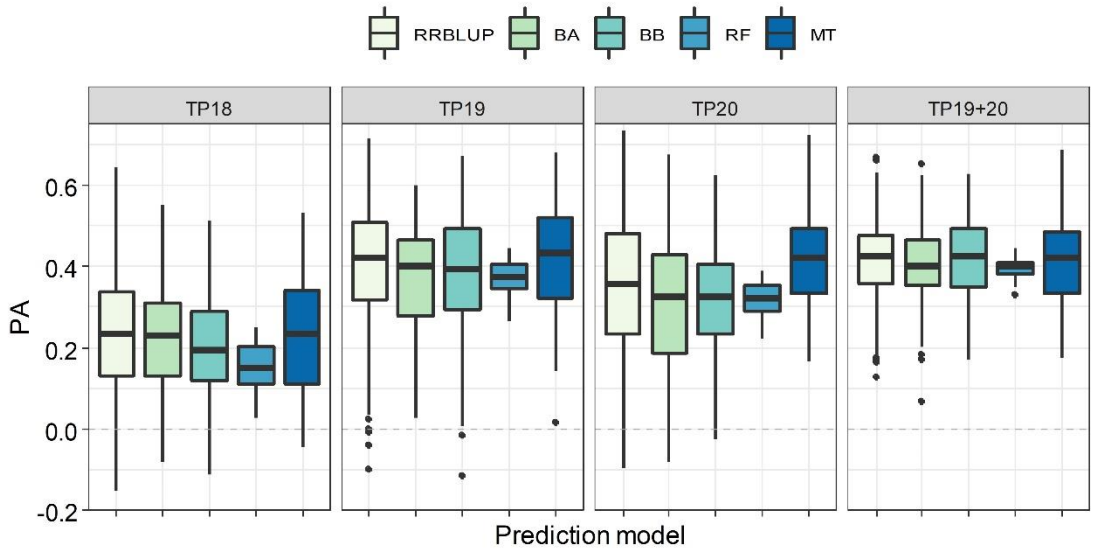


Figure 4.3 The predictive ability (PA) for Fusarium-damaged kernels (FDK) in different sets of training populations (TPs) used in the study. Boxplots compare the PA using five genomic prediction models: rrBLUP, ridge-regression best linear unbiased prediction; BayesA, BA; BayesB, BB; Random Forest, RF; and Multi-trait model, MT. Training population based on 2019 FHB nursery, TP19; Training population based on 2020 FHB nursery, TP20; Training population combining 2019 and 2020 FHB nurseries, TP19+20.

