# LEVERAGING THE GENOMICS REVOLUTION WITH HIGH-THROUGHPUT PHENOTYPING FOR CROP IMPROVEMENT OF ABIOTIC STRESSES

by

JARED LEVI CRAIN

B.S., Oklahoma State University, 2010
M.S., Oklahoma State University, 2012

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

# Abstract

A major challenge for 21st century plant geneticists is to predict plant performance based on genetic information. This is a daunting challenge, especially when there are thousands of genes that control complex traits as well as the extreme variation that results from the environment where plants are grown. Rapid advances in technology are assisting in overcoming the obstacle of connecting the genotype to phenotype. Next generation sequencing has provided a wealth of genomic information resulting in numerous completely sequenced genomes and the ability to quickly genotype thousands of individuals.

The ability to pair the dense genotypic data with phenotypic data, the observed plant performance, will culminate in successfully predicting cultivar performance. While genomics has advanced rapidly, phenomics, the science and ability to measure plant phenotypes, has slowly progressed, resulting in an imbalance of genotypic to phenotypic data. The disproportion of high-throughput phenotyping (HTP) data is a bottleneck to many genetic and association mapping studies as well as genomic selection (GS).

To alleviate the phenomics bottleneck, an affordable and portable phenotyping platform, Phenocart, was developed and evaluated. The Phenocart was capable of taking multiple types of georeferenced measurements including normalized difference vegetation index and canopy temperature, throughout the growing season. The Phenocart performed as well as existing manual measurements while increasing the amount of data exponentially. The deluge of phenotypic data offered opportunities to evaluate lines at specific time points, as well as combining data throughout the season

to assess for genotypic differences. Finally in an effort to predict crop performance, the phenotypic data was used in GS models. The models combined molecular marker data from genotyping-by-sequencing with high-throughput phenotyping for plant phenotypic characterization. Utilizing HTP data, rather than just the often measured yield, increased the accuracy of GS models.

Achieving the goal of connecting genotype to phenotype has direct impact on plant breeding by allowing selection of higher yielding crops as well as selecting crops that are adapted to local environments. This will allow for a faster rate of improvement in crops, which is imperative to meet the growing global population demand for plant products.

LEVERAGING THE GENOMICS REVOLUTION WITH HIGH-THROUGHPUT PHENOTYPING FOR CROP IMPROVEMENT OF ABIOTIC STRESSES

by

JARED LEVI CRAIN

B.S., Oklahoma State University, 2010
M.S., Oklahoma State University, 2012

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:

Major Professor
Jesse A. Poland

# Abstract

A major challenge for 21st century plant geneticists is to predict plant performance based on genetic information. This is a daunting challenge, especially when there are thousands of genes that control complex traits as well as the extreme variation that results from the environment where plants are grown. Rapid advances in technology are assisting in overcoming the obstacle of connecting the genotype to phenotype. Next generation sequencing has provided a wealth of genomic information resulting in numerous completely sequenced genomes and the ability to quickly genotype thousands of individuals.

The ability to pair the dense genotypic data with phenotypic data, the observed plant performance, will culminate in successfully predicting cultivar performance. While genomics has advanced rapidly, phenomics, the science and ability to measure plant phenotypes, has slowly progressed, resulting in an imbalance of genotypic to phenotypic data. The disproportion of high-throughput phenotyping (HTP) data is a bottleneck to many genetic and association mapping studies as well as genomic selection (GS).

To alleviate the phenomics bottleneck, an affordable and portable phenotyping platform, Phenocart, was developed and evaluated. The Phenocart was capable of taking multiple types of georeferenced measurements including normalized difference vegetation index and canopy temperature, throughout the growing season. The Phenocart performed as well as existing manual measurements while increasing the amount of data exponentially. The deluge of phenotypic data offered opportunities to evaluate lines at specific time points, as well as combining data throughout the season

to assess for genotypic differences. Finally in an effort to predict crop performance, the phenotypic data was used in GS models. The models combined molecular marker data from genotyping-by-sequencing with high-throughput phenotyping for plant phenotypic characterization. Utilizing HTP data, rather than just the often measured yield, increased the accuracy of GS models.

Achieving the goal of connecting genotype to phenotype has direct impact on plant breeding by allowing selection of higher yielding crops as well as selecting crops that are adapted to local environments. This will allow for a faster rate of improvement in crops, which is imperative to meet the growing global population demand for plant products.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgements

Compiling a dissertation that involved interdisciplinary research in several different countries would not be possible without the support and efforts of a large number of people. Technicians, scientists, and support staff have played vital roles in making research opportunities accessible to me, as well as assisting with numerous field and laboratory activities. These individuals have also provided insight into data analysis, hardware and software design and implementation along with much needed guidance. Finally, I would like to thank my family and wife Lena for their unconditional love and support.

# Dedication

This work is dedicated to farmers who actually get work done and make life easier for the rest of us.  As some of Nobel Laureate Norman Borlaug's final words "take it to the farmer," it is my hope that this research will find its way back to the farmers through better crops.

# Preface

The world's population is expected to reach 9.7 billion by 2050 and surpass 11 billion by 2100 (United Nations, Department of Economic and Social Affairs, 2015). From 2010-2060, the global population will consume more food than has been consumed in the past 10,000 years (James, 2010). To meet this challenge it is predicted that global agricultural output must double by 2050 (Ray et al., 2013). While the task may be daunting, it is not the first time that agricultural scientists have been mustered to the challenge. For example, from *The Population Bomb*, a popular book in the 1960's:

> The battle to feed all of humanity is over. In the 1970's the world will undergo famines—hundreds of millions of people are going to starve to death in spite of any crash programs embarked upon now (Ehrlich, 1968).

Fortunately such dire predictions failed in response to a "Green Revolution" spearheaded by Nobel Laureate Norman Borlaug (Quinn, 2009). In fact India actually produced enough grains to be self sufficient during the 1970's (Kapila, 2009), during the exact decade that Ehrlich predicted mass starvation.

While many people are familiar with the work and significance of Norman Borlaug, they are often less familiar with the research and innovation that allowed Borlaug to be successful. Directly and indirectly agricultural scientists had been quietly working for years to improve the lives of their fellow man in all facets of agriculture. While a review of all the researchers and contributions to humanity would be far beyond the scope of a preface, the work of a few people should be noted. William and Charles

Saunders worked effortlessly, over twenty years, to develop a new variety of wheat called Marquis, the progeny of Hard Calcutta and Red Fife. Marquis was an early ripening wheat, which was a necessity for northern climates in Canada, but it was also had excellent baking and milling properties (de Kruif, 1928). Marquis would find success in commercial production, but Edgar McFadden would make an improbable cross with Marquis and emmer in 1916 resulting in Hope, both literally and figuratively (Ebert, 2014). Hope wheat provided rust resistance that would be instrumental for Norman Borlaug and his success.

Who generated the idea to cross emmer and wheat in the early 1900's? The answer would belong to a *verdadero* Kansas State Wildcat Alum Mark A. Carleton who suggested in 1901 emmer would be extremely important for improving commercial wheat varieties (Babcok, 1940). Carleton exerted a towering presence in the world of wheat at the turn of the 20$^{th}$ century. Single handedly, he established a durum wheat industry in the US worth over $30,000,000 in grain production in 1907, equivalent to $735,000,000 today. While he influenced durum wheat in the Dakotas, he also introduced Kharkov wheat, which had immense success in the Great Plains and Kansas (de Kruif, 1928).

The challenge to meet food production has been met time and time again by scientist with determination to do the impossible. Today's challenge to provide for the coming nine billion people will be solved with the same determination as Borlaug, McFadden, Saunders, and Carleton displayed. In the words of Borlaug (2007) "Those of us on the food front still have a big job ahead of us. So let's get on with the job."

# Chapter 1 - A match made in heaven—Genomics and Phenomics

The genomics era has long promised to revolutionize the field of plant breeding. In the 1980's, it was suggested that dense genetic maps could be made with restriction fragment length polymorphisms (RFLPs) leading to increased efficiency in manipulating agriculturally important traits (Beckmann and Soller, 1986). By the mid 2000's, several authors have reported that the promise of molecular markers has fallen short of it anticipated goal (Bernardo, 2008; Xu and Crouch, 2008). A myriad of reasons underpin the limited success of exploiting genomic markers for improved plant breeding. One of the largest factors is that reported molecular markers for a trait must be validated in a breeding population of interest, which is often different from the experimental population used to identify the marker. Once a marker is validated in a breeding population, an affordable efficient assay must be developed to screen for the maker (Xu and Crouch, 2008). A large number of markers, hundreds or thousands, should also be available which is well beyond the capacity of traditional markers like RFLPs (Luikart et al., 2003). Additionally, accurate phenotyping has been implicated as hampering genetic progress both in identifying significant marker trait associations as well as the predicting gene to phenotype associations (Campos et al., 2004; Xu and Crouch, 2008).

Technology has provided breakthroughs to the problems of the last century, as DNA sequencing methods have increased exponentially. Next generation sequencing methods, commercially available since 2004, provide massively parallel, high throughput genotyping capabilities (Mardis, 2008). Today, the cost of sequencing the human genome is rapidly closing in on $1,000 (Wetterstrand, 2016) compared to 13 years and $3 billion dollars for the first human genome sequence completed in 2003

(Lange et al., 2014). Agriculturists have been quick to leverage these new technologies for genomic studies as well as plant breeding purposes.

One of the most popular methods in plant genomics is genotyping-by-sequencing (GBS). GBS simplifies marker development and population genotyping, assaying the markers across a group of individuals, into one step. This results in rapid develop of molecular markers as well as avoiding ascertainment bias that accompanies traditional assays. GBS is highly amenable to a large variety of species, and can be adjusted to varying degrees of complexity and multiplexing of samples fitting end users needs (Poland and Rife, 2012). Along with providing robust genotyping GBS has shown to be effective in a variety of genetic studies including genomic selection (Poland et al., 2012b; Crossa et al., 2013) and association mapping (Arruda et al., 2016).

GBS has provided molecular answers to some of the doubts about marker assisted selection (MAS) and the role of genomics in plant breeding. Particularly, GBS can provide thousands to hundreds of thousands of molecular markers (Elshire et al., 2011; Poland et al., 2012a) thus saturating the genome and assuring the best opportunity for finding causal marker trait associations (Poland and Rife, 2012). While the ability to sequence DNA and develop molecular markers has advanced rapidly, phenotyping has greatly lagged behind these genomic advances (Campos et al., 2004; White et al., 2012; Cobb et al., 2013). This dissertation focuses on the areas in which plant phenotyping can be improved and how that phenotypic information can be incorporated into breeding programs, thus meeting one of agricultures biggest challenges of the 21$^{st}$ century predicting cultivar performance based on its genetic architecture (White et al., 2012). Some of the problems and panaceas investigated are:

1.) design and deployment of a flexible phenotyping platforms. 2.) extraction of meaningful information from phenotyping platforms. 3.) methods to combine both the phenomics and genomics into crop breeding programs.

## Phenotyping Platform Development

The trend of decreasing genotyping cost has not been extended to the cost of measuring plant phenotypes. Phenomics, the systematic study of plant phenotypes at the organism scale (Houle et al., 2010), has long advocated for increased screening capacity at reduced cost (White et al., 2012; Cobb et al., 2013). High-throughput phenotyping (HTP) platforms have been developed to help relieve the phenomics bottleneck to plant breeding and genetics (Busemeyer et al., 2013; Andrade-Sanchez et al., 2014). While numerous platforms have been developed in the past few years, with the exception of a push cart by White and Conley (2013), all of them have had extreme limitations namely cost over $100,000 USD (White et al., 2012). Chapter 2 investigates methods to develop high-throughput phenotyping methods that are sufficiently affordable to be used in the developing world. This will allow local breeding programs the world over to utilize HTP techniques (Tester and Langridge, 2010).

## Exploration of HTP Data

While high-throughput phenotyping platforms are instruments that provide ways to assess traits, it should be noted that HTP measurements are not the goal themselves (Granier and Vile, 2014). Additionally, collecting HTP measurements on a single day does not provide insight into the functioning of a plant trait (Granier and Vile, 2014). To examine the value of assessed phenotypes, Chapter 3 looks at how HTP measurements can be collected throughout the growing season, how to determine data quality, and statistical methods that can be used to add value to data taken at multiple times throughout the growing season. HTP measurements and their evaluation can be used to investigate biological processes resulting in grain yield and have the potential to be exploited by plant breeders.

## Combining Phenomics with Genomics

Finally, an effort to integrate the disciplines of genomics and phenomics culminating with enhanced genomic prediction models is developed in Chapter 4. While there are a myriad of ways that HTP traits may be used in practice, linkage and association mapping (Myles et al., 2009), connecting the genotype-to-phenotype (White et al., 2012), genomic selection provides a type of marker assisted selection where genetic markers are located across the entire genome and each physiological trait is associated with at least one marker (Goddard and Hayes, 2007). The ability to match dense genetic markers with high throughput phenotypic data should allow fruition of the promises of genomics to revolutionize plant breeding as well as increasing our knowledge of genotype to phenotype responses in plants.

# Chapter 2 - Development and deployment of a portable field phenotyping platform

This chapter has been published as following journal article:

Crain, J.L., Wei, Y., Barker III, J., Thompson, S.M., Alderman, P.D., Reynolds, M., Zhang, N., Poland, J., 2016. Development and deployment of a portable field phenotyping platform. Crop Sci. doi:10.2135/cropsci2015.05.0290

## Abbreviations

CT, canopy temperature; GNSS, global navigation satellite system; HTP, high-throughput phenotyping; IRT, infrared thermometer; NDVI, normalized difference vegetation index; UTM, Universal Transverse Mercator.

## Abstract

Accurate and efficient phenotyping has become the biggest hurdle for evaluating large populations in plant breeding and genetics. Contrary to genotyping, high-throughput approaches to field-based phenotyping have not been realized and fully implemented. To address this bottleneck, a novel, low-cost, flexible phenotyping platform, named Phenocart, was developed and tested on a field trial consisting of 10 historical and current elite wheat (*Triticum aestivium* L.) breeding lines at the International Maize and Wheat Improvement Center (CIMMYT). The lines were cultivated during the 2013 and 2014 growing cycle in Ciudad Obregon, Mexico, and evaluated multiple times throughout the growing season. The phenotyping platform was developed by integrating several sensors: a GreenSeeker for spectral reflectance, an infrared thermometer (IRT), and a global navigation satellite system (GNSS) receiver into one functional unit. The Phenocart enabled simultaneous collection of normalized difference vegetation index (NDVI) and canopy temperature (CT) with precise assignment of all measurements to plot location by georeferenced data points. Across the set of varieties, the Phenocart temperature measurements were highly correlated to a handheld IRT. In addition, CT and NDVI were both significantly correlated to yield throughout the growing season. The Phenocart is a flexible, low-cost, and easily deployable platform to increase the amount of phenotypic data that crop breeders obtain as well as provide high-resolution phenotypic data for genetic discovery.

