Project 1: Haplotype-Informed Prediction of FHB Resistance in US Wheat Breeding Programs

---

1. What are the major goals and objectives of the research project?

1- Expand the diversity of PHG database by incorporating soft and spring wheat germplasm.
2- Map novel FHB-resistance QTLs using BLUEs by GWAS and develop diagnostic markers for FHB resistance QTLs for use in breeding programs.
3- Build a GS model for FHB resistance for HWW using field-generated phenotypic data
4- Test/optimize the GS model on two sets of breeding lines from Kansas and the Great Plains region.

2. What was accomplished under these goals or objectives? *(For each major goal/objective, address these three items below.)*

1- Expand the diversity of PHG database by incorporating soft and spring wheat germplasm.

a) What were the major activities?

This was completed last year (FY22). We generated ~15x exome coverage sequencing data of 60 soft and 25 spring wheat varieties included into the custom PHG database. This year (FY23), we completed promoter capture sequencing (~1.8 billion Illumina 2 x 151bp reads) data for the 291 lines increasing variants for both mapping and selection objectives. This results in nearly 6.2 million reads per variety, and coverage ~1.5x to the promoter and regulatory target regions (~200 Mb) in the wheat genome.

b) What were the significant results?

With the addition of the ~90 lines (FY22) of exome capture data detected > 1.028 million variants (MAF $\geq$ 0.01) across the wheat genome using the PHG database for variant detection. The promoter capture (FY23) data produced ~142,000 variants across the wheat regulatory space, including targets not captured using exome only, increasing the variants used for GWAS.

c) List key outcomes or other achievements.

There are 291 varieties in the database using exome capture, including representative data from all marker classes. These 291 taxa genotypes are the varieties we are testing in the greenhouse and field experiments used to associate genotype and phenotype data to identify FHB resistance/susceptibility loci. These variants are used in the genomic selection modeling to provide meaningful selection criteria.

2- Map novel FHB-resistance QTLs using BLUEs by GWAS and develop diagnostic markers for FHB resistance QTLs for use in breeding programs.

a) What were the major activities?

FY22 (May 2022 – June 2023) phenotyping data for the second season at the KSU FHB nursery completed. Lines were evaluated for heading date, FHB incidence and severity, FDK, and DON content. FY23 (May 2023 – June 2024) is the third season of field data on this population. Data are being collected this spring and summer scoring heading date, FHB incidence and severity, FDK, and DON. Field conditions exhibited infection and plants were able to proceed with development until harvest, suggesting a successful year of field investigation.

b) What were the significant results?

BLUEs for the first 2 field seasons were generated this past fall and produced 10 significant associations with AUDPC, and 11 significant associations for FHB severity measured at 21 days

post-heading at *p-value* <=10^{-5} using GAPIT FarmCPU method. Significant associations were also detected for FDK, and DON, as well as using other GAPIT models. BLUEs for 3-year data are currently being generated for AUDPC and FHB severity, and post-harvest we will calculate BLUEs for FDK and DON and confirm significant associations.

c) List key outcomes or other achievements.

The 2-year BLUE significant associations for AUDPC and severity (post-heading day 20/21), reside on 11 wheat chromosomes. Most of the detected regions are reported in previous GWAS studies providing replication for other studies and validation for our methods. Four of the GWAS hits appear novel (defined as >100Mb from previous studies based on blasting flanking sequence to CSv2.1), which provide targets for KASP marker or targeted sequencing design to delineate resistance/susceptible alleles without producing whole genome sequencing. We will validate these results with the additional year of data, verify novelty, and generate diagnostic markers this fall.

As a follow-up and validation of the PHG format for generating and storing haplotypes, we extracted haplotypes from a GWAS significant region on chromosome 6A for AUDPC. Five haplotypes are present in our population, and one haplotype increases FHB susceptibility nearly 20% (data presented at FHB forum 2023). These results provide positive evidence that data mining and exploration using the PHG database can identify haplotypes carrying FHB resistance/susceptibility factors that can be incorporated by breeding programs to increase FHB resistance within their programs.

3- Build a GS model for FHB resistance for HWW using field phenotypic data

a) What were the major activities?

