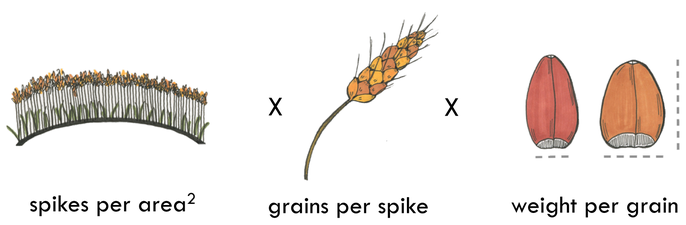
Integrating Multi-omics Data to Fine-map Wheat Grain Weight and Morphology Genes

Ella Taagen, and Dr. Mark Sorrells
Section of Plant Breeding and Genetics
School of Integrative Plant Science
Cornell University, Ithaca, NY, USA

Grain weight and width are significantly associated with **two** nearby regions on chromosome 5A.

1 Yield is a balancing act

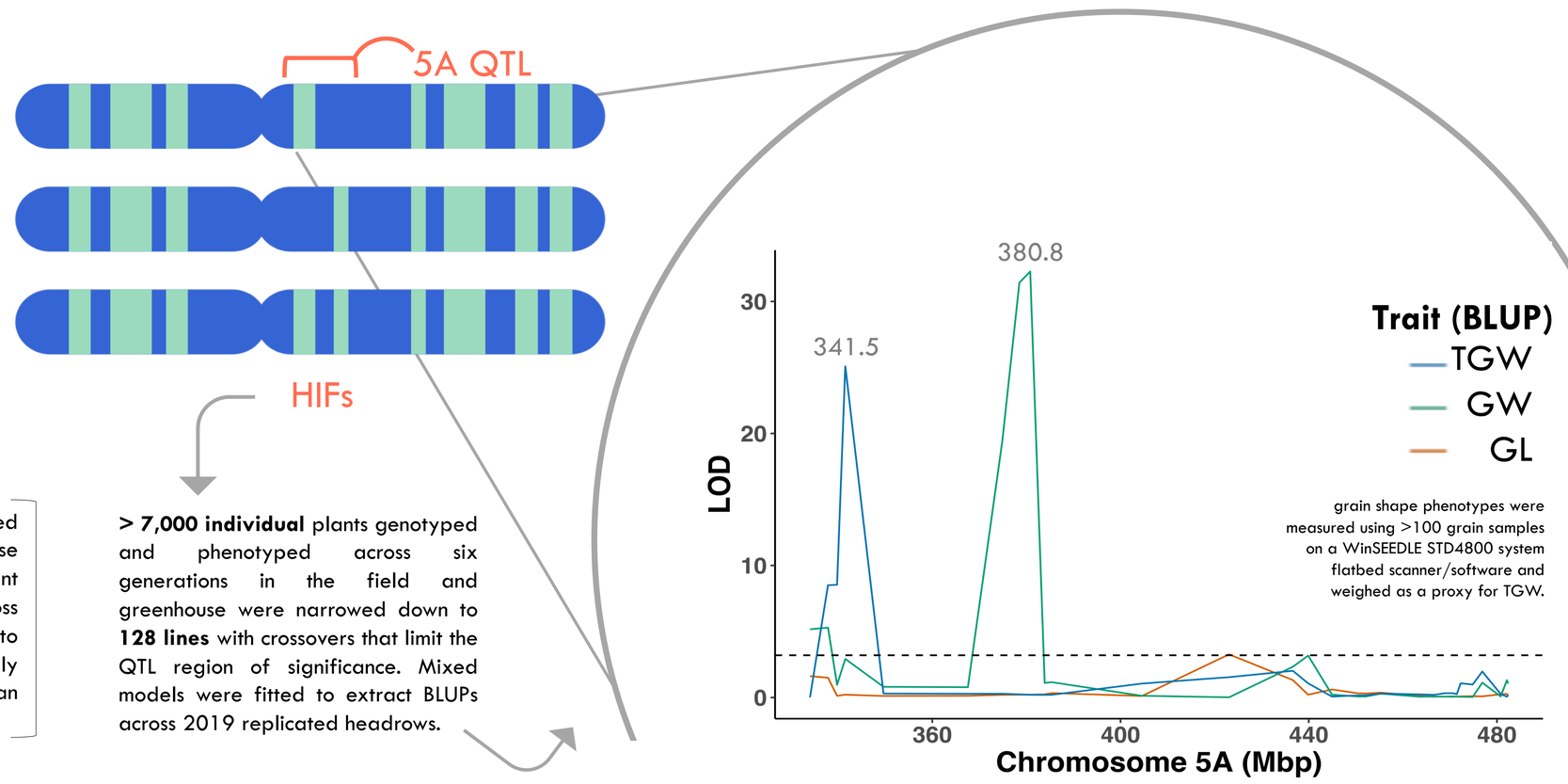
Cereal crop yield = spikes per area^2 X grains per spike X weight per grain