## Introduction

Accurately assessing and recording phenotypic data is essential for plant breeders and geneticists. With the rapid advances of high-throughput genotyping technology—driven in part by the quest to sequence a human genome for less than US$1000—scientists have tremendous opportunity to generate high-density genomic data for crop improvement (Morrell et al., 2011). However, successfully leveraging genotypic data requires large amounts of phenotypic data to connect genotype to phenotype and understand plant genomes (Campos et al., 2004; Cobb et al., 2013). While great progress has been made in decreasing cost and increasing throughput for genotyping and sequencing, phenotyping has lagged behind (Houle et al., 2010; Araus and Cairns, 2014).

Phenotyping is often described as the current bottleneck to applied breeding, association and linkage mapping studies (Myles et al., 2009), development of genomic selection models (Houle et al., 2010; White et al., 2012; Cobb et al., 2013), and the connection of genotype to phenotype (White et al., 2012). Consensus among researchers has indicated that the cost and time associated with phenotyping needs to be decreased (Furbank and Tester, 2011; Fiorani and Schurr, 2013; Dhondt et al., 2013) while realizing concurrent increases in precision and accuracy (Cobb et al., 2013). An increase in phenotyping capacity could lead to better understanding of the genetic basis of complex traits such as yield. Additionally, phenomics, the collecting of high-dimensional phenotypic data (Houle et al., 2010), could play a critical role in rapid adaptation of staple food crops to the stressed environments predicted for the future (Cabrera-Bosquet et al., 2012).

Plant phenotyping is often conducted in controlled or field environments. At the most stringent levels for whole plants, greenhouses and growth chambers are used, allowing for experiments with precisely defined environmental variables including light intensity and temperature. However, controlled environments have limitations for field applicability; by controlling the growth conditions, plants are not subjected to the variable intensity of environmental forces like wind, evaporation rates, soil water regimes, and many other diurnal or daily environmental fluctuations. In addition, space constraints of controlled chambers often limit the size of the populations that can be evaluated. For example, Saint Pierre and Crossa (2012) found a significant response to water use efficiency to wheat screened in greenhouse experiments, but subsequent field trials did not show the advantage of the greenhouse-selected lines under water-limited conditions. While it should be noted that the selected lines from this experiment were equivalent to the controls, this case highlights the need for field-based evaluation and phenotyping. With limited transfer of results from controlled experiments to large-scale field environments, field-based research should be preferred in assessing complex traits like drought tolerance and yield potential (Campos et al., 2004; White et al., 2012). Field-based phenotyping provides the ability to assess plants in real-world conditions and with population sizes consistent with those used in breeding programs and quantitative genetic studies (Yu et al., 2008; White et al., 2012).

Several high-throughput phenotyping (HTP) platforms, capable of generating large quantities of data quickly, have been reported at both the controlled (greenhouse) and field-based levels. One greenhouse-based system, the GROWSCREEN-Rhizo, captures images of plant roots and shoots allowing breeders to select for desired belowground traits (Nagel et al., 2012). A different greenhouse platform, Phenoscope, consists of a robotic system that can move plants, adjust watering, and take images as well as process the images automatically. The Phenoscope has been shown to reduce environmental variability and resulted in mapping power that was limited by genetic complexity rather than inefficient phenotyping (Tisné et al., 2013).

Several field-based platforms have also been described. A tractor-mounted platform using light curtains and spectral reflectance was used for nondestructive, high-throughput phenotyping of maize biomass (Montes et al., 2011). A similar type system named BreedVision was used for nondestructive biomass measurements of triticale (´Triticosecale spp.) using a phenotyping platform pulled by a tractor (Busemeyer et al., 2013a). Andrade-Sanchez et al. (2014) used a platform with four rows of sensors to measure parameters including plant height, spectral reflectance, and CT simultaneously. White et al. (2012) and Deery et al. (2014) reviewed many advantages and disadvantages of current and potential field-based systems. Some systems included tractor-mounted platforms, systems that operate on cranes or mobile vehicles (Haberland et al., 2010; White and Bostelman, 2011), towers similar to sports stadium cameras (Albus et al., 1993; White and Bostelman, 2011), aerial vehicles (Zarco-Tejada et al., 2009; Merz and Chapman, 2011), and push carts (White and Conley, 2013). While all these platforms have a potential for use, there are also limitations associated with each phenotyping scheme. Tractor-mounted vehicles require an experienced operator and could be limited by maneuverability in field. Cranes and cable robots are limited in the amount of area that they can cover. Aerial vehicles are usually weight limited. In addition to these physical limitations, the cost for these platforms is often significant. Equipment costs estimated by White et al. (2012) can be well above US$100,000 for a tractor-mounted phenotyping platform. While the pushcart presented by White and Conley (2013) is affordable, it has the drawback of being less portable between fields.

One advantage of any phenotyping platform—field or glasshouse—is the reliability of automated data collection. Recording data directly from sensors minimizes human error from the system by reducing the number of times that data is manually entered or transcribed from data collection to completed analysis. In addition, each step may be performed by different individuals, further increasing the possibility of introducing errors that are included in the analysis and biasing the results (Taylor, 1987).

While each phenotyping platform has advantages, currently there is a need for a highly mobile, field-based phenotyping platform that could be deployed in locations throughout the world at an affordable cost. This will also enable those working in developing countries or remote field locations to capitalize on the developments in field-based HTP (Tester and Langridge, 2010). To address this scope of phenotyping, we have developed a mobile platform, aptly named Phenocart. The Phenocart integrates spectral reflectance (NDVI), CT, RGB images, and georeferenced (GNSS) data collection. We evaluated the Phenocart across 2 yr of replicated trials consisting of 10 elite breeding lines of wheat and demonstrated the use of this platform for rapid assessment of accurate plant phenotypes.

## Materials and Methods

### Design of Phenocart

The Phenocart integrated multiple pieces of hardware into one functional unit (Fig. 1). At a basic level, the Phenocart is comprised of an IRT sensor to measure CT, an NDVI sensor for spectral reflectance measurements, and a high-precision GNSS (circular error probability, 95% = 10 cm using OmniStar G2 Service) unit to georeference all data. Normalized difference vegetation index and CT were chosen as parameters to measure because of their documented relationship to yield (Amani, 1996; Babar and Reynolds, 2006; Gutierrez et al., 2010). Georeferencing gave each sensor measurement a precise location and time, which prevents incorrect assignment of data and a uniqueness of each data point. In 2014, we made an additional modification by adding a web camera to record color images that were likewise georeferenced. The sensors were connected to a laptop computer serving as the main control unit for the system. The model, technical specifications, and approximate cost for all components are listed in Supplemental Table S1.

The system was operated through a custom-designed LabView 2012 program (National Instruments). The software had two main functions: (i) to log and sync GNSS and sensor data and (ii) to process these data into a useable file format. A separate module was built in LabView for each sensor because the connection parameters and output data for each device differ. Each new set of data was written to a new line in a text file at a rate of 10 Hz, with each line having a unique index number along with sensor observation data. A block diagram of the Phenocart software system and graphical user interface is provided in Supplemental Fig. 1 and 2. The Phenocart software program is provided for download at http://www.wheatgenetics.org/phenocart along with a detailed user manual.

### Phenocart

Initially, the Phenocart was used as a handheld system during the 2013 crop year; however, one of the immediate drawbacks was the overall system weight (~12 kg). While various carts and bicycles have been shown to carry sensors (White and Conley, 2013; Kelly et al., 2015), our design criteria necessitated the cart having a very narrow footprint to accommodate varying planting systems and widths. Most importantly, our cart required high maneuverability to turn in alleys and borders that could be as short as 50 cm between plots, a criteria that had not been addressed by currently proposed carts (White and Conley, 2013). To address this issue, we developed a cart constructed from a bicycle, metal, and miscellaneous hardware. The pedals were removed from the bicycle and then the frame was cut in half. A prebent pipe (walker frame) was attached to the bicycle frame for handlebars, resulting in a hybrid wheelbarrow cart. The seat connection was used to hold a pipe that could be adjusted for height with the end of the pipe threaded to a mounting bracket that held the Phenocart sensors with hose clamps. The entire system was highly flexible, as sensor height, orientation, and handlebars could be adjusted to a variety of different crops and users. The construction and completed Phenocart are outlined in Fig. 2.

## Field Design

A field trial of 10 historical and current elite wheat breeding lines was used to evaluate the Phenocart. The trial was conducted at the Campo Experimental Norman E. Borlaug CIMMYT research center near Ciudad Obregon, Sonora, Mexico. The 10 genotypes were planted in 10 by 2.4 m plots in a three-replicate, *a*-lattice design using raised beds with two rows per bed and an interbed spacing of 0.8 m. The trial was planted on 27 Feb. 2013 and 22 Nov. 2013 for each year (2013 heat trial with full irrigation and 2014 irrigated yield trial, respectively). Irrigation and nutrient levels were maintained at optimal levels with pesticides applied as needed. Grain yield was obtained by using a plot combine. Five days of observations were recorded in 2013; however, in 2014, unexpected rain led to lodging in the plots, and to prevent any additional damage to the plots, only one data set was obtained.

## Phenocart Data Acquisition

In the field, each instrument (GNSS, IRT, and camera) was mounted to the handheld GreenSeeker using a custom fabricated mounting bracket. The bracket enclosed the head of the GreenSeeker sensor without blocking either the light source or the collector (Fig. 3). All sensors were mounted so that they had a nadir field of view, and the GreenSeeker (base of the Phenocart) was held ~80 cm above the crop canopy. Data were collected at a walking speed of ~1 to 2 m $s^{-1}$. Sensor measurements were taken for the entire experiment area, including plot borders and alleys, and assigned observations to plots as described in the Data Analysis section.

Phenocart IRT (Micro-Epsilon CT Series CT-SF02-C1) readings and handheld IRT measurements were taken over the same plots to validate the Phenocart IRT with standard CT methods because of the difference in angle of view of the instruments. For these initial tests, the Phenocart sensor was maintained at a nadir view because of potential problems in georeferencing an oblique angle. The Phenocart IRT had a view angle of 3.8° and, based on the height, had a viewing area of approximately 50 $cm^2$. A Sixth Sense LT300 IRT (Instrumart) was used to collect the handheld IRT data. For temperature measurements, the handheld IRT was positioned at a constant height and angle (~30° from horizontal) over the canopy following methods outlined by Pask et al. (2012). The handheld IRT had a viewing angle of 2.9° and a measurement area of ~12 $cm^2$ (80 cm from the crop canopy). While field of view was not the same for each instrument, we investigated sensor measurements to compare standard techniques (outlined by Pask et al. [2012] and often used by CIMMYT and other research institutes) to potential HTP techniques. Data were only taken on days with cloud-free skies and limited wind (<2 km $h^{-1}$). Measurements for each plot were taken with both instruments in direct succession by passing the Phenocart over the plot followed by the handheld IRT. The plots contained two beds, where a measurement was taken directly over each bed. While the plot coordinates and georeferenced data were developed from the entire experiment, for the IRT handheld comparison, data was only collected over areas directly measured by the handheld IRT to assure that comparisons were as consistent as possible. Comparisons between the handheld IRT and the Phenocart were made at five different intervals during the vegetative and grain-filling growth stages (Feekes 4 and 10.5, respectively [Large, 1954]). For validation, each plot was considered as a unique experimental unit.

In 2014, addition of a web camera to the system provided color images of the plots. The Phenocart recorded three photos per second, and a basic exploratory analysis was conducted with the photos. Using a custom macro in the Fiji platform (Schindelin et al., 2012), the image was adjusted for hue by selecting green color (HUE between 55 and 110). The image was then converted to a binary via green color threshold and used to calculate a percentage area of the photo corresponding to green pixels. This value was used for subsequent analysis.

### Data Analysis

Data analysis was conducted using R software (R Development Core Team, 2014) and the plyr (data manipulation), ggplot2 (graphics) packages (Wickham, 2009, 2011). The rgdal package (Bivand et al., 2013) was used to convert the longitude and latitude measurements to the Universal Transverse Mercator (UTM) coordinates. Pearson correlation was determined between variables that were collected including NDVI, CT, and grain yield. The lme4 (Bates et al., 2015) package was used to fit a mixed linear model to determine variance components and calculate broad-sense heritability ($H^2$), which is also referred to as repeatability (Piepho and Möhring, 2007). The mixed model was fit by first calculating plot-level averages using all measurements from a given plot and then using these plot-level averages for the response variable in the mixed model. Heritability was calculated on a line-mean basis using the formula from Holland et al. (2003):

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}} \quad [1]$$

where $\sigma_g^2$ is genotypic variance, $\sigma_e^2$ is error variance, and $r$ is replications ($r$ = 3).

To assign observations to plots, an algorithm using NDVI was developed to find and georeference plot boundaries. Plot boundaries were determined from a data set of the trial before senescence when NDVI over the experimental plots was high (NDVI over plots >0.5). The alleyways with bare soil had low NDVI (NDVI < 0.2), which could be used to differentiate between adjacent plots. Within the data set, the low-value border NDVI observations were removed by filtering. This resulted in the index values being noncontinuous. The resulting gap in index values was used to determine plot boundaries and assign sequential plots to georeferenced coordinates. The plot boundary was extrapolated by taking the first 20 measurements at the start and end of each new plot and finding the average coordinates for each respective end of the plot. This resulted in removal of the plot ends to avoid possible border affect. Using the two average coordinates for the start and end of the plot, a rectangle (polygon) could be constructed by adding the fixed dimensions of the plot in meters to the UTM coordinates. Figure 4 demonstrates graphically the successive steps in finding and assigning plot boundaries. Once the plot boundaries were determined, all data for a given trial was assigned to a plot if the data fell within the polygon created by the four extrapolated corners (Supplemental Fig. S3). Data were assigned using the sqldf package (Grothendieck, 2012). Current data analysis programs are provided at www.wheatgenetics.org/phenocart.

# Results and Discussion

## Phenotyping Platform

We constructed and deployed a small and relatively inexpensive HTP platform called Phenocart. This platform integrates multiple sensors and high-precision GNSS to precisely tag each sensor measurement within a given field plot. An executable software package and detailed users manual are included as supplemental materials. The Phenocart records NDVI, CT, and GNSS location at 10 Hz and color photos at 3 Hz. By measuring both beds in 10-m plots, at an average walking speed, the Phenocart generated ~100 data points for NDVI, CT, and GNSS location per plot after removing the ends from the plot to eliminate potential border affects. In contrast, the handheld device only generated two data observations per plot. Two of the unique advantages of the Phenocart are the timesaving in data collection and the type of data recorded.

