Genome-wide Association Studies Used to Find Correlations Between Phenotypic Traits of Wheat Plants

Guadalupe Iniguez[1,2], Sheri Rynearson[3], Peter Schmuker[4], Michael Pumphrey[1]

Introduction

Main objective:

Investigate the genetic underpinnings of key traits in the spring wheat breeding populations.

Research Question:

How do wheat phenotypic traits correlate with one another and then how we can optimize selection methods to improve multiple traits in the breeding program simultaneously?

Methods

Phenotypic Data:	Multispectral Data:
Location: Spillman Farm	Location: Spillman Farm Used Sentera 4x multispectral sensor on an DJI inspire 2 drone
- Height - Yield - Test Weight (TW) - Maturity - Protein	- Normalized Difference Vegetation Index [NDVI=(NIR-Red)/(NIR+Red)] - Normalized Difference Red Edge [NDRE=(NIR- Red Edge)/(NIR+Red Edge)] - Normalized Water index [NWI 1=(R_{970}-R_{900})/(R_{970}+R_{900})] - Normalized Water Index [NWI 2=(R_{970}+R_{850})/(R_{970}+R_{850})]

Genotype by sequencing was done by North Carolina State University. The 2021 data was then used in a HapMap file with GWAS in R studio to find associations between SNPs and traits. Also was about to fin correlations between traits.
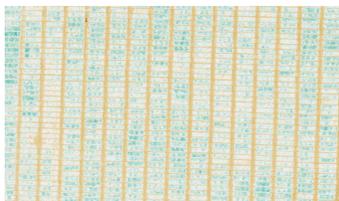
Results

Figure 1: NDVI image taken by multispectral camara on drone. Spillman farm plots.

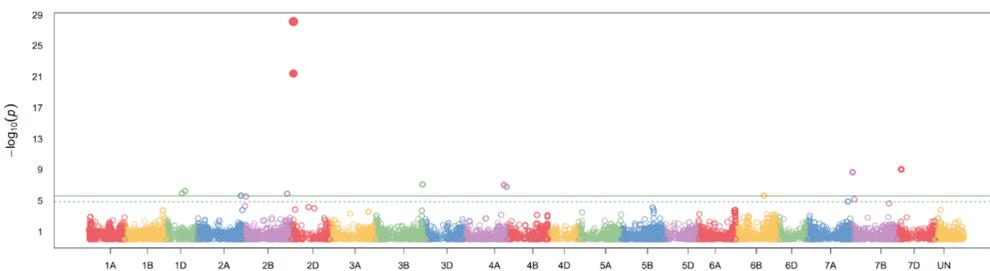### Associations found in Manhattan Plots between SNPs and Traits

Figure 2: Maturity has high association to Chromosome 2D
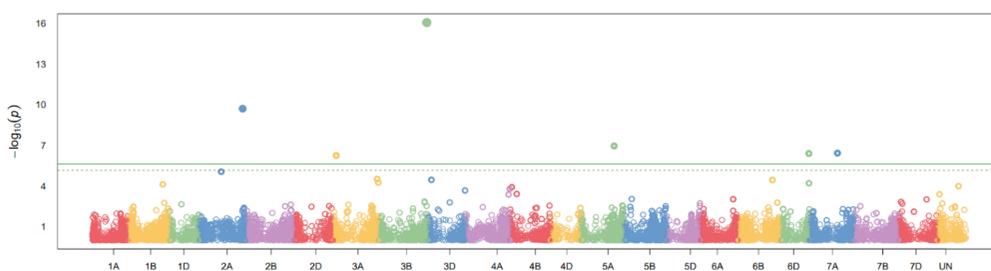