SynOp DH with Opata allele 5A QTL **4.9% heavier** on average compared to W7984 allele

Recent advancement of the wheat reference genome assembly and genome editing tools can help facilitate the characterization of genes underlying quantitative trait loci (QTL) for yield components. Grain weight and morphology are **valuable traits** to consider when releasing a new wheat variety because they can impact the number of kernels it takes to fill a bushel and milling quality.

A QTL for thousand grain weight (TGW) and grain width was mapped using R/qtl (Broman et al., 2003) to a 100 Mbp region on chromosome 5AL in the **W7984 Synthetic x Opata M85** spring bread wheat doubled haploid population (SynOp DH, 145-line subset of 215, Sorrells et al., 2011). Mixed models were fitted to extract BLUPs across 6 site-year combinations.



2 HIFs enhance genetic resolution of QTL

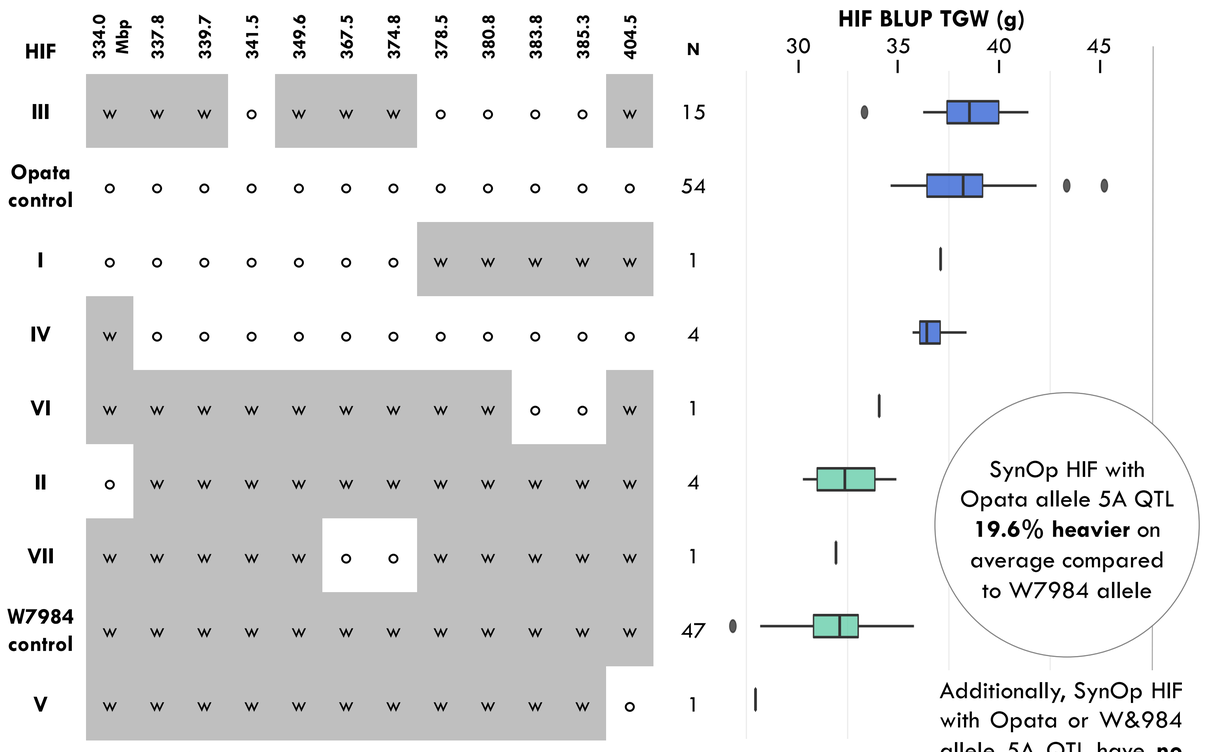
Fine-mapping a causal genetic variant is **limited by crossovers** that disrupt linked markers. We developed three heterogeneous inbred families (HIFs) $F_{6:6}$ to increase genetic resolution of the 5A QTL. HIFs were selected from two SynOp recombinant inbred lines (SynOp RIL 2,039 F_6 lines, Sorrells et al., 2011) based on heterozygosity across the 5A QTL. The individual progeny were phenotyped and genotyped in order to track recombination events across the 5A QTL. The resulting HIFs have highly homogenous background genomes and distinct crossovers across the QTL that can be confidently associated with the line's phenotype (Tuinstra et al., 1997).