The time needed per plot for data collection and processing with the Phenocart was less than the handheld IRT. All Phenocart data collected were written to text files and were immediately ready for data analysis. Handheld IRT readings required additional time to manually record the values and then had to be transcribed into digital files, essentially taking many times longer than the Phenocart. For a typical research program (CIMMYT) to collect this data efficiently, they often employ groups of two people: one person to write the data in the field while a second person takes measurements. In the laboratory, another person enters and checks the data, and based on experience, this takes approximately the same amount of time as field data collection. Using the HTP platform, over three times the number of plots could be covered in the same person hours. In addition to being faster than current handheld IRT measurements, NDVI was also obtained simultaneously, further increasing productivity as manually collected NDVI requires similar commitment as CT measurements. Again, these measurements are often conducted with teams of two people in the field: one to operate sensor and one to note plot locations followed by time curating measurements. This timesaving provides an efficiency boost to research programs in terms of data processing but also by increasing the quantity of data that can be obtained especially when trying to collect dynamic data throughout the growing season.

Along with the time savings, the Phenocart also georeferenced all sensor observations. Using the guidelines in Pask et al. (2012), only the average temperature was recorded for the handheld CT. The Phenocart records all values so that within-plot variation can be determined and included in statistical models, providing a richer and more informative data set.

## Data Processing

Implementation of the Phenocart for phenotyping greatly increased the amount of data that was collected per plot. To effectively handle the quantity of data, we developed an algorithm to efficiently process the data. The first part of the algorithm converted all longitude and latitude to UTM coordinates. Plots were then identified by taking advantage of the bare soil between plots, which resulted in low NDVI values. A rectangle polygon, fit smaller than the actual plot dimensions to avoid border affect, was inscribed onto the data. Subsequent evaluation of this method with 24 random plots of differing sizes resulted in assigning the correct coordinates within 13 cm of the extrapolated plot corner based on NDVI to physically measured coordinates of the plot, which is within the observed accuracy of the GNSS. Coordinates of the plot boundaries were recorded, and any subsequent data located within that experiment was processed by UTM conversion and assigning data observations to the correct plot. This algorithm facilitated fast and accurate data processing, and allowed data analysis and summaries to be calculated immediately following data collection. Rapid assessment of data should enable immediate identification of any abnormalities that may occur and taking corrective actions as needed.

### Validation to Handheld IRT

To validate the readings of the Phenocart to current instruments, we took concurrent measurements with both the Phenocart IRT and a handheld IRT across five time points throughout the vegetative and grain-fill growth periods. We observed a high correlation ($p$ < 0.01) on 4 of the 5 d between instruments (Table 1). These data suggest that the two instruments are consistently measuring the same canopy dynamics, and that the Phenocart provides reliable data that is similar to current methods that are used in many research programs like CIMMYT. This finding is further highlighted by the fact that the measurement area and field of view were different between the two instruments. The Phenocart IRT was maintained at a nadir view because an oblique view would lead to georeferencing problems based on sensor orientation and canopy height. However, it still detected similar differences as the handheld IRT. This suggests that using a nadir field of view could be used in place of a horizontal view, simplifying future HTP data collection.

### Canopy Temperature and Yield

We collected CT data from five different days to evaluate the correlation between CT and grain yield. During the grain-filling period, CT measured with the Phenocart was significantly correlated ($p < 0.05$) to grain yield (Table 1). Correlation between the handheld and grain yield was also calculated, with two of the five measurements having significant ($p < 0.05$) negative correlation with grain yield. As might be expected, the correlations between the two instruments were much more consistent than correlations between yield and the sensor measurements (Fig. 5). Over the five measurement days, correlation coefficient for the Phenocart IRT and grain yield had a large range (−0.02 to −0.55); however, the handheld IRT had a larger range of correlation (−0.13 to −0.73). While this is significant variation, it is consistent with the range observed in previous studies such as Keener and Kircher (1983) ($r^2$ = 0.2–0.79) and Balota et al. (2007) ($r^2$ = 0.05–0.67). Based on these data, the Phenocart IRT is performing as well as current CT measurement techniques.

## NDVI and Grain Yield

Normalized difference vegetation index formed one of the core components of the Phenocart because spectral reflectance indices have been proposed as a selection tool for plant breeding (Prasad et al., 2007). We collected NDVI simultaneously with CT for each date of data observations. The NDVI was significantly correlated to grain yield ($p < 0.001$) during the grain-fill stages (Table 2). The correlation was not as strong during the vegetative growth stages consistent with previous findings that NDVI has higher correlation to yield later in the growing season (Babar and Reynolds, 2006). Given the strong relationship between NDVI and grain yield, the ability to rapidly measure this spectral index with Phenocart will enable more informed selection decisions.

## Heritability of Canopy Temperature and NDVI

Broad-sense heritability for each sensor, NDVI, CT, and color photos was calculated on a line-mean basis. For the same plots, measured at the same time with two different instruments, comparison of broad-sense heritability estimates are a reflection of measurement error, as environmental variance should be negligible and the genetic entries are identical. For comparison of sensors, a higher heritability estimate reflects reduced measurement error. For CT data, the Phenocart had a higher heritability estimate than the handheld IRT for three of the five measurements (Table 3). While heritability varied considerably, it is again within the ranges reported by Andrade-Sanchez et al. (2014) ($H^2$ = 0.01–0.9 for CT measured at multiple time points). Heritability calculated for NDVI was high ($H^2$ = 0.66–0.94; Table 3), which is also consistent with results from Andrade-Sanchez et al. (2014). The heritability of the NDVI and CT indicates that the Phenocart can be used to generate data that is extremely consistent between plots and measurements.

## Potential for Color Images

In 2014, we mounted a web camera on the Phenocart to collect color photos of the plots. We investigated the feasibility of the photos to assess variation within and between plots by calculating heritability (Piepho and Möhring, 2007) of the number of green pixels per plot. The heritability was 0.45 (Table 3), which indicates that the photos could be used by a breeding program as an additional layer of information including the spatial uniformity and ground cover of the plots. Apart from an assessment of green leaf area (for which RGB imaging is not well suited), the photos themselves could be used for more sophisticated data analysis such as plant architecture characteristics (Paulus et al., 2014). Image acquisition could even provide breeders with the opportunity to investigate abnormalities in data after the growing season. By having opportunity to revisit a plot, a breeder could determine if other data readings were correct or in error. As an expandable platform, the addition of NDVI, infrared, or multispectral cameras to the Phenocart opens new doors of possibilities for data analysis and interpretation in a breeding program.

## Recommendations and Future of Phenocart Phenotyping Platforms

We designed the Phenocart to provide dense phenotypic data for genetic analysis and plant breeding. Phenocart is a relatively inexpensive platform that can easily be deployed to field locations around the world. Some of the unique advantages in the Phenocart are that it simultaneously collects multiple measurements including NDVI, CT, and images; additionally, the Phenocart georeferences each measurement with location and time stamp. Georeferencing will lead to less error in data transcription and thus higher data integrity. The addition of image collection to the Phenocart opens an entirely new avenue for high-throughput data collection that could add value to a breeding program. This also highlights the flexibility of the underlying software platform of the Phenocart, which allows for modification and addition of new sensors as needed.

As a first-generation system, the Phenocart platform is considerably less expensive than other phenotyping systems. The current version of Phenocart was ~US$12,000. while high-clearance vehicle phenotyping platforms have been estimated at ³US$100,000 (White et al., 2012). Currently, the complied software provided does not require that the end user purchase the entire LabView software; however, any new development would require a license, and future developments will focus on open-source versions. A large proportion of the expense was related to obtaining high-precision differential correction GNSS (US$6,400) with centimeter-level resolution. Depending on plot size, this level of precision may not be needed for all users. We have used this system in genetic mapping populations that had plot size of 2 by 0.8 m, but in larger yield trials (3.5 by 1.5 m), less accurate and less expensive GNSS units may suffice. In the same manner, other more affordable sensors may also be used in the future to reduce total system cost.

In addition to lower cost georeferencing, development is underway to run the system on Linux on small single-board computers, further reducing cost and weight while still providing sufficient robustness to incorporate an ever expanding array of sensors. A high-level scripting language, Python, is also being used to control sensors and record data. The move to Python will allow users with basic technical abilities to modify the system to fit their own needs. These modifications will reduce the cost of the platform, enabling rapid dissemination and utilization of many units in the breeding programs. Such a dense network of phenotypic data, coupled with genotypic data, will help increase the efficiency of crop breeding.

Along with the software, we have tested various hardware, sensors, and mobility options. While our test in wheat have provided evidence the system is reliable, as users move to other crops, for example maize (Zea mays L.), the users will be well served by conducting their own validations. For example, sensor movement from cart sway or vibrations may be exacerbated if the sensors are mounted 3 m high as would be necessary for maize. Thus basic data quality checks as noted by our work are needed to ensure that HTP methods are providing data as reliable or more reliable than established manual methods. While we have demonstrated specific sensors, our goal is to provide a system that will be flexible enough that users can add their own desired sensor and modify the cart as necessary to meet a wide range of crops, environments, and management practices.

The Phenocart provides the ability to take measurements several times throughout the growing season, which will also allow scientists to better understand dynamic traits. For example, Busemeyer et al. (2013b) integrated HTP data with genomic data to identify quantitative trait loci controlling biomass accumulation in triticale (*´Triticosecale* spp.) that were dynamically expressed at different time points over the growing season. Besides grain yield or biomass accumulation, the suite of traits that could be recorded by multiple measurements is tremendous; for example, Liebisch et al. (2015) recorded early-season plant vigor, senescence, and canopy cover using multiple NDVI image measurements throughout the growing season in maize. Along with crop growth traits, these tools can also be used to determine effect of disease pressure experienced by plants (S.K. Reddy, personal communication, 2015). By using an HTP platform like the Phenocart, existing genomic data can be combined with new and novel phenotypic data to not only increase the efficiency of selection but also increase our understanding of plant processes that are quantitative in nature.

## Conclusions

Phenotyping has always been a cornerstone to plant breeding; however, relative to rapidly developing genomics technologies and tools, the ability to phenotype plants has become an increasingly limiting factor. Developed as one approach for addressing this, the Phenocart integrates two sensors commonly used in plant phenotyping as well as georeferencing, which allows for faster data collection in the field and streamlined data processing. The reduced risk of incorrectly entering information provides better data integrity for researchers. Validation of the Phenocart shows that it performs as well as current methods for CT. In addition, NDVI and CT data from the Phenocart were significantly correlated to grain yield. The Phenocart is a robust, affordable platform that can be modified to fit the user's need and help close the gap between genomics and phenomics. These advances in phenomics, including easily accessible and deployable platforms, are needed to complement the wealth of genomic data with field-based, high-throughput precision phenotyping measurements.

## Supplemental Material Available

1. LabView executable file available at http://www.wheatgenetics.org/phenocart
2. Detailed user manual available at http://www.wheatgenetics.org/phenocart.
3. Supplemental Table S1.
4. Supplemental Fig. S1–S3.

## Acknowledgements

This work was supported through the National Science Foundation–Plant Genome Research Program (IOS-1238187) and the US Agency for International Development Feed the Future Innovation Lab for Applied Wheat Genomics (Cooperative Agreement No. AID-OAA-A-13-00051). J. Crain is supported through a fellowship from the Monsanto Beachell-Borlaug International Scholars program. Josh Sharon made valuable contributions to initial design and hardware testing.

## References

Albus, J., R. Bostelman, and N. Dagalakis. 1993. The NIST Robocrane. J. Robot. Syst. 10:709–724. doi:10.1002/rob.4620100509

Amani, I. 1996. Canopy temperature depression association with yield of irrigated spring wheat cultivars in a hot climate. J. Agron. Crop Sci. 129:119–129. doi:10.1111/j.1439-037X.1996.tb00454.x

Andrade-Sanchez, P., M.A. Gore, J.T. Heun, K.R. Thorp, A.E. Carmo-Silva, A.N. French, M.E. Salvucci, J.W. White, and E.A. Carmo-Silva. 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Funct. Plant Biol. 41:68–79. doi:10.1071/FP13126

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: The new crop breeding frontier. Trends Plant Sci. 19:52–61. doi:10.1016/j.tplants.2013.09.008

Babar, M., and M. Reynolds. 2006. Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. Crop Sci. 46:1046–1057. doi:10.2135/cropsci2005.0211

Balota, M., W. Payne, S. Evett, and M. Lazar. 2007. Canopy temperature depression sampling to assess grain yield and genotypic differentiation in winter wheat. Crop Sci. 47:1518–1529. doi:10.2135/cropsci2006.06.0383

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Statistical Software, 67:1–48. doi:10.18637/jss.v067.i01

Bivand, R., T. Keitt, and B. Rowlingson. 2013. rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.8-10. http://CRAN.R-project.org/package=aqp

Busemeyer, L., D. Mentrup, K. Möller, E. Wunder, K. Alheit, V. Hahn, H.P. Maurer, J.C. Reif, T. Würschum, J. Müller, F. Rahe, and A. Ruckelshausen. 2013a. BreedVision: A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors (Basel) 13:2830–2847. doi:10.3390/s130302830

Busemeyer, L., A. Ruckelshausen, K. Möller, A.E. Melchinger, K. V Alheit, H.P. Maurer, V. Hahn, E.A. Weissmann, J.C. Reif, and T. Würschum. 2013b. Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. Sci. Rep. 3:2442. doi:10.1038/srep02442

Cabrera-Bosquet, L., J. Crossa, J. Von Zitzewitz, M.D. Serret, and J.L. Araus. 2012. High-throughput phenotyping and genomic selection: The frontiers of crop breeding converge. J. Integr. Plant Biol. 54:312–320. doi:10.1111/j.1744-7909.2012.01116.x

Campos, H., M. Cooper, J.E. Habben, G.O. Edmeades, and J.R. Schussler. 2004. Improving drought tolerance in maize: A view from industry. F. Crop. Res. 90:19–34. doi:10.1016/j.fcr.2004.07.003

Cobb, J.N., G. Declerck, A. Greenberg, R. Clark, and S. McCouch. 2013. Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126:867–887. doi:10.1007/s00122-013-2066-0

Deery, D., J. Jimenez-Berni, H. Jones, X. Sirault, and R. Furbank. 2014. Proximal remote sensing buggies and potential applications for field-based phenotyping. Agronomy 5:349–379. doi:10.3390/agronomy4030349

Dhondt, S., N. Wuyts, and D. Inzé. 2013. Cell to whole-plant phenotyping: The best is yet to come. Trends Plant Sci. 18:428–439. doi:10.1016/j.tplants.2013.04.008

Fiorani, F., and U. Schurr. 2013. Future scenarios for plant phenotyping. Annu. Rev. Plant Biol. 64:267–291. doi:10.1146/annurev-arplant-050312-120137

Furbank, R., and M. Tester. 2011. Phenomics: Technologies to relieve the phenotyping bottleneck. Trends Plant Sci. 16:635–644. doi:10.1016/j.tplants.2011.09.005

Grothendieck, G. 2012. sqldf: Perform SQL selects on R data frames. R package version 0.4-6.4. https://cran.r-project.org/web/packages/sqldf/index.html

Gutierrez, M., M.P. Reynolds, and A.R. Klatt. 2010. Association of water spectral indices with plant and soil water relations in contrasting wheat genotypes. J. Exp. Bot. 61:3291–3303. doi:10.1093/jxb/erq156

Haberland, J.A., P.D. Colaizzi, M.A. Kostrzewski, P.M. Waller, C.Y. Choi, F.E. Eaton, E.M. Barnes, and T.R. Clarke. 2010. AgIIS, Agricultural Irrigation Imaging System. Appl. Eng. Agric. 26:247–253. doi:10.13031/2013.34847

Holland, J., W. Nyquist, and C. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 22:9–112.

