N-IWYP800 NIFA CAP: LEVERAGING HIGH-THROUGHPUT GENOTYPING AND PHENOTYPING TECHNOLOGIES TO ACCELERATE WHEAT IMPROVEMENT AND MITIGATE THE IMPACTS OF CLIMATE CHANGE.

Jorge Dubcovsky

Goals / Objectives

Overall goal: The overall goals of this project are to accelerate breeding cycles to mitigate the impact of climate change on wheat productivity and to train a new generation of plant breeders.

The specific objectives of the project are to: 1) Develop cost-effective, medium-density, single nucleotide polymorphism (SNP) assays incorporating both functional and haplotype-tagging SNPs for effective genome-wide imputation and link these assays with practical haplotype graphs (PHG) to establish a centralized automated platform for genomic selection (GS) in public wheat breeding programs.2) Implement a centralized high-throughput phenotyping platform based on unmanned aerial systems (UAS-HTP). This centralized BRAPI-compliant platform will support data processing, analysis and management and is aimed at accelerating the adoption of UAS-HTP in public wheat breeding programs. Phenotypic data will be deposited into the public T3/BreedBase database.3) Develop improved breeding decision support tools based on T3/BreedBase capable of maintaining centralized upload and processing of genotypic and agronomic data from public breeding programs. Promote the exchange of data and analyses tools with equivalent public wheat breeding databases used by CIMMYT and the UK.4) Develop a publicly accessible genomic catalogue of natural and induced genetic variants that regulate gene expression in wheat, assess the role of regulatory diversity in controlling pathways underlying variation in growth and grain yield, and integrate information about regulatory variants into the T3/BreedBase database.5) Train a new cohort of 20 plant breeders within an active Community of Practice in plant breeding, by integrating them into active wheat breeding programs that combine field research, UAS-HTP, and GS to accelerate the development of improved wheat varieties. In rapidly changing environments, a species' chance of survival depends on two factors: levels of genetic variation and efficiency of selection for beneficial variants. Our project plans to use strategies affecting both factors to mitigate the impact of climate change on wheat productivity.

Increasing genetic diversity: The ability to develop varieties adapted to changing climatic conditions will strongly depend on the availability of novel climate-adaptive diversity. Several teams in this project will explore wild relatives of wheat selected from diverse geographic regions with distinct and extreme climatic conditions as a source of climate-adaptive diversity for future wheat breeding. This introgressed diversity will be systematically evaluated in multiple environments for adaptive potential by collecting image data using the UAS-based sensory technologies. In addition to natural variation, this project will also incorporate novel induced diversity generated by chemical mutagenesis and CRISPR. Mutants that increase grain size and weight (e.g. gw2) and grain number (e.g. ful2, vrt2, and svp1) identified in our previous grant will be incorporated into elite lines in the public breeding programs.

Accelerating breeding cycles: The second strategy for mitigating the negative effects of climate change on wheat productivity is based on accelerating breeding cycles. Shortening of breeding cycles allows for effective selection of superior allelic complexes better adapted to the climatic conditions. The GS pipelines proposed in this project can reduce the time from crosses to released varieties by facilitating

selection in off season nurseries or in "rapid breeding" greenhouses and by identifying valuable parental lines combining multiple favorable alleles for inclusion in crossing blocks at earlier generations.

Project Methods

Methods used in this project are presented below organized by objectives.

Methods for Objective 1. Development of a Genomic Selection (GS) pipeline.

For this objective we will first develop a cost-effective medium-throughput SNP platform incorporating functional and haplotype-tagging SNPs. The focus will be on sequence-based targeted methods, which offer low levels of missing data, sufficient marker density for genotype imputation, the ability to incorporate functional SNPs and the possibility to discover new allelic variants. The project will support a GS coordinator at T3 that will work with the four genotyping laboratories to establish a centralized GS data analysis pipeline. Genotyping laboratories will offer DNA extraction and genotyping of up to 500 lines per year per breeding program with the medium-density genotyping platform. The GS coordinator will receive and deposit the SNP calls, the UAS-HTP data, and the phenotypic data from the breeders. A Practical Haplotype Graphs (PHG) will be applied to impute a high-density, uniform marker set that will be used to obtain genomic predictions. The PHG will also be used to impute alleles for known genes affecting grain yield and biomass.

Methods for Objective 2. Development of a UAS-HTP data analyses pipeline.

