Project 1: Metagenomics Informed Trait Development for Breeders

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

In this research, we aim to unravel the complex interactions between barley genotypes and microbial communities in the context of Fusarium Head Blight (FHB) infection. In the first year, our objective was to identify specific microbes that are recruited by FHB and those that are responsive to different barley genotypes. This involves extensive sampling from four U.S. nurseries, measuring Fusarium biomass and DON levels, and profiling the microbial communities associated with diseased and non-diseased spikes. We discovered that 2021 was an extreme drought year and needed additional analysis from the 2022 sample collection, along with employing metagenomics to resolve gene and taxa level changes caused by FHB. In the second year, we decided to delve deeper into the amplicon sequencing of barley genotypes with an additional year and now have analyzed the bacterial microbiome for 2022. We have been developing advanced analysis pipelines to identify disease and genotype-responsive taxa from the 2-year amplicon sequencing. We aim to augment these data with deep metagenomic sequencing of select samples to elucidate genetic and genomic responses within the microbiome to FHB. These data could enable the development of biomarkers that could facilitate breeding for enhanced microbiome associations that may reduce FHB.

FY22 Year 1 TSCI

Obj 1: Identify FHB recruited and genotype-responsive microbes, under non-drought conditions (2022)

- Rate and collect 2022 Training Population (TP)
  - 10 random genotypes x 4 nurseries x 5 diseased-spikes x 5 nondseased-spikes
- DNA extraction/ Fusarium biomass measurements/ DON measurements
- ITS and 16s Amplicon microbial community profiling
- Data analysis

FY23 Year 2 TSCI

Obj 2: Metagenome analysis of barley genotypes with identified FHB recruited and genotype-responsive microbes

- Sample preparation for metagenome sequencing
- In-depth Metagenomic sequencing on Novaseq
- Develop workflow and data analysis to identify gene and genome response to FHB and genotype

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

What were the major activities?

Major activities of this project include the following:

1. The Yr2 barley material (from the 2022 planting in Fargo, St Paul and Ithaca) was sequenced using the new NextSeq technologies, and simultaneously compared to previous MiSeq technologies. The NextSeq technologies led to substantial improvements in read depth and recovery of rare microbial taxa from plant material, with up to 6x the sampling depth across samples. Both the bacterial and fungal kingdoms of the microbiome were processed using the same set of primers as in yr1 for inter-year comparison of common microbial taxa.
2. The same 10 genotypes from 2021 and 2022 were planted and harvested at a single FHB nursery in ID in 2023. The samples from this nursery will provide for a second year of microbiome data for ID, which had poor disease in 2022. The ID 2023 samples are currently being processed.
3. We resequenced Yr1 samples that had previously shown high levels of mitochondrial and chloroplast reads with higher PNA concentrations to improve data quality.
4. Additionally, we refined the core alpha and beta diversity analysis from the Y1 bacterial and fungal datasets. The improved data quality for the resequenced samples in Yr1 also allowed us to perform more in-depth analyses.
5. We performed several experiments to optimize a protocol for metagenomic sequencing of barley head microbiomes, including a pilot sequencing run, and optimization of a sonication workflow for microbiome extraction from heads prior to sequencing.

What were the significant results?

In 2021, the observed richness (# of unique taxa) for the bacterial and fungal kingdoms varied by a significant three-way interaction with Location, Disease Status, and Line. Interestingly, the locations with the greatest richness varied by location. Specifically, bacterial richness was highest in ID and MN, while the fungal richness was highest in NY.

Location had a dominating effect on both bacterial and fungal community structure (i.e., beta-diversity) - with the ID location showing the most divergent community structure compared to the more eastern locations (**Figure 1**). Because location was such a driving force in the community structure of both microbiomes, we separated the data by location to further assess the importance of disease and line on community structure. We found a significant two-way interaction of disease and line on microbiome structure for all locations and kingdoms except for the bacteria in the NY samples (**Figure 2**).

Additionally, we performed a variation partitioning analysis on the microbiome data separated by location. Variation partitioning is performed using distance-based redundancy analysis, a type of constrained ordination. This analysis showed that disease status explained the most variation in community structure in 2021 at the MN location (6.0% for bacteria and 11.0% for fungi; **Table 1**). For all other locations in 2021, barley line explained more variation in community structure than disease. Interestingly, this also appeared to correlate with overall disease pressure - which was substantially higher in the MN nursery relative to the other locations in 2021. Additionally, disease

tended to explain more variation in disease in the fungal communities, than in the bacterial communities. Once processed, the Yr2 data is expected to help us further test these qualitative patterns and determine the significance of overall disease pressure and kingdom type on community structure patterns.