Houle, D., D.R. Govindaraju, and S. Omholt. 2010. Phenomics: The next challenge. Nat. Rev. Genet. 11:855–866. doi:10.1038/nrg2897

Keener, M.E., and P.L. Kircher. 1983. The use of canopy temperature as an indicator of drought stress in humid regions. Agric. Meteorol. 28:339–349. doi:10.1016/0002-1571(83)90010-9

Kelly, J., J.L. Crain, and W.R. Raun. 2015. By-plant prediction of corn (Zea mays L.) grain yield using height and stalk diameter. Commun. Soil Sci. Plant Anal. 46:564–575.

Large, E.C. 1954. Growth stages in cereals illustration of the Feekes scale. Plant Pathol. 3:128–129. doi:10.1111/j.1365-3059.1954.tb00716.x

Liebisch, F., N. Kirchgessner, D. Schneider, A. Walter, and A. Hund. 2015. Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. Plant Methods 11. doi:10.1186/s13007-015-0048-8

Merz, T., and S. Chapman. 2011. Autonomous unmanned helicopter system for remote sensing missions in unknown environments. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVIII:143–148. doi:10.5194/isprsarchives-XXXVIII-1-C22-143-2011

Montes, J.M., F. Technow, and B.S. Dhillon. 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. F. Crop. Res. 121:268–273. doi:10.1016/j.fcr.2010.12.017

Morrell, P.L., E.S. Buckler, and J. Ross-Ibarra. 2011. Crop genomics: Advances and applications. Nat. Rev. Genet. 13:85–96.

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich, and E.S. Buckler. 2009. Association mapping: Critical considerations shift from genotyping to experimental design. Plant Cell 21:2194–2202. doi:10.1105/tpc.109.068437

Nagel, K.A., A. Putz, F. Gilmer, K. Heinz, A. Fischbach, J. Pfeifer, M. Faget, S. Blossfeld, M. Ernst, C. Dimaki, B. Kastenholz, A.K. Kleinert, A. Galinski, H. Scharr, F. Fiorani, and U. Schurr. 2012. GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. Funct. Plant Biol. 39:891. doi:10.1071/FP12023

Pask, A., J. Pietragalla, D. Mullan, and M. Reynolds, editors. 2012. Physiological breeding II a field guide to wheat phenotyping. CIMMYT, Mexico City, Mexico.

Paulus, S., J. Behmann, A.K. Mahlein, L. Plümer, and H. Kuhlmann. 2014. Low-cost 3D systems: Suitable tools for plant phenotyping. Sensors (Basel) 14:3001–3018. doi:10.3390/s140203001

Piepho, H.P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. doi:10.1534/genetics.107.074229

Prasad, B., B. Carver, and M. Stone. 2007. Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under Great Plains conditions. Crop Sci. 47:1426–1440. doi:10.2135/cropsci2006.07.0492

R Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Saint Pierre, C., and J. Crossa. 2012. Phenotyping transgenic wheat for drought resistance. J. Exp. Bot. 63:1799–1808. doi:10.1093/jxb/err385

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. 2012. Fiji: An open-source platform for biological-image analysis. Nat. Methods 9:676–682. doi:10.1038/nmeth.2019

Taylor, J. 1987. Quality assurance of chemical measurements. CRC Press, New York.

Tester, M., and P. Langridge. 2010. Breeding technologies to increase crop production in a changing world. Science 327:818–822. doi:10.1126/science.1183700

Tisné, S., Y. Serrand, L. Bach, E. Gilbault, R. Ben Ameur, H. Balasse, R. Voisin, D. Bouchez, M. Durand-Tardif, P. Guerche, G. Chareyron, J. Da Rugna, C. Camilleri, and O. Loudet. 2013. Phenoscope: An automated large-scale phenotyping platform offering high spatial homogeneity. Plant J. 74:534–544. doi:10.1111/tpj.12131

White, J., P. Andrade-Sanchez, M.A. Gore, K.F. Bronson, T.A. Coffelt, M.M. Conley, K.A. Feldmann, A.N. French, J.T. Heun, D.J. Hunsaker, M.A. Jenks, B.A. Kimball, R.L. Roth, R.J. Strand, K.R. Thorp, G.W. Wall, and G. Wang. 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133:101–112. doi:10.1016/j.fcr.2012.04.003

White, J., and R. Bostelman. 2011. Large-area overhead manipulator for access of fields. Proc. 4th Int. Multi-Conference Eng. Technol. Innov. (IMETI 2011), Orlando, FL. 19–22 July, 2011. NIST, Gaithersburg, MD.

White, J.W., and M.M. Conley. 2013. A flexible, low-cost cart for proximal sensing. Crop Sci. 53:1646–1649. doi:10.2135/cropsci2013.01.0054

Wickham, H. 2009. ggplot2: Elegant graphics for data analysis. Springer, New York.

Wickham, H. 2011. The split-apply-combine strategy for data analysis. J. Stat. Software 40:1–29.

Yu, J., J.B. Holland, M.D. McMullen, and E.S. Buckler. 2008. Genetic design and statistical power of nested association mapping in maize. Genetics 178:539–551. doi:10.1534/genetics.107.074245

Zarco-Tejada, P.J., J.J. Berni, L. Suárez, G. Sepulcre-Cantó, F. Morales, and J.R. Miller. 2009. Imaging chlorophyll fluorescence with an airborne narrow-band multispectral camera for vegetation stress detection. Remote Sens. Environ. 113:1262–1275. doi:10.1016/j.rse.2009.02.016

**Figure 2-1. Integrated hardware components of Phenocart.**

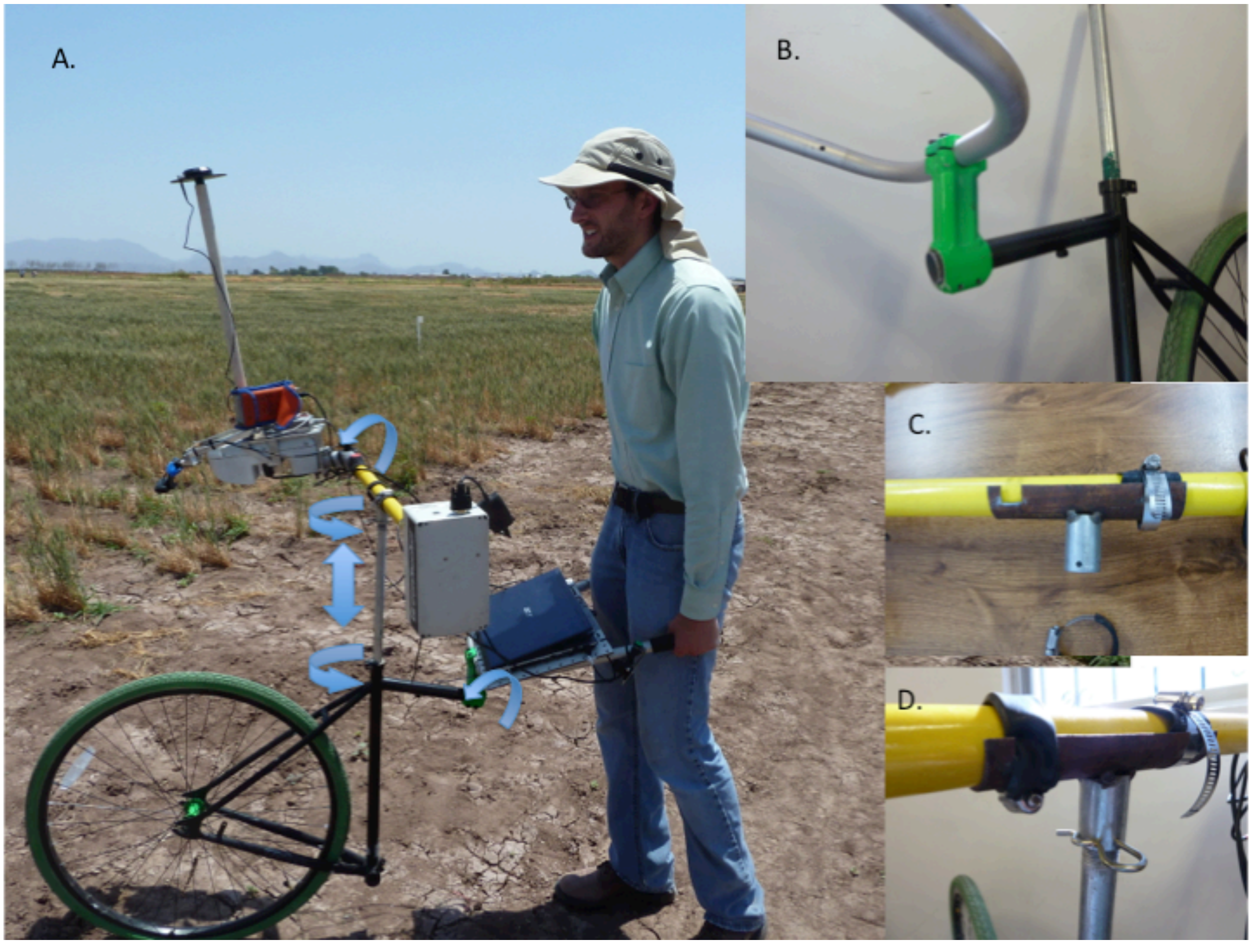


**Figure 2-2. Phenocart to carry instruments in the field. (A) Completed Phenocart with arrows showing adjustable areas of the cart. (B) Attaching a handle to the cart. (C) Bracket used to hold the GreenSeeker sensor in combination with hose clamps. (D) Bracket and GreenSeeker attached to the cart.**

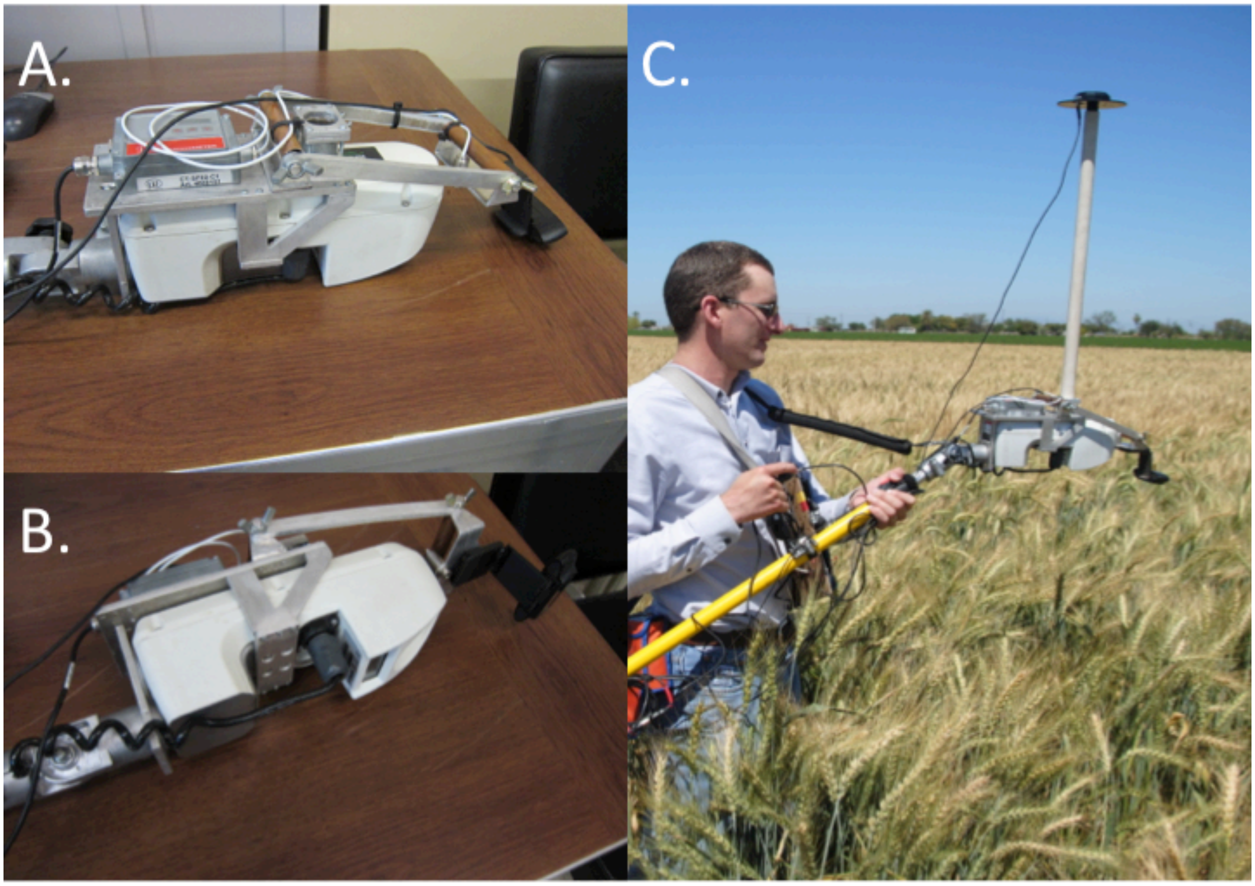


**Figure 2-3. Combined sensor design of the Phenocart. (A, B) Custom mounting bracket to hold individual sensors. (C) Using the Phenocart as a handheld platform in the field, Ciudad Obregon, Sonora, Mexico.**