Unmanned Aircraft System (UAS) will be used to obtain high-throughput phenotyping (HTP) data. We will use a pipeline already developed by the TAMU/Purdue UAS-HTP group that includes an automated data processing workflow. This pipeline uses UAS data collected over the growing season to extract plant height, canopy cover, canopy volume, and vegetation indices for wheat. The WheatCAP will provide support for a full-time person at TAMU to coordinate these efforts. Participating programs will collect raw data using their UAS and sensors and upload it to a central UAS Hub in TX. The Hub will process the raw data into geospatial Level 1 and Level 2 data and send the plot/grid level phenotypic data back to the breeder. Breeding programs will be able to submit data for 2 acres/year without charge. The Texas UAS group will store data for the duration of the project and Level 2 data will be deposited in T3. During the first year of the project the UAS coordinator at TAMU will organize an online workshop for breeders and graduate students to coordinate the UAS-HTP activities in the project. During the 2nd year, the UAS coordinator will organize a hands-on workshop for WheatCAP students to train them in UAS-HTP data collection and analysis.

Methods for Objective 3. Implementation of BreedBase in breeding programs and transfer data to T3.

T3 will transition all datasets to the BreedBase backend. This backend is currently used by applied breeding programs globally and has strong external software development support for breeding management functions that WheatCAP can leverage and contribute to. Through BreedBase, WheatCAP breeders will have access to integrated functions for experimental design, managing barcodes, mobile data capture through the Android Field Book, seed inventory, and submission and tracking of samples for DNA marker analysis. T3 will provide an online workshop for the students on BreedBase. T3 will work with the UK and CIMMYT wheat database groups to federate the three databases using BrAPI implementations that will help researchers push/pull datasets in and out of the cloud. In addition, T3 will develop decision support tools to facilitate integration of GS into the public wheat breeding

programs, including selection of model training datasets, genomic estimated breeding values (GEBVs) and identification of cross combinations with high mean and variance among progeny.

Methods for Objective 4. Regulatory diversity in the wheat genome

Regulatory diversity will be characterized in a panel of 200 lines representing global and US-regional wheat genetic diversity. All lines will be sequenced with the new wheat Regulatory Capture Assay and will be skim-sequenced at low coverage. Reads will be mapped to 15 PanGenome wheat lines, and used for variant calling and imputation using PHG. The generated variants with MAF > 0.02 will be applied for mapping. To map regulatory variants, we will perform association mapping using transcript abundance levels estimated for 10 tissues from each of the 200 wheat lines using the cost-effective Quant-Seq approach. All tissues will be collected in three biological replicates that will be combined in one RNA sample per tissue per accession. Quant-seq of the 2,000 RNA samples will be performed at the KSU Integrated Genomics Facility. Illumina single-end 150 bp reads will be quality trimmed, aligned to the wheat PanGenome, and read count data will be generated using the HTSeq-count. The association between variants and gene expression will be performed by Matrix eQTL. SNPs with the strongest association signal within an interval will be defined as an eQTL. The identified regulatory variants will be deposited into T3 and incorporated into the GS models. Both cis- and trans-eQTL will be used to infer regulatory interactions between genes in wheat gene co-expression networks. In addition, we will use GWAS and eQTL summary-level Mendelian randomization (SMR) data analysis to identify SNPs that control both gene expression and trait variation data. The identified SNPs will be incorporated into the public medium-density genotyping assay.

Methods for Objective 5

This project will provide support and integrated training in plant breeding to a minimum of 20 PhD students. WheatCAP students will be involved in the wheat breeding activities to gain hands-on experience in managing a breeding program. Students will also interact with stakeholders through participation in field days, grain grower meetings and conferences. WheatCAP students will participate in implementing the GS and UAS-HTP platforms and in integrating field trial data into BreedBase. Students will be formally trained in experimental design, marker-assisted selection, genomic selection and basic bioinformatics, all important skills for modern plant breeders. Students will receive soft skill training to learn about their strengths, improve interpersonal communication skills, team work, mentorship and how to lead people. Students will organize workshops and symposiums. In addition to online workshops, students will attend one face-to-face WheatCAP workshop each year, which will help to build a well-integrated cohort of students and to promote a culture of collaborative research. During the first half of the project, courses will be focused on training students in the different skills required for their experiments and in the use of domestic and international wheat database resources. Courses offered during the latter half of the project will focus on providing students with opportunities to network with representatives from industry to learn how to prepare for job interviews and to understand the expectations for successful careers. To familiarize students with the plant breeding industry, we will organize workshops with the private plant breeding companies' recruiting and training personnel during the last two years of this project. During the same period, the educational coordinator will work with industry to provide students with opportunities for short term internships in industry.