> 7,000 individual plants genotyped and phenotyped across six generations in the field and greenhouse were narrowed down to **128 lines** with crossovers that limit the QTL region of significance. Mixed models were fitted to extract BLUPs across 2019 replicated headrows.

3 Comparing HIF crossovers further explains two QTL peaks

2019 HIFs grown in headrows and genotyped with KASP markers across the 5A QTL were analyzed with R/qtl, narrowing the significant sequence to a **9.9 and 10.5 Mbp region.**

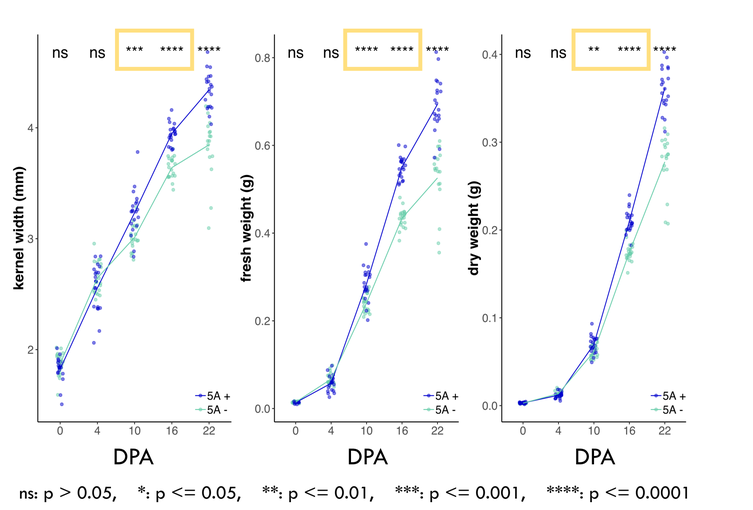
Tests for epistasis between the most significant markers were inconclusive because a simple linear interaction model cannot estimate an interaction from the present genotype frequencies (341.5 : 380.8 Mbp, 73 o:o, 54 w:w, 1 o:w, 0 w:o). The two QTL peaks could be due to multiple causal variants, or linkage. **HIF I is** the only line with a recombination event resulting in o:w genotype for the most significant markers and will be very useful for comparison and understanding the two-peak phenomenon in a gene expression study.



SynOp HIF with Opata allele 5A QTL **19.6% heavier** on average compared to W7984 allele

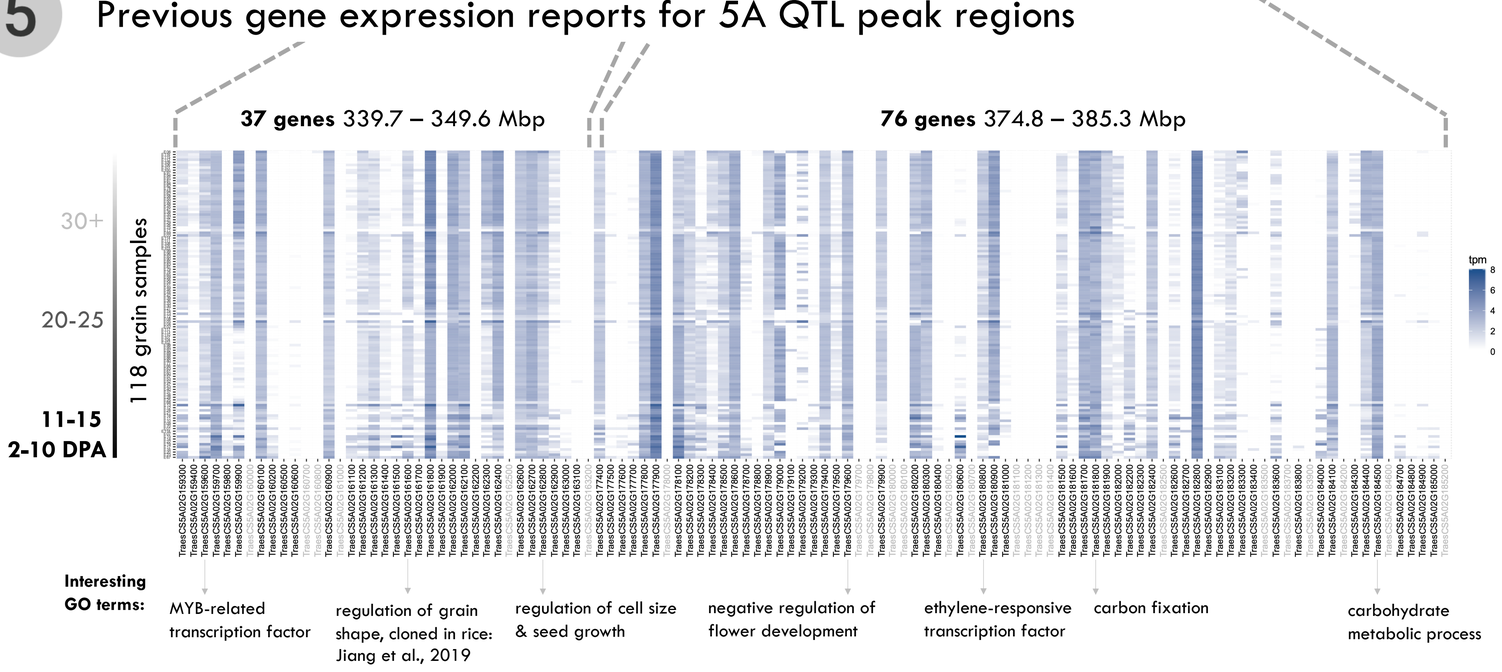
Additionally, SynOp HIF with Opata or W&984 allele 5A QTL have **no significant difference** in grain fill duration or spikelets / spike.