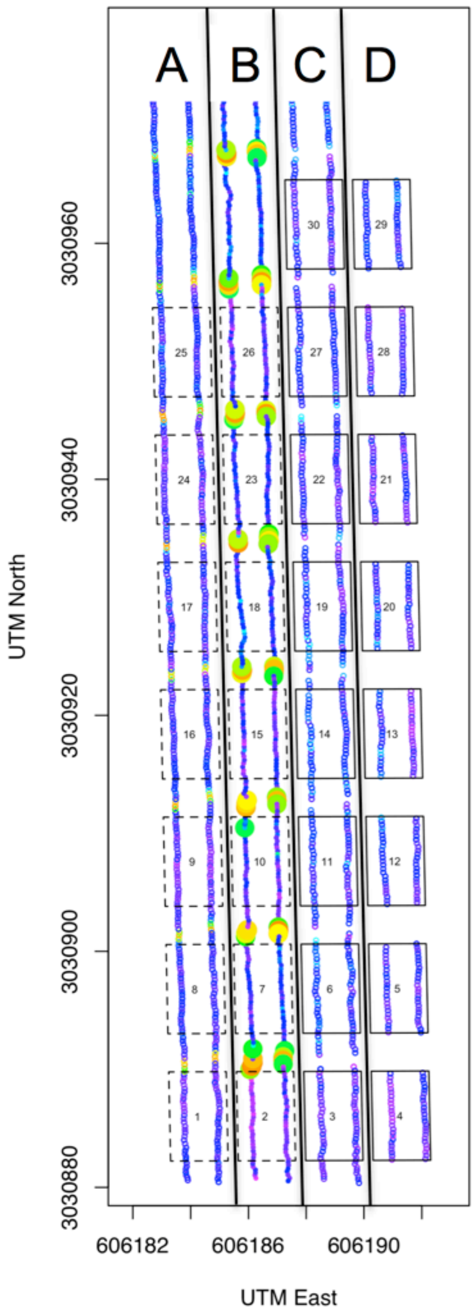


**Figure 2-4. Layout of experiment and workflow for assigning plot coordinates. Workflow is shown in progressive columns A–D; each step occurred for all experimental plots. (A) All data as recorded by Phenocart; plot boundaries are in dashed lines. (B) Removing low normalized difference vegetation index (NDVI) values occurring in the border. Low NDVI values identified by circles of increased size. (C) Spatially assigning plots by trimming data close in proximity to the borders. (D) Assigned data to coordinates within plot boundaries.**

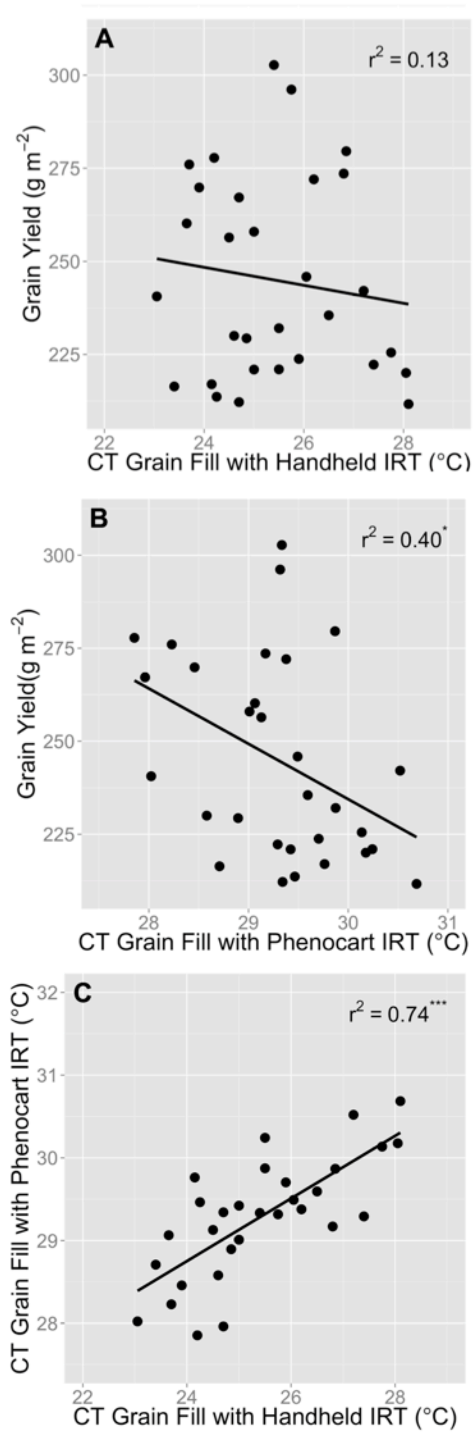


**Figure 2-5. Correlation between Phenocart infrared thermometer (IRT), handheld IRT measurements, and grain yield in wheat plots, Ciudad Obregon, Sonora, Mexico, 2013. (A) Relationship between handheld IRT and grain yield. (B) Relationship between Phenocart IRT and grain yield. (C) Relationship between handheld IRT and Phenocart IRT. Single asterisk (*) indicates model significant at the 0.05 level of probability. Triple asterisks (***) indicates model significant at the <0.001 level of probability**

**Table 2-1. Correlation coefficient between Phenocart infrared thermometer (IRT), handheld IRT, and grain yield for vegetative and grain-filling growth stages in wheat, Ciudad Obregon, Sonora, Mexico.**

| Date | Growth stage | Canopy temperature vs. yield<br>Phenocart IRT | <br>Handheld IRT | Between instruments<br>Phenocart IRT vs. handheld IRT |
|---|---|---|---|---|
| 29 Mar. 2013 | Vegetative | −0.02 | −0.38* | −0.08 |
| 1 Apr. 2013 | Vegetative | −0.34 | −0.29 | 0.84*** |
| 4 Apr. 2013 | Vegetative | −0.22 | −0.73*** | 0.52** |
| 10 May 2013 | Grain fill | −0.55** | −0.21 | 0.67*** |
| 14 May 2013 | Grain fill | −0.40* | −0.13 | 0.74*** |

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

**Table 2-2. Correlation between Phenocart normalized difference vegetation index (NDVI) and grain yield for vegetative and grain-filling growth stages in wheat, Ciudad Obregon, Sonora, Mexico.**

| Date | Growth stage | NDVI vs. yield |
|---|---|---|
| 29 Mar. 2013 | Vegetative | 0.09 |
| 1 Apr. 2013 | Vegetative | 0.24 |
| 4 Apr. 2013 | Vegetative | 0.46* |
| 10 May 2013 | Grain fill | 0.65*** |
| 14 May 2013 | Grain fill | 0.65*** |

* Significant at the 0.05 probability level.

*** Significant at the 0.001 probability level.

**Table 2-3. Broad-sense heritability ($H^2$) for instruments, Phenocart normalized difference vegetation index (NDVI) and canopy temperature, and handheld infrared thermometer (IRT) for each date measured in 2013 and 2014, Ciudad Obregon, Sonora, Mexico.**

| | $H^2$ | | | |
|---|---|---|---|---|
| | NDVI | Canopy temperature | | Photos |
| Date | Phenocart | Phenocart | Handheld IRT | Phenocart |
| 29 Mar. 2013 | 0.66 | 0.69 | 0.35 | NA† |
| 1 Apr. 2013 | 0.80 | 0.24 | 0.51 | NA |
| 4 Apr. 2013 | 0.83 | 0.31 | 0.05 | NA |
| 10 May 2013 | 0.92 | 0.77 | 0.76 | NA |
| 14 May 2013 | 0.94 | 0.33 | 0.51 | NA |
| 6 Mar. 2014 | 0.82 | 0.47 | NA | 0.45 |

† NA, data not collected for particular instrument and date.

# Chapter 3 - Utilizing high-throughput phenotypic data for improved phenotypic selection of stress adaptive traits in wheat

This chapter has been submitted to Crop Science as the following journal article:

Crain, J.L., Reynolds, M., Poland, J., 2016. Utilizing high-throughput phenotypic data for improved phenotypic selection of stress adaptive traits in wheat.

## Abbreviations

BLUE, best linear unbiased estimator; CENEB, Campo Experimental Normal E. Borlaug; CIMMYT, International Maize and Wheat Improvement Center; CT, canopy temperature; H2, broad-sense heritability; HTP, high-throughput phenotyping; i.i.d., independent and identically distributed; NDVI, normalized difference vegetation index; RIL, recombinant inbred line; VPD, vapor pressure deficit

## Abstract

Efficient phenotyping methods are key to increasing genetic gain and precisely mapping genetic variation. Recent phenotyping developments have resulted in high-throughput phenotyping (HTP) platforms that utilize proximal sensing to simultaneously measure multiple physiological traits. However, there has been limited exploration of multiple dimensions of this high-resolution data. To address this, two wheat (*Triticum aestivum)* bi-parental populations were grown for three years under drought and heat stress at the International Maize and Wheat Improvement Center, Ciudad Obregon, Mexico. The lines were evaluated at multiple time points throughout the growing season with the "Phenocart" a portable field phenotyping platform that integrates precision GPS, spectral reflectance and thermal sensors. Both normalized difference

vegetation index (NDVI) and canopy temperature (CT) were correlated to final grain yield. We found that broad-sense heritability ($H^2$) and correlation to yield for both NDVI and CT had a regular pattern over the growing season. The maximum correlation and $H^2$ existed during mid-grain fill stage while correlations were low for early and late season measurements. We also found that the $H^2$ of CT on a given day was a good indication of how well that dataset correlated to yield. In addition, the temporal NDVI data from heading to senescence was modeled to evaluate stay-green and senescence differences between genotypes. Based on the repeatable correlations, HTP platforms can be used to assist with indirect selection through rapid collection of physiological measurements that could provide a 3-13% increase in gain from selection compared to direct selection for grain yield alone.

## Introduction

Wheat is one of the most important cereal crops in the world accounting for 26% of world cereal production and 44% of world cereal consumption in 2011 (Food and Agriculture Organization of the United Nations). Although advances in wheat yield production have been made, the gain is currently less than 1% a year (Reynolds et al., 2012), and this progress is expected to be reduced by increasing climatic variability (Wiebe et al., 2016). For example, one climate scenario calls for re-classifying up to 51% of the Indo-Gangetic Plains as heat stressed by 2050, an area that currently produces 15% of the world’s wheat supply (Ortiz et al., 2008). Given that 81% of wheat consumed in developing countries is produced domestically (CIMMYT, 2005) this plausible scenario poses a real threat to hundreds of millions of people.

Numerous studies have reported the impact of increasing temperature on wheat yield with estimates for losses ranging from 3-50% based on severity of increasing temperatures (Wardlaw and Wrigley, 1994; You et al., 2009; Asseng et al., 2011, 2014). In addition to yield loss due to increasing temperatures, changing rainfall conditions could also lead to more drought affected areas further depressing yield. Fischer and Maurer (1978) reported grain yield under varying drought conditions ranging from 37-86% of the control plot yields. Globally, it is estimated that the effect of climate change on cereal production from 1980-2008 have decreased wheat yields nearly 5.5% compared to scenarios without climate trends (Lobell et al., 2011). Additionally, current global crop model estimates suggest a 6% yield decrease for each one degree Celsius increase in temperature (Asseng et al., 2014). With a growing global population and the threat of climate change novel methods are needed to ensure the safety of this food supply.

Efficient and accurate phenotyping methods are needed for breeding programs to fully realize the benefits of plant breeding. In particular, phenotyping is often considered the bottle neck of crop breeding (Araus and Cairns, 2014). To address this constraint, several field based high-throughput phenotyping (HTP) platforms have been developed in the past few years to successfully measure phenotypic traits (Busemeyer et al., 2013a; Andrade-Sanchez et al., 2014; Crain et al., 2016) as well as provide opportunity for genetic studies (Busemeyer et al., 2013b). Even though HTP offers solutions for increasing genetic gain (Cobb et al., 2013), phenomics is often costly (White et al., 2012; Cobb et al., 2013) and data analysis poses a great challenge. Given that HTP technologies need to be affordable and readily accessible for developing countries to implement and realize genetic gains (Tester and Langridge, 2010) low-cost methods have been developed (White and Conley, 2013; Crain et al., 2016). By deploying affordable instruments and high-throughput phenotyping methods developing countries should be able realize more success in their breeding programs.

Several physiological traits that are amenable to high-throughput phenotyping have been proposed for use in breeding programs. For example, spectral reflectance has long been advocated as a selection tool in wheat breeding programs (Babar and Reynolds, 2006; Babar et al., 2006; Prasad et al., 2007). Another trait that has been suggested for selection is canopy temperature, with several studies finding correlation between canopy temperature and grain yield (Amani, 1996; Balota et al., 2007). With low cost infrared thermometers and spectral reflectance measuring devices both of these traits can be measured quickly and effectively.

In addition to providing a single measurement to evaluate lines, there is also the possibility to combine multiple data sets to assess temporal dynamics of plant growth and development. Lopes and Reynolds (2012) used multiple measurement dates of normalized difference vegetation index (NDVI) to estimate stay-green in wheat. Stay-green is measured as delayed leaf senescence (Thomas and Smart, 1993; Thomas and Howarth, 2000), and has been shown to be beneficial to crop yield in a variety of crops (ex. Borrell et al., 2000, 2014), including wheat (Vijayalakshmi et al., 2010). Environmental factors such as high temperature, may induce early senescence in plants resulting in lower grain yield and quality (Xu et al., 2000). While Vijayalakshimi *et al.* (2010) extensively studied stay-green in wheat using Gompertz nonlinear curves to model senescence and completed QTL mapping, it is relevant to note that stay-green was characterized by visual observations and plants were grown in a growth chambers. Lopes and Reynolds (2012) suggested using spectral reflectance sensors (e.g. GreenSeeker, Trimble Navigation Ltd., Sunnyvale, CA) that provide a continuous value for estimating stay-green rather than discrete visual scoring. Thus opportunity exists to model stay-green under field conditions, and this trait could be very relevant for breeders working to develop stress-adapted material.

To simultaneously explore abiotic tolerance to drought and heat and the ability to collect high-density, low cost phenotypic data, we studied two bi-parental populations of wheat that were derived from parents contrasting in abiotic stress tolerance at the International Maize and Wheat Improvement Center (CIMMYT). We deployed an affordable field phenotyping platform, Phenocart, which utilizes precision GPS to geo-reference each sensor measurement to collect high-throughput data (Crain et al., 2016). Phenocart uses commercial sensors to collect NDVI and canopy temperature (CT). Our objectives were to: 1) Characterize the physiological populations with respect to NDVI and CT, 2) Evaluate the potential of HTP to provide novel meaningful dense phenotypic data, 3) Determine the effect of stay-green by modeling senescence curves of the lines, and 4) Identify selection strategy for grain filling and senescence in wheat.

## Materials and Methods

### Genetic Material

Two bi-parental populations between parents contrasting for heat and drought stress were developed at CIMMYT. The recombinant inbred lines (RILs) had 168 entries in Family 8 and 162 entries in Family 5. The populations were selected for no more than a 10-day range in flowering date, and were derived from parents that did not show GxE for flowering time in 15 contrasting temperature and photoperiod environments.