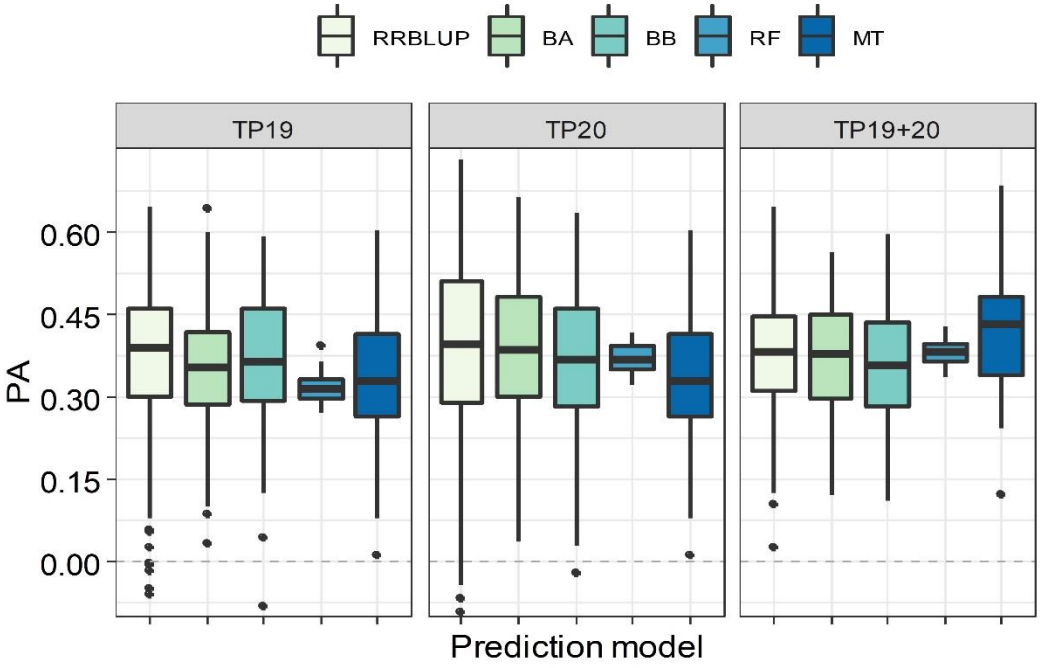


Figure 4.4 Boxplots comparing the predictive ability of ST model and the MT model with different combinations of secondary traits. DIS, FHB disease Index; FDK, fusarium damaged kernel percentage; DTH, days to heading; PH, plant height; DON, deoxynivalenol content.

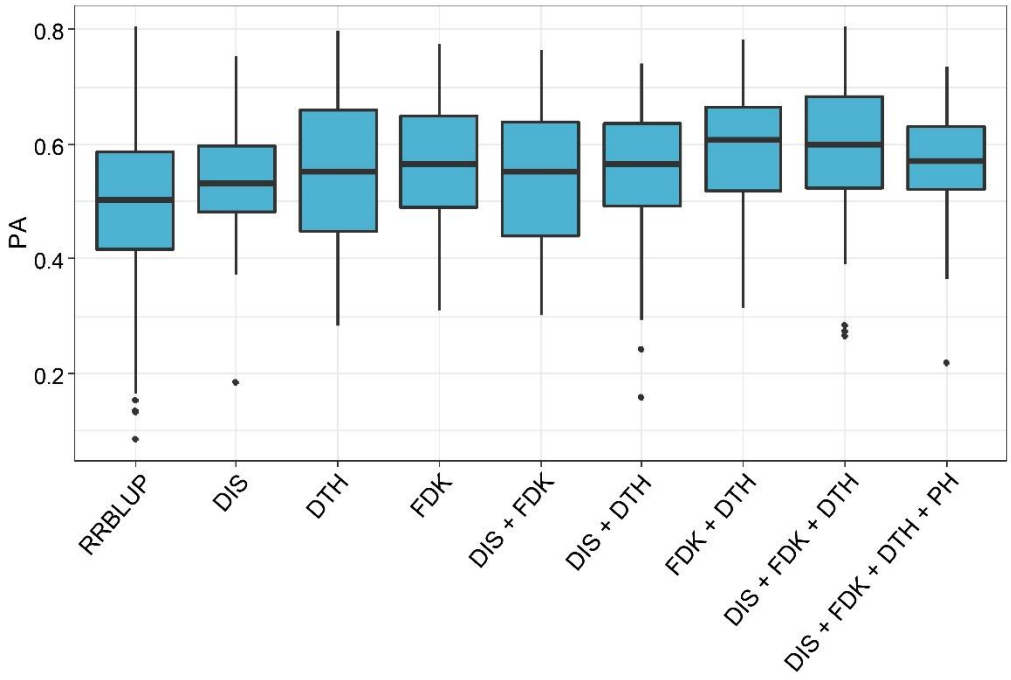
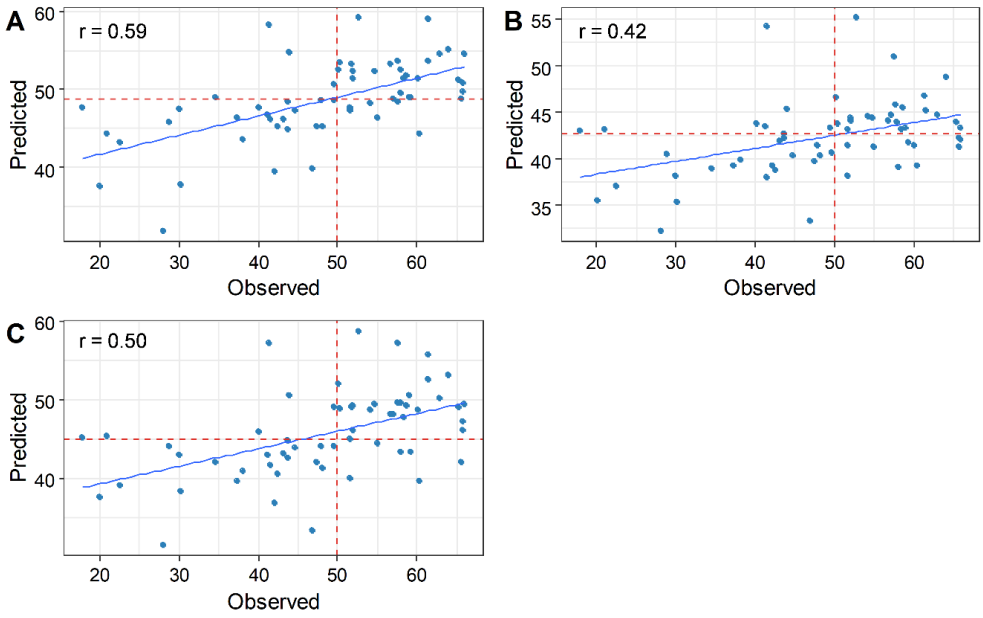


