Genome Wide Association Analysis and Prediction of *Fusarium* Head Blight Resistance in Soft Red Winter Wheat

Dylan L. Larkin[1], R. Esten Mason[1], Amanda L. Holder[2], David E. Moon[1] Gina Brown-Guedira[3,4], and Stephen A. Harrison[5]

[1]Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR, USA 72701

[2]Department of Animal and Food Sciences, Oklahoma State University, Stillwater, OK, USA 74075

[3]Department of Crop and Soil Sciences, North Carolina State University, Raleigh, NC, USA 27607

[4]USDA-ARS Plant Science Research, Raleigh, NC, USA 27607

[5]School of Plant, Environmental, and Soil Sciences, Louisiana State University Agricultural Center, Baton Rouge, LA, USA 70803

*Presenting author email: dllarkin@uark.edu

Introduction

Fusarium head blight (FHB) is a disease in wheat (*Triticum aestivum* L.) caused by the fungal pathogen *Fusarium graminearum*. FHB poses potential economic losses and health risks due to the accumulation of the mycotoxin deoxynivalenol (DON) on infected seed heads. The objectives of this study were to identify novel FHB resistance loci using a genome wide association (GWAS) approach and determine if the use of markers as fixed effects in genomic selection models improve prediction accuracies for FHB resistance traits in a training population consisting of 354 soft red winter wheat (SRWW) lines. The population was evaluated in inoculated, misted FHB nurseries in Fayetteville and Newport, AR and Winnsboro, LA in a randomized complete block design from 2014-2017. Lines were phenotyped for resistance traits including DON accumulation, *Fusarium* damaged kernels (FDK), incidence, and severity. Fifty SNP markers were significantly ($p \leq 0.0001$) associated with the resistance traits across 19 chromosomes using the FarmCPU model. Thirteen significant SNPs were identified for DON, notably on chromosome 7AS. Ten were identified for FDK, notably on chromosomes 3AL, 3BL, 4BL, and 7DL. Twelve were identified for incidence, notably on chromosomes 2BS, 5DS, 4AL, and 7BL. While 15 were identified for severity, notably on chromosomes 3BL and 4BL. The naïve genomic selection model outperformed the fixed effect model for all four traits. Genomic prediction accuracies (r) for the naïve model were 0.607, 0.535, 0.342, and 0.550 for DON, FDK, incidence, and severity, respectively. Results from this study will facilitate the development of SRWW cultivars with improved resistance to FHB.

Objectives

1) Evaluate SRWW genotypes for four FHB resistance traits over multiple location-years.
2) Identify marker trait associations for four traits using a GWAS approach.
3) Perform a genomic prediction, cross-validation analysis to determine if a genomic selection model with marker fixed effects can improve prediction accuracy compared to a naïve model.

Materials and Methods

Plant material

- 354 soft red winter wheat lines.
- 238 lines from Arkansas, 40 lines from Georgia, and 38 lines each from Louisiana and North Carolina.

Phenotype data

- Lines were evaluated in inoculated, misted FHB nurseries in Fayetteville and Newport, AR from 2013-2017 and Winnsboro, LA in 2017.
- Planted in two row plots with two replications in a randomized complete block design.
- Inoculated with *F. graminearum* infected maize (*Zea mays* L.) and overhead misted for a total of 480 and 520 minutes, for Fayetteville and Newport, respectively, to provide optimal conditions for FHB infection.
- Field ratings (taken one week after heading).
  - Incidence rating: Scale of 0-100; rated twice one week apart.
  - Severity rating: Scale of 0-100; rated twice one week apart.
- Post Harvest ratings:
  - *Fusarium* damaged kernels (FDK): Scale of 0-100, once per rep on 100g of seed.
  - Deoxynivalenol (DON) analysis: gas chromatography.

Statistical analysis

- For each of the four FHB-associated phenotypic traits, best linear unbiased predictions (BLUPs) were calculated using the following model in SAS v9.4:

$$Y_{ijk} = \mu + env_i + rep\ (env)_{ij} + line_k + (env\ x\ line)_{ik} + \varepsilon_{ijk}$$

where Y_{ijk} is the observed phenotype, μ is the overall mean, env_i is the random effect of the ith environment, $rep(env)_{ij}$ is the random effect of jth rep within the ith environment, $line_k$ is the random effect of the kth line, $(env\ x\ line)_{ik}$ is the random effect of the interaction between the ith environment and the kth line, and ε_{ijk} is the random error term.

Genotypic Data

- Genotyping by Sequencing (GBS) was performed at the USDA-ARS Eastern Regional Small Grains Genotyping Laboratory in Raleigh, NC.
- Single nucleotide polymorphism (SNP) calling was performed using the TASSEL 5.0 pipeline and the Wheat IWGSC RefSeq assembly v1.0 as a reference genome.
- Lines were filtered based on greater than 50 percent missing genotypic data, greater than ten percent heterozygosity and minor allele frequency of less than five percent.
- Missing data imputed using LD*k*NN*i* function in TASSEL 5.0.
- Final genotypic dataset consisted of 72,634 GBS SNPs.