## Field Design and Management

The two populations were grown for three seasons (2012-2013, 2013-2014, and 2014-2015) at the International Maize and Wheat Improvement Center, Campo Experimental Normal E. Borlaug (CIMMYT, CENEB). Each season, the two populations were grown under reduced water (drought) and heat environments. The drought experiment was sown in early December at optimal planting time with irrigation ceased at the booting growth stage. This resulted in a terminal drought stress on the plants. The heat environment was induced by a late planting date in mid March leading to higher than optimal temperatures throughout the growing season. Full irrigation was maintained throughout the heat trial.

The trials were planted as an alpha lattice design with two replications. The experimental units were 2m x 0.8m in size. Nutrient levels were maintained at optimum rates, and pesticides were applied as needed. Grain yield was harvested with a plot combine, except for the drought 2014 trial that was harvested by hand to prevent losses from lodging.

## Phenotypic and High-Throughput Phenotypic Data Collection

Throughout the growing season various phenotypic data was recorded. Data recorded included heading date, plant height, and grain yield for all trials, and anthesis date in the drought trials. Anthesis can occur before heading under heat stress so only heading date was recorded in the heat trials (Pask et al., 2012).

In addition to these phenotypic data, we employed Phenocart to collect geo-referenced NDVI and canopy temperature (Crain et al., 2016). The Phenocart records data at 10Hz, which resulted in 12-15 measurements per plot per date after trimming plot ends to avoid border effect. This data was taken at multiple time points throughout the growth cycle from tiller elongation though senescence. We used similar algorithms to those presented in Crain *et al.* (2016) for assigning data to plots and curating the data. Briefly the analysis of raw data consists of 1) Defining plot boundaries, 2) Assigning data that fell within plot boundaries to that plot, 3) Curating the data in a database.

## Data Analysis

All data analysis was conducted in R (R Core Team, 2015). Principal component analysis used the stats package (R Core Team, 2015) to investigate relationships among the numerous phenotypic observations observed. Minimal filtering was conducted on the data sets as a whole, but some filtering was used to discard outliers as noted. Only a subset of 40-90 days from planting was used for relationship between heritability to correlation of grain yield, as this removed extremely early and late observations, and significantly improved correlations.

### *Trait heritability*

The broad-sense heritability, often termed repeatability (Piepho and Möhring, 2007) was calculated as:

$$Equation\ 1: \quad H^2 = \frac{\sigma_g^2}{\sigma_p^2}$$

where $\sigma_g^2$ is the genotypic variance, and $\sigma_p^2$ is the phenotypic variance. For grain yield, plant height, and heading date broad-sense heritability was calculated across all years and environments according to equation 2 (Holland et al., 2003):

$$Equation\ 2: \quad H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{gy}^2}{y} + \frac{\sigma_{ge}^2}{e} + \frac{\sigma_{gye}^2}{ye} + \frac{\sigma_{error}^2}{yer}}$$

where $\sigma_{gy}^2$ is genotype by year variance, $\sigma_{ge}^2$ is genotype by environment interaction, $\sigma_{gye}^2$ is genotype by year by environment interaction, and $\sigma_{error}^2$ is residual error variance. The number of years, environments, and replications are signified by *y*, *e*, and *r* respectively. Heritability was also calculated within each environment by dropping the environment terms in Equation 2. We also calculated heritability for each HTP data set for the individual trial and date using equation 3:

$$Equation\ 3: \quad H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{plot}^2}{r} + \frac{\sigma_{error}^2}{rs}}$$

Equation 3 takes into account the multiple measurements of each parameter (pseudoreplication) by fitting a variance term for plot level and subsample level where s is the number of subsamples per plot, $\sigma_{plot}^2$ is the variance among plots, and $\sigma_{error}^2$ is the residual variance (within plot variance) and r is the number of replications. To fit equations all components (year, environment, rep, block, entry, plot) were entered as random effects into a mixed model using the lme4 (Bates et al., 2014) and asreml (Butler, 2009) packages in R.

*Senescence Curves*

Senescence was modeled as a logistic growth curve using the following nonlinear regression model (Fox and Weisberg, 2011):

$$Equation\ 4:\ \ y = \frac{\theta_1}{1 + e^{(-(\theta_2 + \theta_3 x)}} + \epsilon$$

Where *y* is the response, $\theta_1$ is the maximum value (asymptote), $\theta_2$ is a parameter that models the onset of senescence, $\theta_3$ is the rate of senescence, $\epsilon$ is error, and x is time in days after (before) heading. To fit the model, only observations that were 7 days before heading or later were used. This ensured that the model started at maximum NDVI and then NDVI values declined as time progressed throughout senescence. The nlme (Pinheiro et al., 2015) R package was used for model fitting.

*Parameter Estimation*

Best linear unbiased estimators (BLUEs) were calculated for each line using the following mixed model:

$$Equation\ 5:\ y_{ijklm} = \mu + \ g_i + t_j + d_{k(j)} + r_{l(j,k)} + b_{m(j,k,l)} + \epsilon_{ijklm}$$

Where $y_{ijklm}$ is the phenotype for trait of interest, $\mu$ is the overall mean of the population, $g_i$ is the fixed genotype effect of entry *i*, $t_j$ is random effect for each year of measurement with independent and identically distributed (i.i.d.) $t_j \sim N(0, \sigma_j^2)$, $d_{k(j)}$ is the random effect for the k environment nested within year with i.i.d. $d_{k(j)} \sim N(0, \sigma_k^2)$, $r_{l(j,k)}$ is the random effect for the *l* rep nested within year and environment distributed as i.i.d. $r_{l(j,k)} \sim N(0, \sigma_l^2)$, $b_{m(j,k,l)}$ is the random effect for the m block nested within year, environment, and rep distributed as i.i.d. $b_{m(j,k,l)} \sim N(0, \sigma_m^2)$, and $\epsilon_{ijklm}$ is the residual distributed as i.i.d. $\epsilon_{ijklm} \sim N(0, \sigma_\epsilon^2)$. BLUEs calculated included grain yield, days to heading, plant height, and the three parameters from the senescence curve. Once BLUEs were calculated Pearson correlation was used to determine the relationship between grain yield and measured phenotypic traits.

*Gain from Indirect Selection*

We estimated the gain that could be achieved from using indirect selection for traits correlated to yield. The formulas for correlated response to selection are provided by Falconer and MacKay (1996).

$$Equation\ 6\colon\ \frac{CR_X}{R_X} = \frac{i_Y h_Y r_A \sigma_{AX}}{i_X h_X \sigma_{AX}}$$

Where $CR_X$ is the correlated response of character X resulting from selection on character Y, $R_X$ is the response of character X by applying direct selection on X, $i$ is selection intensity applied to X and Y characters respectively, $h$ is square root of heritability for character X and Y, $\sigma_{AX}$ is the additive genetic standard deviation of trait X, and $r_A$ is the additive genetic correlation between characters X and Y. The additive genetic correlation can be related to the phenotypic correlation as:

$$Equation\ 7\colon\ r_P = h_X h_Y r_A$$

When selection intensity is not changed, rearrangement and substitution of equation 6 and 7 result in the gain for indirect selection

$$Equation\ 8\colon\ \frac{r_P}{h_X^2}$$

Equation 8 was used to determine what types of genetic gains could be achieved by utilizing HTP data assuming the same selection intensity for both direct and indirect selection. This equation utilizes narrow-sense heritability, which only captures the additive portion of genetic variance. Because narrow-sense heritability is equal to or less than broad-sense heritability equation 8 provides a conservative estimate of the gain from indirect selection. If selection intensity was increased for the correlated trait ($i_Y$) the gain from selection would further escalate.

# Results and Discussion

## Data Collection and Modeling

In order to evaluate the suitability of HTP for phenotypic selection, we deployed a small field based platform, Phenocart, to collect geo-referenced NDVI and canopy temperature across three years and two environments in two different RIL populations. For each environment we collected multiple HTP data sets, with the majority of the data collected from booting through senescence. The HTP platform was efficient in data collection measuring each family trial (approximately 300 plots) in less than 15 minutes, and each trial was measured at multiple time points throughout the growing season (Table 1). In order to objectively measure plant senescence, we measured NDVI extensively between heading and maturity. We fit a logistic growth curve to this data using nonlinear least squares regression. This model decomposed our season long measurements (throughout senescence) into three parameters to model the growth. The parameters related to maximum NDVI value ($\theta_1$), a time of initiation in senescence ($\theta_2$), and the rate of senescence ($\theta_3$). Our final data set was large with greater than 500,000 phenotypic observations for each family with which to determine correlation to grain yield, NDVI, canopy temperature, and senescence parameter characteristics.

## Phenotypic Trait Heritability

Broad-sense heritability on a line mean basis was calculated for grain yield, plant height, heading date, senescence parameters, and for each time point of NDVI and canopy temperature. Broad-sense heritability is a reflection of both the genetic variance and the level of precision that can be achieved within and across trials (Piepho and Möhring, 2007). We computed heritability across all environments and years, within environments, and within each trial (Tables 2, 3 and 4). For both populations the heritability for grain yield, plant height, and days to heading was high for each stratification level of measurement and maintained relatively constant magnitudes. Trait heritability for fitted senescence curves parameters ($\theta_{1,2,3}$) was much lower and there was less consistency between populations. It appears that within individual trials these parameters show moderate heritability, but upon combining years or environments the heritability decreases rapidly suggesting high genotype-by-environment interaction.

We also calculated the broad-sense heritability of NDVI and CT by each measurement day. The heritability at different time points varied widely with a range from a low of 0 to a high of 0.87 (Figure 1). NDVI spectral reflectance consistently had higher heritability than CT. The $H^2$ range agrees with other studies using HTP data ($H^2$ =0.28–0.90 for NDVI and $H^2$ =0.01–0.90 for canopy temperature (Andrade-Sanchez et al., 2014)). The higher heritability from canopy reflectance could be expected as the GreenSeeker sensor is an active sensor and is not influenced by ambient conditions. In contrast, canopy temperature measurements are strongly influenced by daily environmental conditions (Pask et al., 2012), and while efforts were taken to mitigate environmental factors, it is difficult to completely eliminate them. For spectral reflectance, the general trend was for heritability measurements to begin at low values and then increase until reaching a plateau around mid-grain filling stage. For late season NDVI measurements, there was a slight decline in heritability, which would be expected, as all lines would approach full senescence at which point NDVI values for each genotype would be at the baseline value.

The heritability for canopy temperature was highly variable, rarely exceeding $H^2 > 0.5$, and had a much less consistent trend throughout the growing season and across years and environments. In addition, there were wide swings from one measurement day to the next. Even when heritability was high (Family 5 and 8, heat trials 2014) there were extreme variation in the heritability from day to day measurements, confirming that useful CT measurements must be taken on an optimal day with full sun, low humidity and no wind (Pask et al., 2012). Based on this daily variability, it would appear more difficult to use this trait with certainty in climates with unstable conditions.

## HTP Trait Relationships to Yield

The extensive number of phenotypic parameters evaluated, NDVI and CT for each date of observation, created a complex data set with potentially large amounts of redundancy. To evaluate which variables were contributing unique information, we performed a principal component analysis on all of the phenotypic parameters (NDVI and CT for each data measurement, senescence parameters, plant height, and heading date). We only used principal components accounting for more than two percent of the total variation, which were the first ten components for each population. For the selected principal components variation within each component attributable to CT and NDVI ranged from 18-57% and 29-69% respectively. Each population displayed similar trends within the PCA analysis (Table 5). The results from the PCA show that both CT and NDVI are providing novel information that can be exploited for breeding purposes.

To assess the value of using HTP measurements for variety selection we investigated the relationship between phenotypic traits to grain yield. For both populations, correlations between grain yield, plant height and days to heading were significant in both populations and relatively consistent in magnitude (Table 6). Even though heritability for the $\theta$ parameters was low, there were still moderate correlations between the senescence parameters and grain yield, but the magnitude and significance of the parameter varied between populations. This lack of consistency in the correlation may be expected as the traits had lower heritability than plant height and planting date. However, there was a near perfect correlation between $\theta_2$ and $\theta_3$ in each of the populations. Because of the populations studied, it is difficult to discern if this correlation is biological relevant or an artifact of the narrow genetic diversity, but it may warrant further investigation in breeding populations.

Along with assessing senescence parameters over all years and environments, we also examined each individual trial for correlation between yield and senescence parameters (Table 7). Upon further investigation, a general trend was that plant height and days to heading were usually more highly correlated to final grain yield than any of the senescence parameters. For example, within the heat trial in 2013, $\theta_1$was significantly related to yield. $\theta_1$ being a parameter for maximum NDVI value, this would suggest that selection could occur by finding the highest NDVI genotypes. However, for the same population and environment for the following year $\theta_1$ was not significantly correlated, but in contrast $\theta_2$ and $\theta_3$ were correlated. Based on these observations, there is no conclusive evidence for improved selection based on modeling senescence curves with NDVI. If there were a common trend in the data, then selection could target certain parameters, whether that be maximum NDVI value ($\theta_1$) or rate of senescence ($\theta_3$) as each parameter could be related to underlying physiology. Maximum NDVI could occur with large amount of green biomass, leading to a non-stressed high yielding variety, while a fast rate of senescence could allow the plant to maintain maximal carbohydrate translocation to the seed and then rapidly senesce after completion of grain filling. Another option would be to delay senescence, the function of $\theta_2$, and suggested by Thomas and Howarth (2000).

For each day of HTP data measurement, we correlated the measured trait to final grain yield (Figure 2). Across both populations, there were several consistent trends. Consistent with previous studies, NDVI was positively correlated with yield, while canopy temperature was negatively correlated with yield (Babar et al., 2006; Balota et al., 2007; Pinto and Reynolds, 2015). In general, the correlation between grain yield and the HTP traits were higher under the heat trials than the drought trials, suggesting that physiological HTP measurements may be more useful as a selection tool under heat trials. CT (under well-watered conditions) is driven largely by vapor pressure deficit (VPD), and VPD increases with air temperature which improves the expression and resolution of CT under heat environments (Amani, 1996). Canopy temperature was quite variable from day to day with little discernable trend. Taken as a whole canopy temperature usually had less correlation to yield than NDVI, thus more observation days where indirect selection would be less efficient due to the lower correlation.

### Utilization of HTP Traits for Indirect Yield Selection

Based on the results of trait heritability and correlation to yield as well as the ability to take rapid measurements with HTP, NDVI and canopy temperature could be used for indirect selection for grain yield under heat and drought. To see if there was any way to determine optimal data quality, we examined the relationship between heritability of the HTP traits for each measurement day and the correlation coefficient between grain yield and the HTP measurements on those respective days (Figure 3). We used a subset of measurement days, 40-90 days from planting, and observed that as heritability of canopy temperature increased, the strength of the correlation with yield increased. Based on the strength of the trend, it would suggest when canopy temperature on a given day has a high broad-sense heritability, that the predictive power of final grain yield would be high. NDVI did not present as significant trend in relationship between heritability and correlation to grain yield, although both populations had similar trends, after the removal of outliers in Family 8. Based on these data, CT can be optimized by selectively using dates that have a high heritability. For canopy temperature, broad-sense heritability provides a good assessment of data quality and predictive ability for estimating final grain yield.