Figure 4.5 Scatterplots showing observed v/s predicted values for FHB disease index (DIS). The observed DIS estimates were based on phenotypic evaluation of 60 independent lines, and the predicted values are GEBVs using **(A)** TP19, **(B)** TP20, and **(C)** TP19+20. The red dashed line represents the cutoff (50%) to discard genotypes based on the observed data and estimated GEBVs. Genotype to the right side of the red dashed line would be discarded based on observed phenotype, and genotype above the red dashed line would be discarded based on GEBVs.



Tables

Table 4.1. Descriptive statistics of two Fusarium head blight (FHB) resistance traits, disease index (DIS), and Fusarium-damaged kernels (FDK) for advanced winter wheat breeding lines evaluated in three independent FHB nurseries from 2018, 2019, and 2020.

Trait	Year of valuation	N[a]	Mean	Min.	Max.	H^2
DIS (%)	2018	153	42.9	21.0	59.7	0.54
(0-100)	2019	169	49.1	16.0	91.2	0.76
	2020	154	42.0	12.0	81.5	0.77
FDK	2019	169	80.0	38.1	99.0	0.75
(%)	2020	153	58.6	22.2	89.5	0.66

a The number of advanced breeding lines evaluated in each nursery is represented by *N*; Min. and Max. refers to the Minimum and Maximum trait values; H^2 is the broad-sense heritability for the respective trait.

Table 4.2. Mean prediction accuracy and standard error for two Fusarium head blight (FHB) resistance traits, disease index (DIS) and Fusarium-damaged kernels (FDK) with cross-validation in different Training Populations using different genomic prediction models. rrBLUP, ridge-regression best linear unbiased prediction; BayesA, BA; BayesB, BB; Random Forest, RF; and Multi-trait model, MT.

		Genomic prediction model				
Trait	TP[a]	rrBLUP	BA	BB	RF	MT
DIS	TP18	0.23 ± 0.01	0.22 ± 0.01	0.20 ± 0.01	0.15 ± 0.01	0.22 ± 0.02
	TP19	0.41 ± 0.01	0.38 ± 0.01	0.37 ± 0.02	0.38 ± 0.01	0.41 ± 0.02
	TP20	0.35 ± 0.01	0.32 ± 0.02	0.32 ± 0.01	0.32 ± 0.01	0.42 ± 0.02
	TP19+20	0.42 ± 0.01	0.41 ± 0.01	0.41 ± 0.01	0.40 ± 0.01	0.41 ± 0.01
FDK	TP19	0.37 ± 0.01	0.35 ± 0.01	0.36 ± 0.01	0.32 ± 0.01	0.34 ± 0.02
	TP20	0.40 ± 0.01	0.37 ± 0.01	0.37 ± 0.01	0.37 ± 0.01	0.34 ± 0.02
	TP19+20	0.38 ± 0.01	0.37 ± 0.01	0.35 ± 0.01	0.38 ± 0.01	0.41 ± 0.02

[a]Different training populations used in cross-validation analysis. Training population based on 2018 FHB nursery, TP18; Training population based on 2019 FHB nursery, TP19; Training population based on 2020 FHB nursery, TP20; Training population combining 2019 and 2020 FHB nurseries, TP19+20.