Metagenomic sequencing of barley head DNA directly produces insufficient microbial DNA for metagenomic analysis (>90% of reads mapping to the barley genome). We found that sonication is effective at isolating bacterial and fungal communities from barley heads and optimized the methodology to retrieve sufficient DNA for metagenomic sequencing.

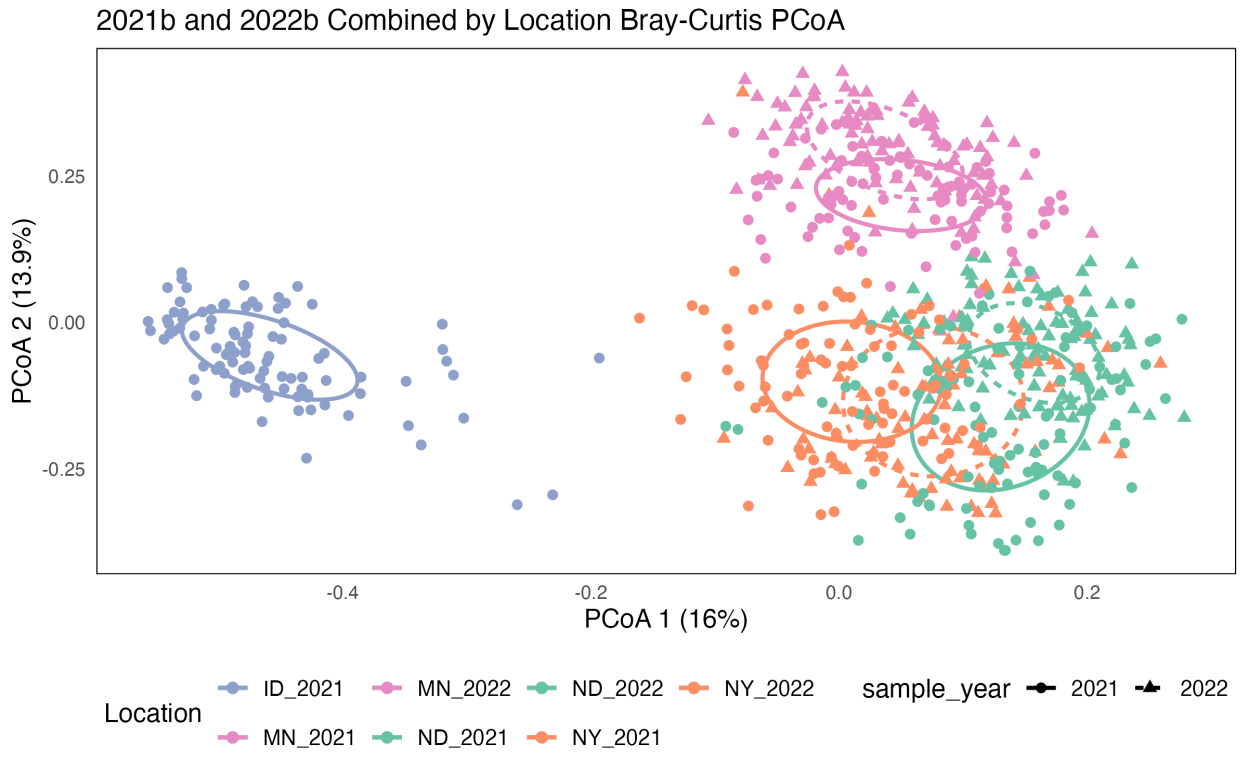

Figure 1. Location is a major driver of bacterial microbiome composition across years.