Finally, we investigated the effect of indirect selection using CT or NDVI on grain yield. Using the phenotypic correlations between CT and NDVI measurements and yield along with the heritability of yield, we estimated gain from indirect selection using HTP measurements. From equation 8, it is apparent that indirect selection will perform better than direct selection when the phenotypic correlation is higher than the square of heritability of grain yield. Additionally, if selection intensity can be increased in the correlated response (HTP) trait the response to selection can be increased as well. We had 122 and 120 HTP data sets for Family 5 and Family 8 respectively which were evaluated for indirect selection. Using indirect selection based on HTP measurements alone produced very little gain. Less than 10% of the observation sets alone produced an increased gain from indirect selection with the average gain of 5%. However, if selection intensity could be increased, HTP traits provided larger gains. Many breeding programs use a selection intensity of 15% (Ornella et al., 2014), and if selection intensity could be increased to 10% HTP would provide more genetic gain. Using this higher selection intensity resulted in 13% and 9% higher gains than selection for yield alone for Family 5 and 8, respectively. It should be noted that we used known values for heritability and phenotypic correlation, in practice this will not be known until harvest so the amount of gain realized may be less. Nonetheless, the magnitude of the percent gain suggest that indirect selection could play an important role in selecting improved cultivars. Using HTP will allow breeders to identify superior lines, discarding inferior lines sooner, resulting in more efficient use of resources.

Not every phenotypic observation days were useful, and it appears that the data sets most likely to increase genetic gains had high heritability, Figure 4, confirming that calculating broad-sense heritability is an important aspect of data quality assessment. While there were days where indirect selection failed to enhance genetic gain, HTP provides a method to easily measure correlated traits. As Falconer and MacKay (1996) note indirect selection may be useful when the trait of interest is difficult to measure. Grain yield is a variable that can be measured only once at the end of the growing season, thus using HTP methods throughout the growing season could be used to assist in identify superior varieties.

## Conclusions

Phenotyping is essential to plant breeding, and enhancements in phenotyping capabilities will give breeders new options to ensure sufficient food, fiber, and fuel production. Using Phenocart, we generated dense phenotypic data and applied various methods to extract novel and meaningful information from spectral reflectance and canopy temperature. Both canopy temperature and NDVI were related to final grain yield, and using heritability of measurements assisted with prediction accuracy. While the senescence models were not predictive of yield across environments and genotypes, this could be a result of the extensive abiotic stresses applied. Additionally, the methods used collected large quantities of data, which are not normally captured in public breeding programs. This data provided increased selection accuracy as well as more opportunities to evaluate genotypes. Future applications of these types of phenotyping efforts will assist the genomics arena in increasing selection efficiency and accuracy.

## Acknowledgements

This work was supported through the National Science Foundation - Plant Genome Research Program (IOS-1238187) and the US Agency for International Development Feed the Future Innovation Lab for Applied Wheat Genomics (Cooperative Agreement No. AID-OAA-A-13-00051). J. Crain is supported through a fellowship from the Monsanto Beachell-Borlaug International Scholars program. SK Reddy provided valuable comments on an earlier version of this paper.

## References

Amani, I. 1996. Canopy temperature depression association with yield of irrigated spring wheat cultivars in a hot climate. J. Agron. Crop Sci. 129: 119–129.

Andrade-Sanchez, P., M.A. Gore, J.T. Heun, K.R. Thorp, A.E. Carmo-Silva, A.N. French, M.E. Salvucci, J.W. White, and E.A. Carmo-Silva. 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Funct. Plant Biol. 41(2007): 68–79.

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19(1): 52–61.

Asseng, S., F. Ewert, P. Martre, R.P. Rötter, D.B. Lobell, D. Cammarano, B. A. Kimball, M.J. Ottman, G.W. Wall, J.W. White, M.P. Reynolds, P.D. Alderman, P.V. V. Prasad, P.K. Aggarwal, J. Anothai, B. Basso, C. Biernath, A. J. Challinor, G. De Sanctis, J. Doltra, E. Fereres, M. Garcia-Vila, S. Gayler, G. Hoogenboom, L. A. Hunt, R.C. Izaurralde, M. Jabloun, C.D. Jones, K.C. Kersebaum, A-K. Koehler, C. Müller, S. Naresh Kumar, C. Nendel, G. O'Leary, J.E. Olesen, T. Palosuo, E. Priesack, E. Eyshi Rezaei, A. C. Ruane, M. A. Semenov, I. Shcherbak, C. Stöckle, P. Stratonovitch, T. Streck, I. Supit, F. Tao, P.J. Thorburn, K. Waha, E. Wang, D. Wallach, J. Wolf, Z. Zhao, and Y. Zhu. 2014. Rising temperatures reduce global wheat production. Nat. Clim. Chang. 5(2): 143–147.

Asseng, S., I. Foster, and N.C. Turner. 2011. The impact of temperature variability on wheat yields. Glob. Chang. Biol. 17(2): 997–1012.

Babar, M. A., M. Van Ginkel, A. R. Klatt, B. Prasad, and M.P. Reynolds. 2006. The potential of using spectral reflectance indices to estimate yield in wheat grown under reduced irrigation. Euphytica 150(1-2): 155–172.

Babar, M., and M. Reynolds. 2006. Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. Crop Sci. 46(3): 1046–1057.

Balota, M., W. Payne, S. Evett, and M. Lazar. 2007. Canopy temperature depression sampling to assess grain yield and genotypic differentiation in winter wheat. Crop Sci. 47(4): 1518–1529.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4.

Borrell, A.K., G.L. Hammer, and A.C.L. Douglas. 2000. Does maintaining green leaf area in sorghum improve yield under drought? I. Leaf growth and senescence. Crop Sci. 40(4): 1026–1037.

Borrell, A.K., E.J. van Oosterom, J.E. Mullet, B. George-Jaeggli, D.R. Jordan, P.E. Klein, and G.L. Hammer. 2014. Stay-green alleles individually enhance grain yield in sorghum under drought by modifying canopy development and water uptake patterns. New Phytol. 203(3): 817–830.

Busemeyer, L., D. Mentrup, K. Möller, E. Wunder, K. Alheit, V. Hahn, H.P. Maurer, J.C. Reif, T. Würschum, J. Müller, F. Rahe, and A. Ruckelshausen. 2013a. BreedVision--a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors (Basel). 13(3): 2830–47.

Busemeyer, L., A. Ruckelshausen, K. Möller, A.E. Melchinger, K. V Alheit, H.P. Maurer, V. Hahn, E. A Weissmann, J.C. Reif, and T. Würschum. 2013b. Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. Sci. Rep. 3: 2442.

Butler, D. 2009. asreml: asreml() fits the linear mixed model.

CIMMYT. 2005. CIMMYT Business Plan 2006-2010. Translating the Vision of Seeds of Innovation into a Vibrant Work Plan. El Batan, Mexico.

Cobb, J.N., G. DeClerck, A. Greenberg, R. Clark, and S. McCouch. 2013. Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126(4): 867–887.

Crain, J.L., Y. Wei, J. Barker III, S.M. Thompson, P.D. Alderman, M. Reynolds, N. Zhang, and J. Poland. 2016. Development and deployment of a portable field phenotyping platform. Crop Sci.

Falconer, D.S. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group.

Falconer, D.S., and T.F.C. Mackay. 1996. Quantitative Genetics. 4th ed. Pearson, New York.

Fischer, R., and R. Maurer. 1978. Drought resistance in spring wheat cultivars. I. Grain yield responses. Aust. J. Agric. Res. 29(5): 897.

Food and Agriculture Organization of the United Nations. FAOSTAT database (FAOSTAT, 2015). Available at http://faostat3.fao.org/ (verified 18 January 2016).

Fox, J., and S. Weisberg. 2011. Nonlinear Regression and Nonlinear Least Squares in R.

Holland, J., W. Nyquist, and C. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: an update. Plant Breed. Rev. 22: 9–112.

Lobell, D.B., W. Schlenker, and J. Costa-Roberts. 2011. Climate trends and global crop production since 1980. Science 333(6042): 616–620.

Lopes, M.S., and M.P. Reynolds. 2012. Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. J. Exp. Bot. 63(10): 3789–3798.

Ornella, L., P. Pérez, E. Tapia, J.M. González-Camacho, J. Burgueño, X. Zhang, S. Singh, F.S. Vicente, D. Bonnett, S. Dreisigacker, R. Singh, N. Long, and J. Crossa. 2014. Genomic-enabled prediction with classification algorithms. Heredity (Edinb). 112(6): 616–26.

Ortiz, R., K.D. Sayre, B. Govaerts, R. Gupta, G. V. Subbarao, T. Ban, D. Hodson, J.M. Dixon, J. Iván Ortiz-Monasterio, and M. Reynolds. 2008. Climate change: Can wheat beat the heat? Agric. Ecosyst. Environ. 126(1-2): 46–58.

Pask, A., J. Pietragalla, D. Mullan, and M. Reynolds. 2012. Physiological breeding II a field guide to wheat phenotyping. Mexico City, Mexico.

Piepho, H.-P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177(3): 1881–8.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Core Team. 2015. nlme: Linear and Nonliner Mixed Effects Models.

Pinto, R.S., and M.P. Reynolds. 2015. Common genetic basis for canopy temperature depression under heat and drought stress associated with optimized root distribution in bread wheat. Theor. Appl. Genet. 128(4): 575–585.

Prasad, B., B. Carver, and M. Stone. 2007. Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under Great Plains conditions. Crop Sci. 47(4): 1426–1440.

R Core Team. 2015. R: a language and environment for Statistical Computing. https://www.r-project.org/.

Reynolds, M., J. Foulkes, R. Furbank, S. Griffiths, J. King, E. Murchie, M. Parry, and G. Slafer. 2012. Achieving yield gains in wheat. Plant, Cell Environ. 35(10): 1799–1823.

Tester, M., and P. Langridge. 2010. Breeding technologies to increase crop production in a changing world. Science 327(5967): 818–822.

Thomas, H., and C.J. Howarth. 2000. Five ways to stay green. J. Exp. Bot. 51 Spec No(February): 329–337.

Thomas, H., and C.M. Smart. 1993. Crops that Stay Green. Ann. Applies Biol. 123: 193–219.

Vijayalakshmi, K., A.K. Fritz, G.M. Paulsen, G. Bai, S. Pandravada, and B.S. Gill. 2010. Modeling and mapping QTL for senescence-related traits in winter wheat under high temperature. Mol. Breed. 26(2): 163–175.

Wardlaw, I., and C. Wrigley. 1994. Heat Tolerance in Temperate Cereals: an Overview. Aust. J. Plant Physiol. 21(6): 695.

White, J., P. Andrade-Sanchez, M.A. Gore, K.F. Bronson, T.A. Coffelt, M.M. Conley, K.A. Feldmann, A.N. French, J.T. Heun, D.J. Hunsaker, M.A. Jenks, B.A. Kimball, R.L. Roth, R.J. Strand, K.R. Thorp, G.W. Wall, and G. Wang. 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133: 101–112.

White, J.W., and M.M. Conley. 2013. A Flexible, Low-Cost Cart for Proximal Sensing. Crop Sci. 53(4): 1646–1649.

Wiebe, K., D. Mason-D'Croz, T. Sulser, N. Cenacchi, S. Islam, R. Robertson, S. Robinson, T. Thomas, T. Zhu, and M.W. Rosegrant. 2016. Agriculture and Food Security in 2050: Socioeconomic and Climate Challenges and Options. IFPRI.

Xu, W., D.T. Rosenow, and H.T. Nguyen. 2000. Stay green trait in grain sorghum: Relationship between visual rating and leaf chlorophyll concentration. Plant Breed. 119: 365–367.

You, L., M.W. Rosegrant, S. Wood, and D. Sun. 2009. Impact of growing season temperature on wheat productivity in China. Agric. For. Meteorol. 149(6-7): 1009–1014.

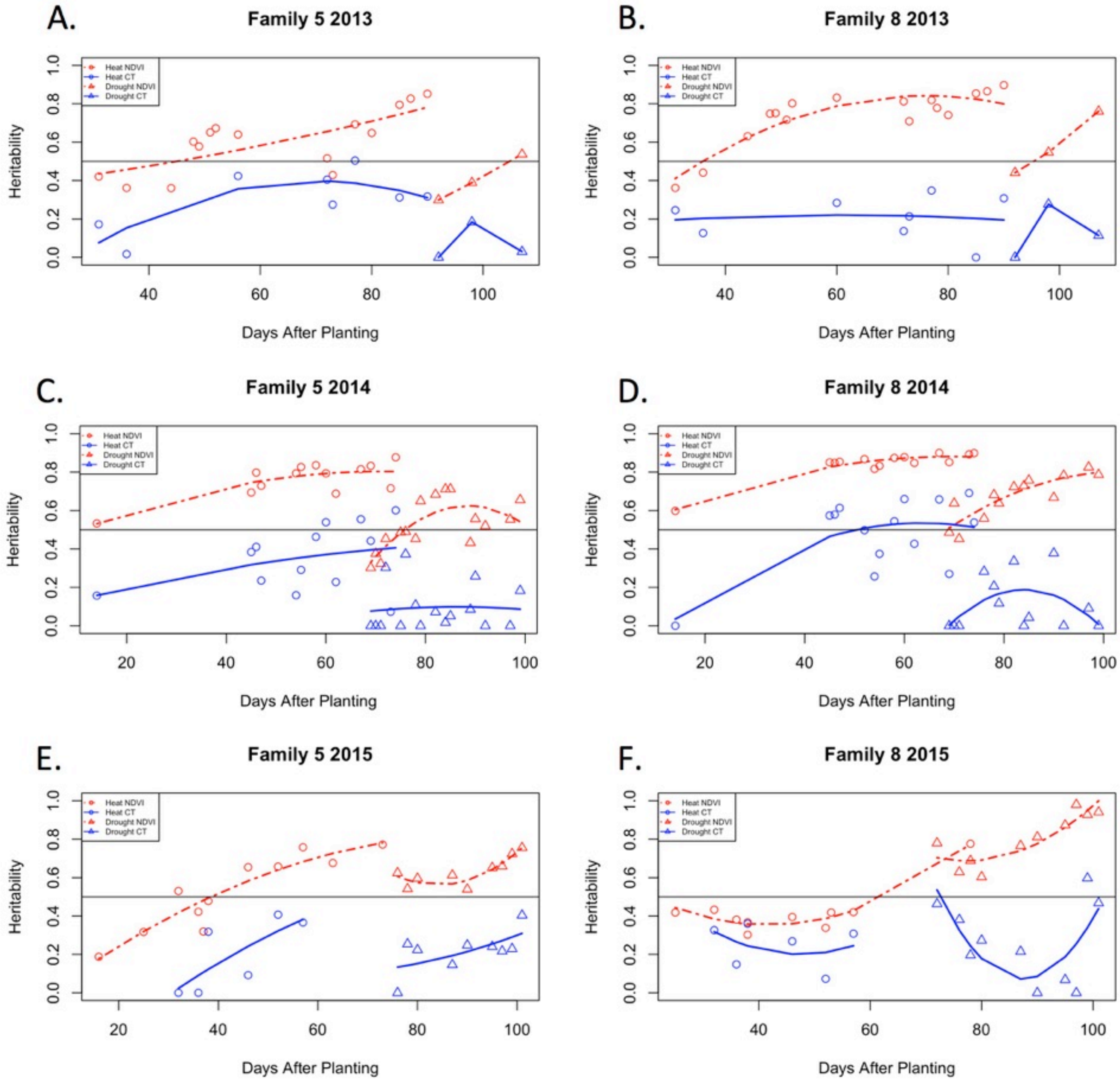


**Figure 3-1. Broad-sense heritability of normalized difference vegetation index (NDVI) and canopy temperature (CT) for each measurement day for two populations. Each figure a-f corresponds to one year for each family, with drought environment observations in triangles, heat observations in circles. Days after planting are on the x-axis with heritability for each day on the y-axis. Points are observations, with lines being smoothed to multiple observations. NDVI plotted with a dashed red line and canopy temperature with a solid blue line. The horizontal line represents 0.5.**