Table 4.3. Prediction accuracy for two Fusarium head blight (FHB) resistance traits, disease index (DIS) and Fusarium-damaged kernels (FDK) in forward prediction scheme. Three different training populations were used to predict 200 untested preliminary yield trial lines using rrBLUP model. The prediction accuracy was estimated by phenotyping a random set of 60 untested lines for given traits.

Trait	Training population		
	TP19	TP20	TP19+20
DIS	0.59	0.42	0.50
FDK	0.50	0.49	0.54

Supplementary Table S1. Information of the three SDSU FHB nurseries evaluated in the growing seasons of 2017-18, 2018-19, and 2019-20.

Training population	Year of evaluation	Location of FHB nursery	Date of Planting	Date of Harvesting	Lines evaluated	Lines used for GS.
TP18	2018	Brookings, SD	10/15/2017	08/18/2018	153	152
TP19	2019	Brookings, SD	10/08/2018	08/13/2019	169	161
TP20	2020	Brookings, SD	10/03/2019	08/17/2020	154	153

Supplementary Table S2. Analysis of variance for FHB traits from the linear mixed model analysis for individual FHB nurseries. DIS, FHB disease Index; FDK, fusarium damaged kernel percentage; CV, coefficient of variation. Statistically significant differences are denoted by an asterisk (*) where * denotes $P \leq 0.05$, ** denotes $P \leq 0.01$, and *** denotes $P \leq 0.001$.

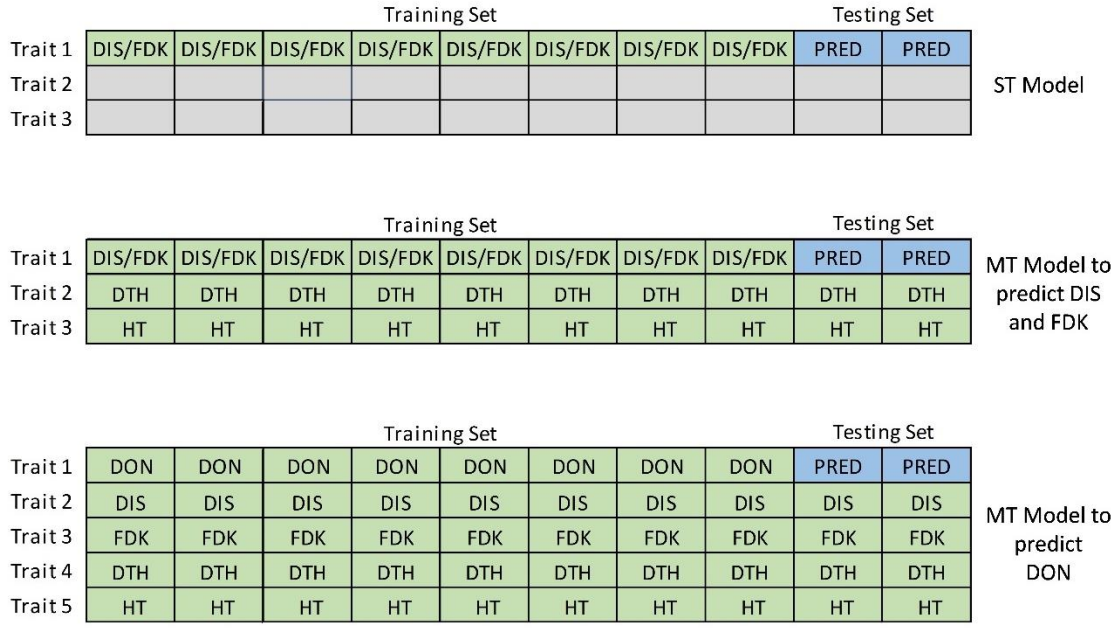
Trait	Year of evaluation	Mean square		
		Genotype	Residual	CV
DIS	2018	21.72***	36.08	13.97
	2019	170.02***	103.21	20.69
	2020	153.74***	84.12	21.84
FDK	2019	94.64***	59.84	9.66
	2020	151.77***	149.89	20.88

Supplementary Table S3. Mean prediction accuracy along with standard deviation (SD) and standard error (SE) for deoxynivalenol content (DON) with cross-validation in TP20 using the ST (rrBLUP) and MT genomic prediction models.