4 grain development time sequence analysis



grain morphology for 10 HIF genotypes with positive or negative 5A QTL causal variant alleles were tracked during the 2019 field season (Days post anthesis, DPA). grain width, fresh weight and dry weight were measured with 2 replicates / genotype, 10 spikes / time point and 10 kernels / spike. **Significant difference in grain morphology is measurable 10-16 DPA for 5A QTL HIFs grown in Ithaca, NY.** The difference in phenotypes for the positive and negative causal variant genotypes is likely due to a gene expression event prior to 10 DPA, providing us with a high confidence window for timing RNA extraction.

5 Previous gene expression reports for 5A QTL peak regions



High confidence gene annotation (Alaux et al., 2018) and gene expression reports from wheat-expression.com indicate that 24 out of the 113 candidate genes have **zero gene expression** in the kernels prior to 15 DPA.

36 gene expression studies
118 grain tissue samples
65 different varieties of wheat
Grey text: 0.0 tpm before 15 DPA

Borrill et al., 2016; Ramírez-González et al., 2019
Figure code credit Dr. Shantel A. Martinez

6 Next steps: RNA-seq

Relying on large populations over many generations to detect crossovers and capture finer resolution of the QTL is resource limiting. Our attention has now turned to gene expression to facilitate characterization of the underlying causal grain weight and morphology gene, in a greenhouse environment.

RNA-seq experimental design

HIF grain tissue sample **4 & 8 DPA**	334.0 Mbp	337.8	339.7	341.5	349.6	367.5	374.8	378.5	380.8	383.8	385.3	404.5
7-956-2-19-1-31-3	o	o	o	o	o	o	o	o	o	o	o	o
7-956-2-19-1-44	w	w	w	w	w	w	w	w	w	w	w	w
7-956-2-19-1-31-5	o	o	o	o	o	o	o	w	w	w	w	w
7-956-2-12-1-69-07	w	w	w	o	w	w	w	o	o	o	o	w

40 spikes for the 4 DPA timepoint and randomly sampled sets of 10 spikes across 4 biological reps. In addition, 8 spikes were tagged for the 8 DPA timepoint randomly sampled sets of 2 spikes across 4 biological reps (fewer needed based on increased grain size). The RNA was extracted using a modified hot borate method, and 3/4 biological reps were sent to Novogene for extraction (24 samples).

References

Alaux et al., 2018 doi: 10.1186/s13059-018-1491-4
Borrill et al., 2016 doi: 10.1104/pp.15.01667
Broeman et al., 2003 doi: 10.1093/bioinformatics/btg112
Jiang et al., 2019 doi: 10.1186/s12284-019-0308-8
Ramírez-González et al., 2019 doi: 10.1126/science.aar6089
Sorrells et al., 2011 doi:10.1139/G11-054
Tuinstra et al., 1997 doi: 10.1007/s001220050654
Wan & Layton, 2006 isbn: 978-1-59745-130-7

Stay in touch:
et395@cornell.edu
@etaagen
ella-taagen

This project is funded by the Wheat CAP grant from the USDA-NIFA (no. 2017-67007-25939)