Preliminary models for GS utilizing all variants and our two-year phenotypic data were generated using the rrBLUP package in R this past fall. Data were split into 80% training population and 20% testing populations, and models were permuted 100 times. Preliminary models were used to get acquainted with the types of data needed to generate models. We are expanding this section to generate models for other traits, such as FDK and DON, and increase the permutations to 1000.

b) What were the significant results?

The accuracy of our model using FHB severity as AUDPC was predicted at 63%. We also split data into 50% training and 50% testing populations, and found AUDPC predictive ability to be 0.56, suggesting a trade-off in amount of data used to train and predictive accuracies, but lends credibility to using these models on related populations and across different years.

c) List key outcomes or other achievements.

We will test our models using the previous two years of field data, to predict the phenotypes for this year's (FY23) field data. We anticipate ~60-65% predictive ability based on preliminary results but hope to increase accuracy within the same location across years. Once an acceptable model is generated, we will test predictive abilities in related breeding material from the Great Plains (Objective 4).

4- Test/optimize the GS model on two sets of breeding lines from Kansas and the Great Plains region.

a) What were the major activities?

We collaborated with Allan Fritz (KSU) to get phenotype information for ~ 400 lines in his breeding program, which were scored in the Rocky Ford KSU FHB nursery for years 2022 and 2023. We genotyped this population using a skim-exome sequencing approach in February 2024. This generated ~500 million paired-end reads (2 x 151bp Illumina), resulting in 1.2 million reads per breeding line, producing ~0.5x target coverage. These sequencing reads are an input to the

PHG imputation pipeline to generate an imputed set of markers to test predictive abilities on our model from Objective 3.

In addition, hard winter wheat material from the central Plains regional nurseries were phenotyped this spring and summer 2023 (FY2023) by Jessica Rupp (KSU). Seeds from this material were sent to genotype in the central small grains genotyping facility (Guihua Bai, Manhattan, KS) using a mid-density (~5,000 markers) AgriSeq array.

b) What were the significant results?

Genotype data need to be returned from AgriSeq to run models on central Plains germplasm. Genotype data from A. Fritz program are currently being imputed with PHG developed in Objective 1.

c) List key outcomes or other achievements.

Waiting on data to learn outcomes. We will run the optimal model from Objective 3 on these data to test predictive abilities in both Kansas and Great Plains breeding material when data are available. We will share the outcomes with USWBSI. We are very interested in comparing approaches on predictive accuracies, (predicting with imputed PHG skim-sequenced data), or a mid-density targeted genotyping approach from hard winter wheat germplasm.

3. What opportunities for training and professional development has the project provided?

Graduate Student, Lawrence Tidakbi began his studies at Kansas State University in the Plant Pathology Department in August 2022 (FY22). He presented a poster and a Flash and Dash talk at the FHB Forum in Cincinnati in December 2023, and a poster in San Deigo at PAG in January 2024. This builds on his previous presentations at both conferences. Each opportunity gave him a chance to discuss our project with research scientists and develop collaborative relationships beyond what he presented his first year. Lawrence also attended the 2024 Fusarium workshop hosted at Kansas State University and presented his research at the north-central American Phytopathological Society (APS) Meeting in Manhattan Kansas this summer.

PI Katherine Jordan attended the FHB Forum in Cincinnati in December 2023, where we were able to talk future collaborations with this genotype database with other SCAB Initiative scientists, including future collaborations.

4. How have the results been disseminated to communities of interest?

In addition to the posters presented at professional conferences (FHB Forum and PAG), we have shared PHG variant data with the breeders that supplied lines used in our field study for their use in their own breeding programs for other phenotypic traits. We are waiting on third year results (for replication and BLUEs) before we publish the association results formally.

5. What do you plan to do during the next reporting period to accomplish the goals and objectives?

Not applicable, this is the final report for the agreement. We anticipate publishing results from GWAS and haplotype analysis to a peer-reviewed journal within the next year (waiting on FDK and DON data from FY23). We also anticipate publishing the genomic selection results and models in a peer-reviewed journal hopefully next summer.