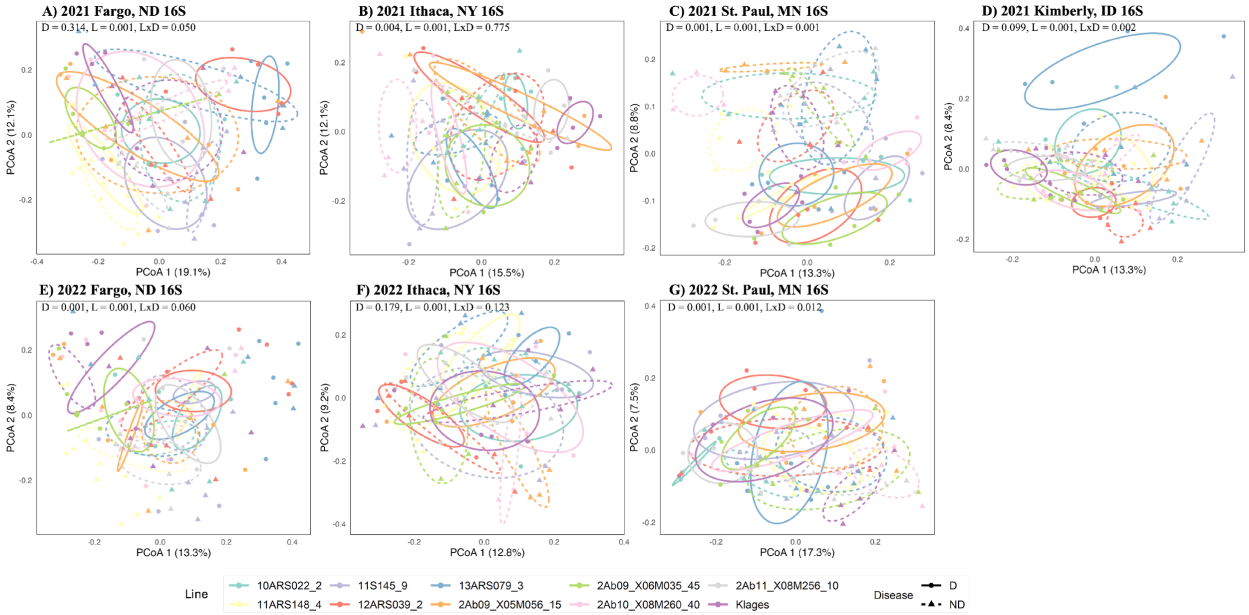

Figure 2. Impact of disease and genotype on bacterial microbiome composition in 2021 and 2022.

		ID	MN	ND	NY
Bacteria	**Disease**	0.2%	6.0%	0.0%	0.6%
	Line	10.4%	5.6%	22.1%	11.4%
Fungi	**Disease**	1.9%	11.0%	2.1%	1.4%
	Line	5.7%	7.3%	17.2%	9.6%

Table 1. Variation partitioning analysis results indicating importance of disease and genotype on bacterial and fungal microbiome composition in 2021.

List key outcomes or other achievements.

- Sequencing of 2022 samples (~300 samples from 3 locations)
- Refined data analysis of community composition
- Completed pilot metagenomics sequencing experiments

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

- Presented at 2023 Scab initiative with an invited oral presentation by Dr. Barney Geddes "The Impact of Environment, Host Genotype and Fusarium Head Blight on Microbiome Assembly in a Barley Breeding Population"
- Oral presentation by Dr. Briana Whitaker at Ecological Society of America 2023 "Host Genotype & Fusarium head blight filtering in the barley head microbiome"
- Two 2023 Scab initiative posters by students "Experimentally Tractable Systems for Investigating Fusarium Head Blight-Microbiome Interactions on Barley: A Pilot Study" and "Host Genotype and Fusarium Head Blight Status Impact Microbiome Assembly of a Barley Breeding Population Across Four Locations"
- IS-MPMI 2023, Providence, RI (Poster)
- Plant Health 2023, Denver, CO (Poster and Travel Award for Microbiome workshop)

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

   The results have been disseminated to the scientific communities through the posters and oral presentations mentioned above. Farmers and industry have been approached in conversation from those meetings and commodity group meetings on the progress of this research. Letters of support for this research were received from both the American Malt Barley Association and North Dakota Barley Council to continue and expand this work by applying to federal grants.

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

   We are nearing complete data collection for both Year 1 and Year 2 bacterial/fungal amplicon sequencing datasets. To be done are 1) sequencing of 2023 Idaho samples (2022 samples were compromised and recollected the following year in this location), 2) DON analysis of Year 2 samples.

   A robust data analysis pipeline has now been developed and used for Year 1 samples. This pipeline will be used to analyze the complete Year 1/2 dataset and finalize determinations of 1) the impact of disease on the barley head microbiome, 2) the impact of barley line on the disease response in the head microbiome, and 3) prediction of genotype/disease-responsive taxa.

   We plan to rerun metagenomic sequencing on altered sample collection methodologies designed to minimize barley DNA to enable the depth of sequencing required for robust metagenomic analysis. Using metagenomics data we will attempt to assemble metagenome-assembled genomes (MAGs) belonging to line/disease responsive taxa identified from amplicon sequencing data. Alternatively, we will use metagenomic read recruitment to map reads to genomes of isolates of these taxa collected during sampling. In either case, we will 1) evaluate the abundance of reads matching the corresponding genomes in disease and no-disease across a subsample of lines, and 2) identify genomic signatures from responsive genomes/genomic loci that can be used as qPCR targets to test the enrichment or depletion of the line/disease responsive taxa in training populations.