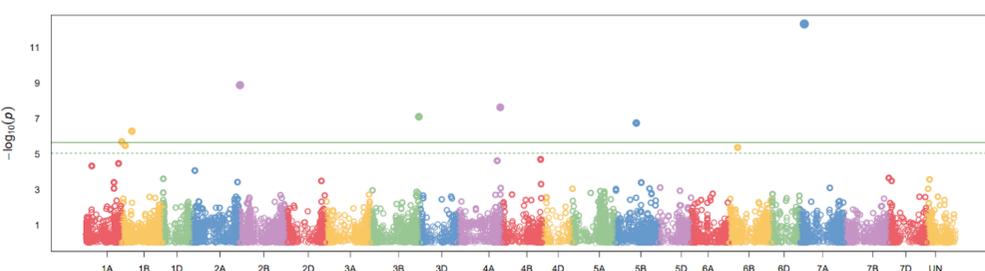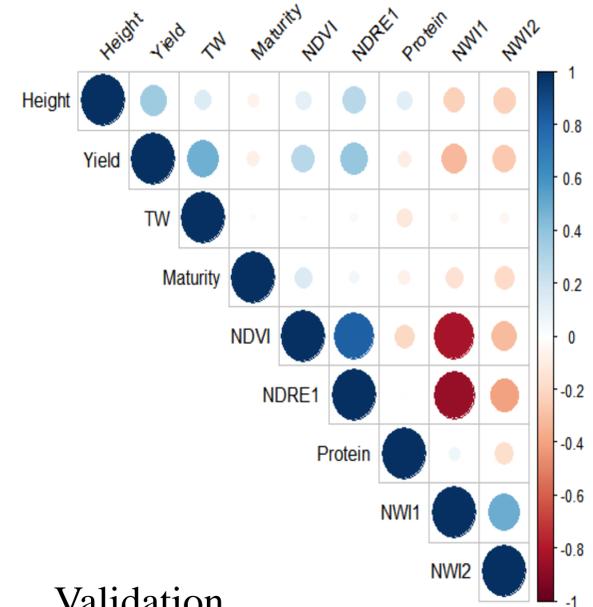Figure 3: Yield has high association to chromosome 3B

Figure 4 Test Weight has high association to chromosome 7A

Correlations Between Traits

Figure 5: Correlation Heat Map. Correlations of 0.36 between Yield and Height. Correlation of 0.49 between Test Weight and Yield. Correlation of 0.38 between NDRE and Yield.

Validation

Single-nucleotide polymorphism (SNPs)	Gbrowser Grain Genes Similar Findings
4A Maturity	Found on T3/Wheat: Marker WCSS1_contig7118743_4AL-3392
7D Maturity	Found on T3/Wheat: Marker gbsHWWAMP30278
1B Protein	Found on T3/Wheat: Marker Tdurum_contig31094_386
2A Yield	Found on T3/Wheat: Marker gbsHWWAMP19292
2A Test Weight	Found on T3/Wheat: Marker gbsHWWAMP14546
2B Test Weight	Found on T3/Wheat: Marker gbsHWWAMP14863

Figure 6: Shows that other research groups have found similar traits show some relationship to the SNPs found

Conclusion

When looking through the data many correlations stood out. There was positive correlations between yield and NDVI and NDRE 1, and when they are positive higher yield is expected. We also had a positive correlation between test weight and yield implying that as yield went up so did test weight, we can then look for the associated markers that were seen in the Manhattan plots and focus on determining whether these wheat varieties yield the most and least. We also found a negative correlation between both NWI 1 and 2 with yield which means that yield was increasing in this case because lower NWI levels are better.

Future Directions

We can see if last years drought affected the 2021 preliminary data and our results by completing the same experiment with 2022's collected data. Repetition in future years can also increase the validation of the data found during this summers research. Furthermore, learning how the correlations between certain traits and their associated markers can be used for wheat selection.

Acknowledgements

I would like to acknowledge Washington State University and their Research and Extension Experiences for Undergraduates (REEU) program. I would also like to thank Dr. Sankaran Sindhuja for guiding us through the research experience. I am honored to work in Dr. Mike Pumphrey's Spring wheat breeding program and I appreciate all the help I got from Morgan Hasler, Peter Schmuker, Tyson Koepke, Ruth Uwugiaren, and Sheri Rynearson.

References

Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. PLOS Genetics 12(2): e1005767. https://doi.org/10.1371/journal.pgen.1005767

RStudio Team (2022). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Home. Sentera. (2022, May 19). Retrieved July 26, 2022, from https://sentera.com/

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

Yao E., Blake V.C, Cooper L., Wight C.P., Michel S., Cagirici H.B., Lazo G.R., Birkett C.L., Waring D.J., Jannink J-L., Holmes, I., Waters A.J., Eickholt D.P., Sen, T.Z., "GrainGenes: A Data-Rich Repository for Small Grains Genetics and Genomics", Database, Volume 2022, 2022, baac034, https://doi.org/10.1093/database/baac034

Zhang, Y., Cheng, Y., Jiang, W., Ye, Y., Lu, Q., & Zhao, H. (2021, January 26). *Comparison of methods for estimating genetic correlation between complex traits using GWAS summary statistics*. OUP Academic. Retrieved July 26, 2022, from https://academic.oup.com/bib/article/22/5/bbaa442/6120325