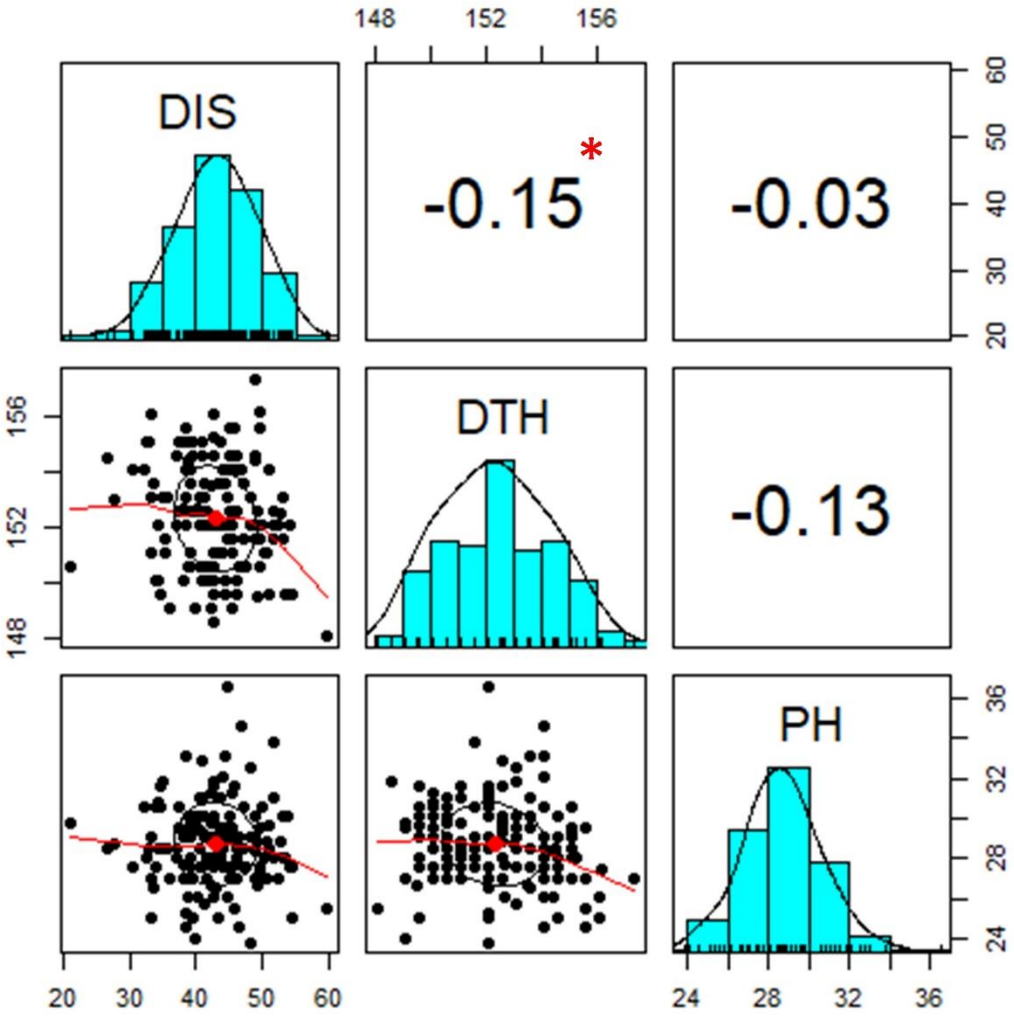
Model	Trait Combinations[a]	Predictive Ability	SD	SE
rrBLUP	-	0.49	0.13	0.01
MT	DIS, FDK, DTH, PH	0.56	0.10	0.01
MT	DIS, FDK, DTH	0.59	0.13	0.02
MT	DIS, FDK	0.54	0.11	0.02
MT	DIS, DTH	0.55	0.12	0.02
MT	FDK, DTH	0.58	0.11	0.02
MT	DIS	0.54	0.11	0.01
MT	FDK	0.56	0.11	0.02
MT	DTH	0.55	0.13	0.02

[a] Trait combinations refers to a different set of trait(s) used as secondary trait(s) in the MT model for prediction of DON. DIS, FHB disease Index; FDK, fusarium damaged kernel percentage; DTH, days to heading; PH, plant height.