Genome Wide Association Study (GWAS)

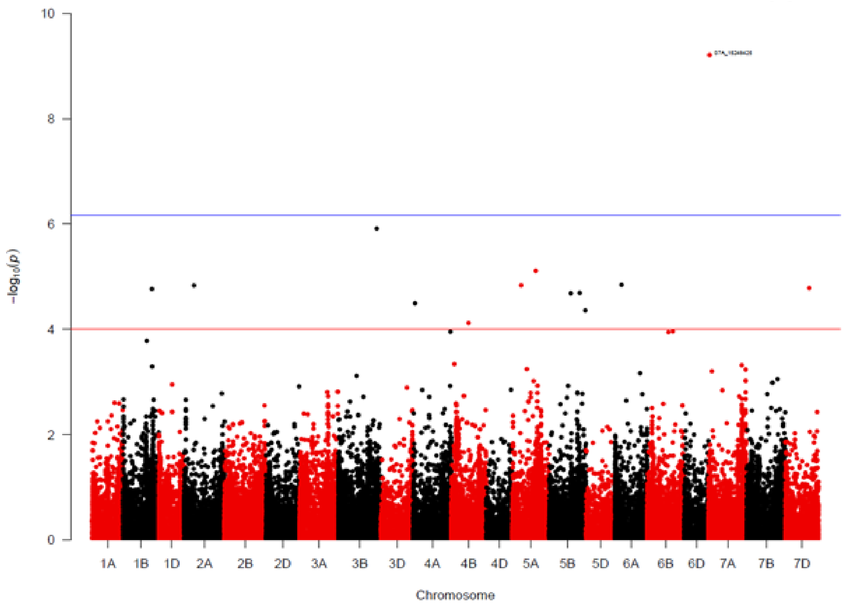
- GWAS was performed using the FarmCPU model within GAPIT Version 2 using the iPat interface with source code from R v3.6.1.
- Top seven principle components contributed to greater than two percent of variation and were included as model covariates to control for population structure.
- Marker trait associations significant at $p \leq 0.0001$.
- JBrowse tool on Wheat@URGI Portal used to identify genes from IWGSC RefSeq v1.1 annotation within 10.0 kbp of SNP.

Genomic Prediction Cross-Validation

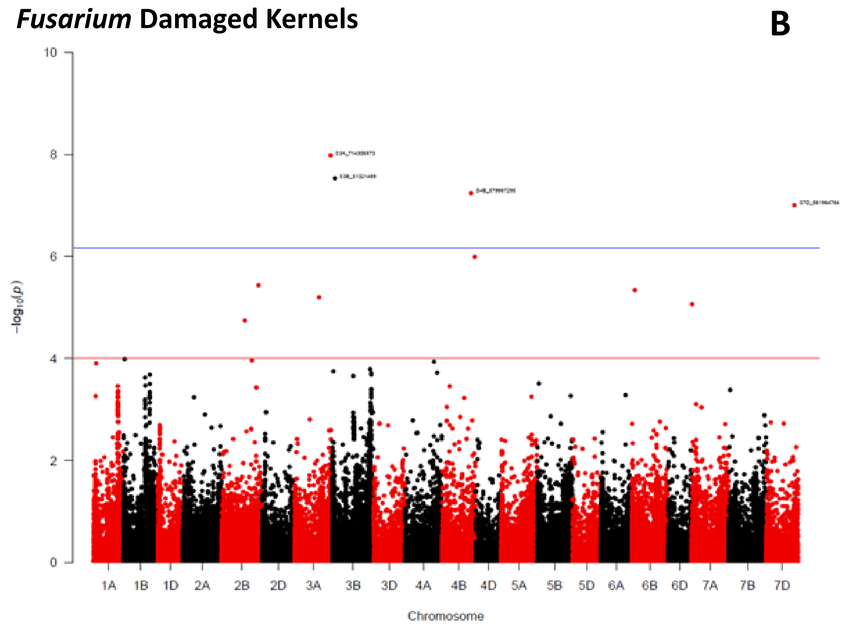
- Cross-validation analysis performed for all four traits using GBLUP model in rrBLUP package in iPat using source code from R v3.6.1 over 10 cycles.
- Marker fixed effects selected using FarmCPU GWAS of each training subpopulation ($p \leq 6.88 \times 10^{-8}$, $-\log_{10}(p) \geq 6.16$).
- Mean prediction accuracy compared between naïve and fixed effect GBLUP models using Fisher's LSD (α=0.05) in SAS v9.4.