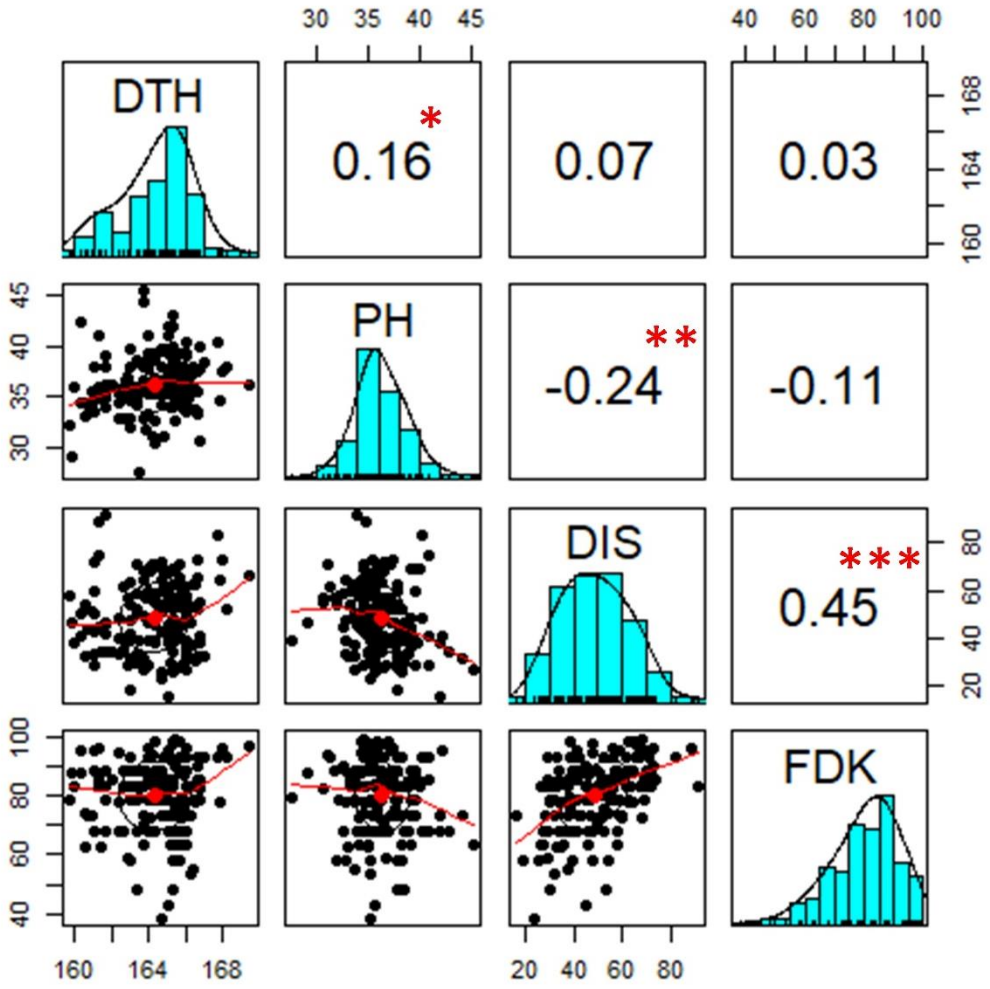
Supplementary Figure S1. Representation of the different cross-validation (CV) schemes used in this study. The Single trait (ST) Model was evaluated using a 80:20 CV scheme where four sets were used to train the model and the remaining set was used as a testing/validation set. The multi-trait (MT) model was used to predict DIS/FDK by using DTH and PH as secondary traits. For DON predictions, we used different combinations of secondary traits including DIS, FDK, DTH, and PH. The MT model used the genotypic and phenotypic information (primary and secondary traits) of individuals in the training set. DIS, FHB disease Index; FDK, fusarium damaged kernel percentage; DTH, days to heading; PH, plant height; DON, deoxynivalenol content.