Results and Discussion

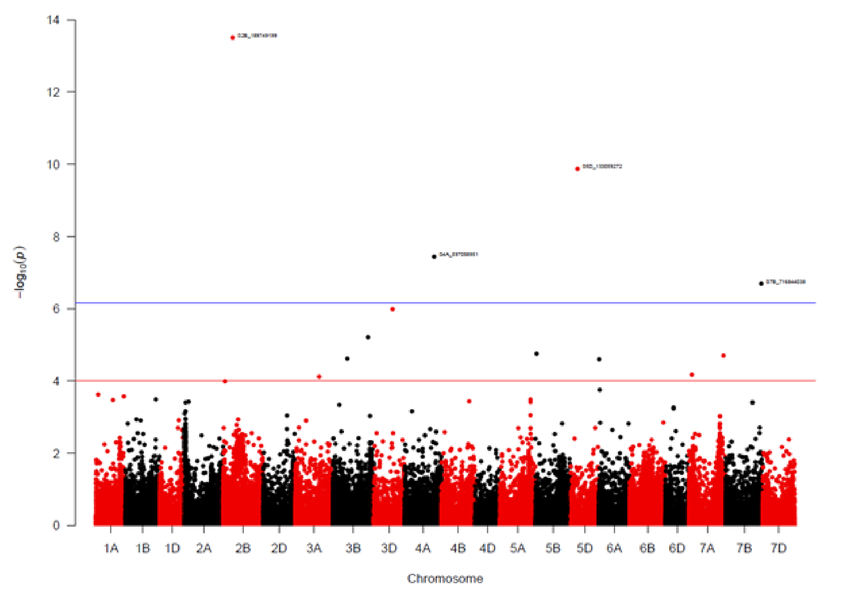
DON Accumulation **A**



***Fusarium* Damaged Kernels** **B**



Incidence **C**



Severity **D**

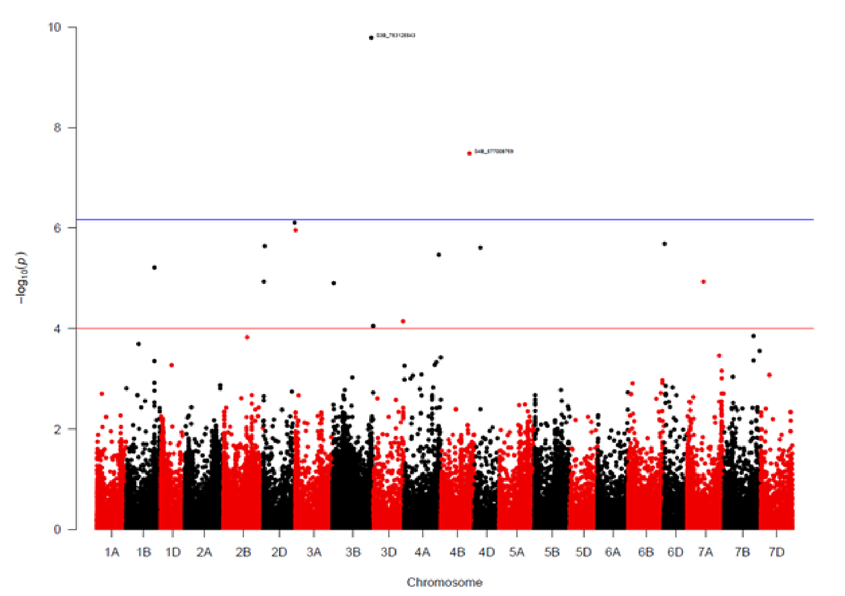


Figure 1. Manhattan plots showing genome-wide SNP loci associated with four traits impacting FHB resistance. The blue horizontal line represents the Bonferroni significance threshold ($p \leq 6.88 \times 10^{-8}$, $-\log_{10}(p) \geq 6.16$). The red horizontal line represents the threshold for declaring a marker significant based on $p \leq 0.0001$ ($-\log_{10}(p) \geq 4.00$). MTAs determined using the FarmCPU model with seven principle components. **(A)** Plot of genome-wide markers associated with DON accumulation using BLUPs. **(B)** Plot of genome-wide markers associated with FDK using BLUPs. **(C)** Plot of genome-wide markers associated with incidence using BLUPs. **(D)** Plot of genome-wide markers associated with severity using BLUPs.

Results and Discussion

Genome Wide Association Study

- GWAS identified a total of 50 significant SNP markers ($p \leq 0.0001$) for the four traits (Fig. 1).
- Significant SNP markers were identified on 19 of 21 wheat chromosomes.
- Chromosome 3B contained the highest number of significant SNPs with eight, followed by chromosomes 7A (5), 3A and 5B (4).
- **Incidence:** SNP on 2BS within 10.0 kbp of two high confidence genes and SNP on 4AL within 10.0 kbp of three high confidence genes.
- **Severity:** SNPs on 3BL and 4BL were both within 10 kbp of three high confidence genes.
- **FDK:** SNP on 3AL was present on the high confidence gene *TraesCS3A02G484800*, related to acetyl-CoA carboxyltranferase activity, contributing to cuticle formation and plant defense.
- A QTL on 4BL has been associated with severity in hard red winter wheat (Clinesmith et al., 2019).
- FHB resistance has been identified on 3BL in native soft red winter wheat germplasm in multiple QTL studies (Liu et al., 2013; Islam et al., 2016; Petersen et al., 2017).

Table 1. Highly significant SNPs ($-\log_{10}(p) \geq 6.16$) for four FHB resistance traits in wheat.

Trait	Chromosome	Position (bp)	Allele	*p*-value	LOD	MAF	Effect
DON (ppm)	7A	15248425	A/C	6.17E-10	9.21	0.093	-1.05
FDK (%)	3A	714306573	A/G	1.05E-08	7.98	0.225	-2.52
	3B	51321489	T/A	2.95E-08	7.53	0.202	2.90
	4B	579987295	G/A	5.77E-08	7.24	0.410	2.77
	7D	561964764	T/A	9.87E-08	7.01	0.116	3.60
Incidence (%)	2B	189749199	A/G	3.13E-14	13.50	0.175	-2.46
	5D	133059272	C/T	1.34E-10	9.87	0.184	-0.54
	4A	597058951	C/T	3.63E-08	7.44	0.236	-0.43
	7B	716844038	C/A	2.03E-07	6.69	0.064	0.69
Severity (%)	3B	783125543	T/A	1.63E-10	9.79	0.381	1.70
	4B	577008759	C/G	3.28E-08	7.48	0.417	1.63

Genomic Prediction Cross-Validation

- DON accumulation and severity had the highest mean prediction accuracies (Table 2).
- The naïve model significantly outperformed the fixed effect model for three out four traits (Table 2).
- Strongly associated markers accounted for less than ten percent variation (Bernardo, 2014).
- Fixed effects usually do not improve prediction accuracy if heritability greater than 50 percent (Bernardo, 2014).

Table 2. Comparison of mean prediction accuracies between a naïve and fixed effect genomic selection model.

Trait	H^2	Naïve Model (r)	Fixed Effects Model (r)	*p*-value (α = 0.05)
DON	0.79	**0.61**	0.54	0.0030
Severity	0.78	0.55	0.50	0.0800
FDK	0.82	**0.54**	0.48	0.0036
Incidence	0.38	**0.34**	0.27	0.0384

Conclusions

- Location-years were successfully analyzed for four FHB traits using data collected in inoculated screening nurseries.
- GWAS identified marker-trait associations for incidence, severity, FDK and DON.
- The fixed effect model did not improve prediction accuracy compared to the naïve model.
- AMP could be used as a training population to determine genomic estimated breeding values for FHB resistance.

Future Work

- Modify the population for forward prediction of advanced breeding lines in the University of Arkansas wheat breeding program.
- Compare the selection accuracy between genomic selection and phenotypic selection of advanced breeding lines within the University of Arkansas wheat breeding program.

References

Bernardo, R. 2014. Genomewide selection when major genes are known. Crop Science 54: 68-75.

Clinesmith, M.A., A.K. Fritz, C. Lemes da Silva, W.W. Bockus, J.A. Poland, F.E. Dowell, et al. 2019. QTL Mapping of Fusarium Head Blight Resistance in Winter Wheat Cultivars 'Art'and 'Everest'. Crop Science 59: 911-924.

Islam, M.S., G. Brown-Guedira, D. Van Sanford, H. Ohm, Y. Dong and A.L. McKendry. 2016. Novel QTL associated with the Fusarium head blight resistance in Truman soft red winter wheat. Euphytica 207: 571-592.

Liu, S., C.A. Griffey, M.D. Hall, A.L. McKendry, J. Chen, W.S. Brooks, et al. 2013. Molecular characterization of field resistance to Fusarium head blight in two US soft red winter wheat cultivars. Theoretical and applied genetics 126: 2485-2498.

Petersen, S., J.H. Lyerly, A.L. McKendry, M.S. Islam, G. Brown-Guedira, C. Cowger, et al. 2017. Validation of Fusarium head blight resistance QTL in US winter wheat. Crop Science 57: 1-12.

Acknowledgments

This work was supported by the United States Wheat and Barley Scab Initiative and in collaboration with SunGrains. DON analysis was conducted by the USDA Mycotoxin Diagnostic Laboratory in the Department of Plant Pathology Department at the University of Minnesota under the direction of Dr. Yanhong Dong.