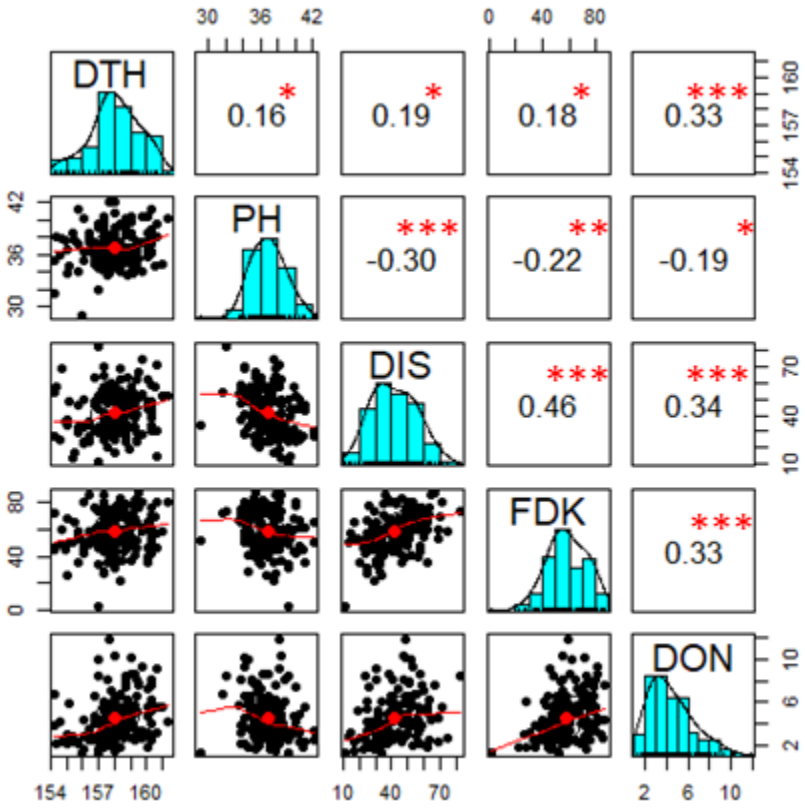
Supplementary Figure S2. Correlation coefficients among different traits in FHB nursery of 2017-18. DIS, FHB disease Index; DTH, days to heading; PH, plant height. Statistically significant correlations are denoted by an asterisk (*) where * denotes P ≤ 0.05, ** denotes P ≤ 0.01, and *** denotes P ≤ 0.001.



Supplementary Figure S3. Correlation coefficients among different traits in FHB nursery of 2018-19. DIS, FHB disease Index; FDK, fusarium damaged kernel percentage; DTH, days to heading; PH, plant height. Statistically significant correlations are denoted by an asterisk (*) where * denotes $P \leq 0.05$, ** denotes $P \leq 0.01$, and *** denotes $P \leq 0.001$.



Supplementary Figure S4. Correlation coefficients among different traits in FHB nursery of 2019-20. DIS, FHB disease Index; FDK, fusarium damaged kernel percentage; DTH, days to heading; PH, plant height; DON, deoxynivalenol content. Statistically significant correlations are denoted by an asterisk (*) where * denotes $P \leq 0.05$, ** denotes $P \leq 0.01$, and *** denotes $P \leq 0.001$.



Supplementary Figure S5. Scatterplots showing observed v/s predicted values for fusarium damaged kernel (FDK) percentage. The observed DIS estimates were based on phenotypic evaluation of 60 independent lines, and the predicted values are GEBVs using (A) TP19, (B) TP20, and (C) TP19+20. The red dashed line represents the cutoff (50%) to discard genotypes based on the observed data and estimated GEBVs. Genotype to the right side of the red dashed line would be discarded based on observed phenotype, and genotype above the red dashed line would be discarded based on GEBVs.