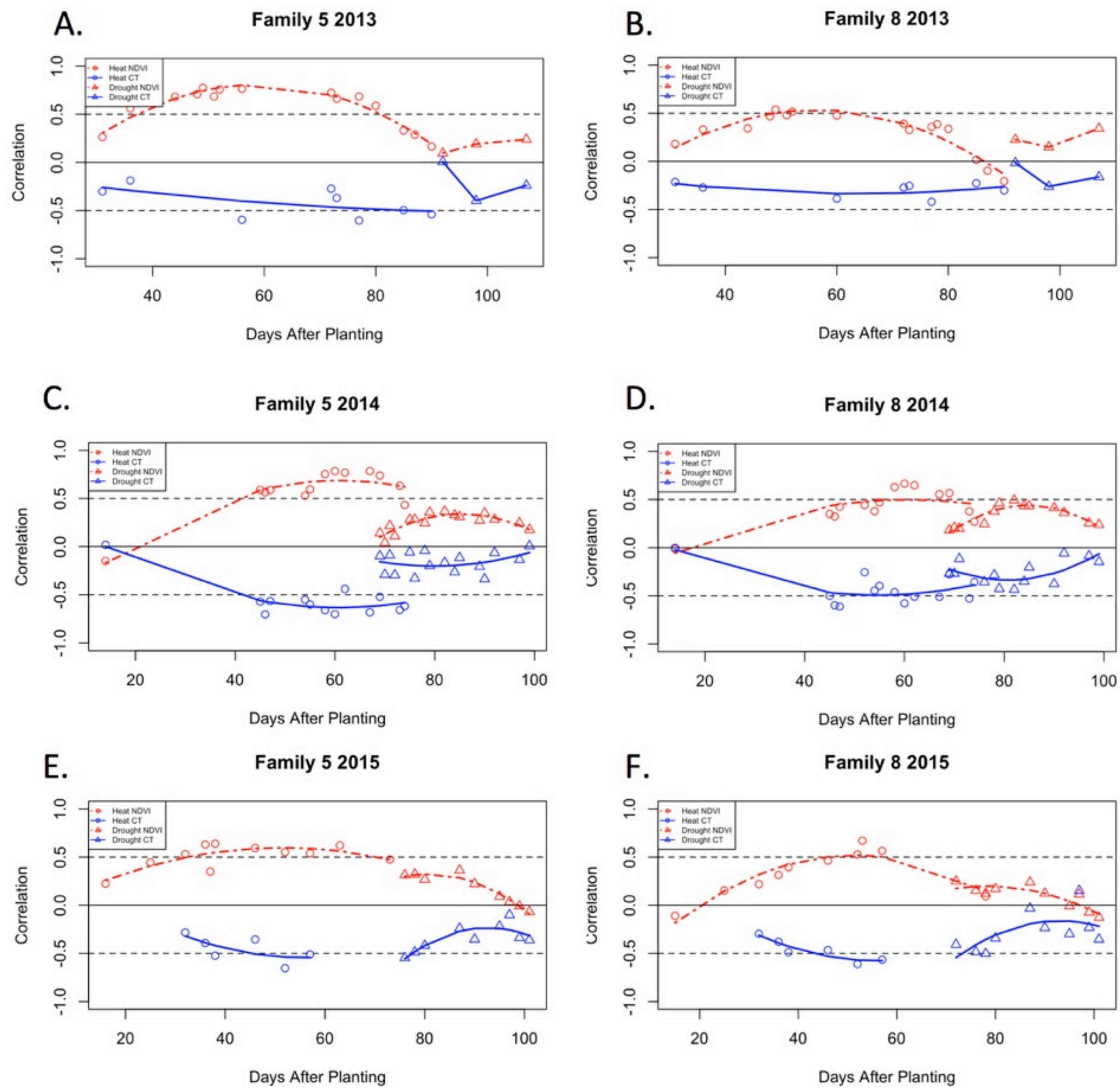


**Figure 3-2. Correlation between normalized difference vegetation index (NDVI) and canopy temperature (CT) measurements and yield for each measurement days. Each figure a-f corresponds to one year for each family, with drought environment observations in triangles, heat environment observations in circles. Dashed red line is smoothed curve to NDVI data, and solid blue line is smoothed fit to canopy temperature data. Days after planting are on the x-axis with correlation to yield on the y-axis. The horizontal lines represent -0.5, 0, 0.5 respectively.**

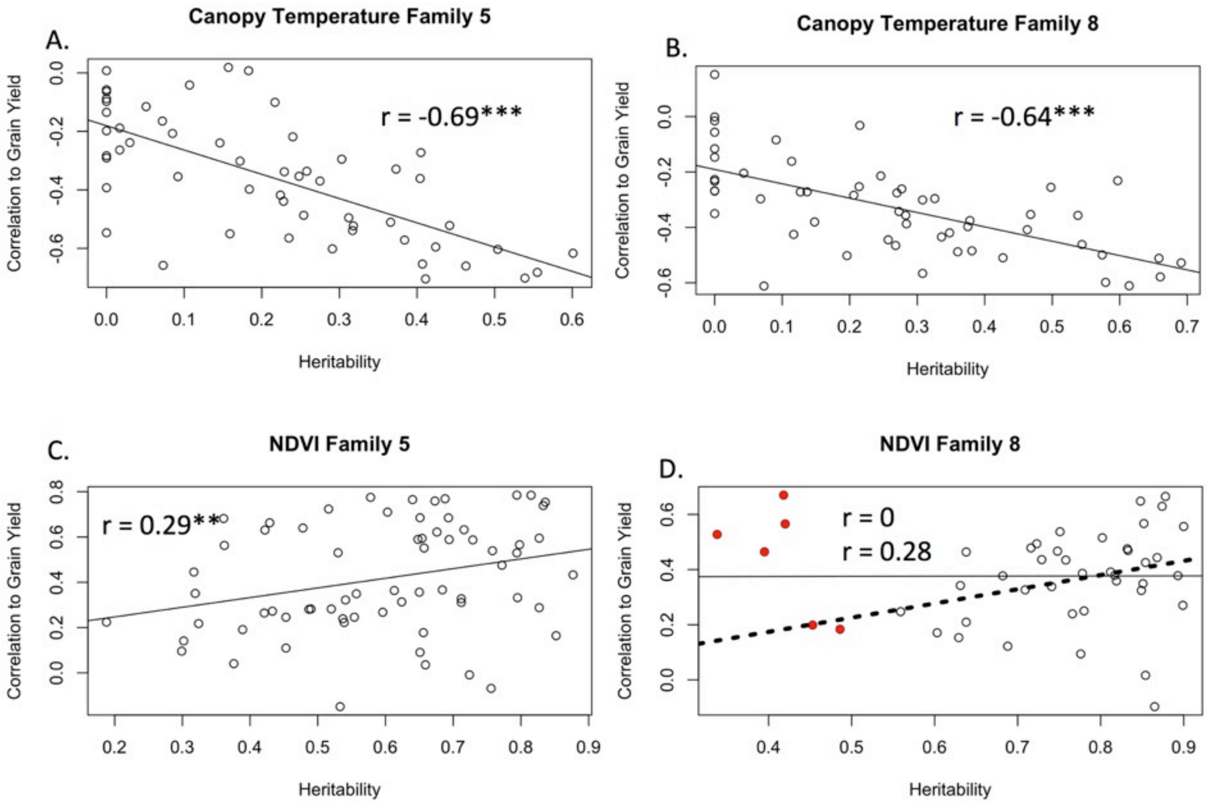


**Figure 3-3. Relationship between calculated heritability for HTP traits and correlation to grain yield for each measurement day. Each figure A-D shows one instrument and one population. Measured heritability is on the x-axis with correlation on the y-axis. Correlation coefficient is noted in each panel with linear regression line. In figure D, two fits are shown, all data, which resulted in no trend, and removing data with heritability under 0.5, data shown in colored circles.**

**Model significant at the 0.01 level of probability

*** Model significant at the <0.001 level of probability.

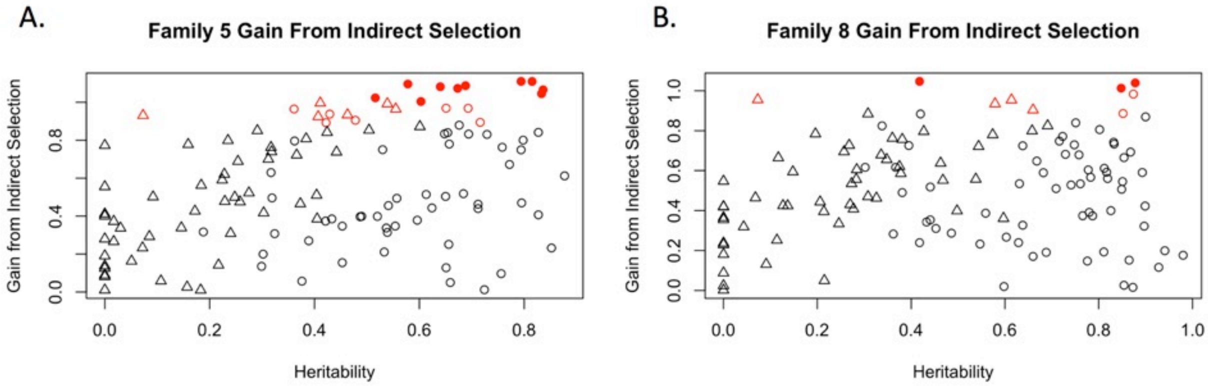


**Figure 3-4. Scatter plot of HTP heritability and genetic gain from indirect selection of CT and NDVI. Each figure A-B shows one population. NDVI is shown in circles and CT is shown in triangles. Filled symbols indicate that indirect selection resulted in more genetic gain that direct selection for yield on these observation dates with equal selection intensity. Symbols with red border are days where increasing selection intensity results in enhanced gain from selection.**

**Table 3-1. HTP observations per trial including number of observation days throughout the growing season, average number of observations per plot, and total number of observations.**

| Population | Year | Environment | Days of Observation | Average Observations per plot per day† | Total Observations per Trial |
|---|---|---|---|---|---|
| Family 5 | 2013 | drought | 3 | 30.1 | 29267 |
| Family 5 | 2014 | drought | 16 | 30.6 | 158452 |
| Family 5 | 2015 | drought | 9 | 18.5 | 53844 |
| Family 5 | 2013 | heat | 15 | 17.6 | 85407 |
| Family 5 | 2014 | heat | 13 | 28.2 | 118790 |
| Family 5 | 2015 | heat | 11 | 15.9 | 56540 |
| Family 8 | 2013 | drought | 3 | 29.8 | 30058 |
| Family 8 | 2014 | drought | 13 | 26.3 | 114979 |
| Family 8 | 2015 | drought | 10 | 15.8 | 52984 |
| Family 8 | 2013 | heat | 16 | 18.6 | 100204 |
| Family 8 | 2014 | heat | 14 | 37.7 | 177346 |
| Family 8 | 2015 | heat | 10 | 18 | 60342 |

†Includes both NDVI and CT, thus 30.1 would be 15 measurements of NDVI and 15 measurements of CT.

**Table 3-2. Heritability for two populations for measured parameters across three years and two environments.**

| | Heritability | |
|---|---|---|
| | **Family 5** | **Family 8** |
| Grain Yield | 0.707 | 0.64 |
| Plant Height | 0.825 | 0.812 |
| Days to Heading | 0.821 | 0.92 |
| $\Theta_1$ | 0.4 | 0.248 |
| $\Theta_2$ | 0.315 | 0.42 |
| $\Theta_3$ | 0.177 | 0.359 |

**Table 3-3. Heritability for two populations for traits measured within each environment.**

| | | Heritability | |
|---|---|---|---|
| **Trait** | **Environment** | **Family 5** | **Family 8** |
| Grain Yield | Drought | 0.657 | 0.491 |
| Plant Height | Drought | 0.766 | 0.861 |
| Days to Heading | Drought | 0.942 | 0.955 |
| $\Theta_1$ | Drought | 0.292 | 0 |
| $\Theta_2$ | Drought | 0.395 | 0.043 |
| $\Theta_3$ | Drought | 0.082 | 0 |
| Grain Yield | Heat | 0.843 | 0.698 |
| Plant Height | Heat | 0.875 | 0.88 |
| Days to Heading | Heat | 0.902 | 0.911 |
| $\Theta_1$ | Heat | 0.361 | 0.766 |
| $\Theta_2$ | Heat | 0.163 | 0.443 |
| $\Theta_3$ | Heat | 0.228 | 0.452 |

**Table 3-4. Heritability for two populations for traits measured within each trial.**

| | | | Trait Heritability | | | | | |
|---|---|---|---|---|---|---|---|---|
| Population | Year | Environment | Grain Yield | Plant Height | Days to Heading | $\Theta_1$ | $\Theta_2$ | $\Theta_3$ |
| Family 5 | 2013 | drought | 0.4314 | 0.7115 | - | - | - | - |
| Family 5 | 2014 | drought | 0.4089 | 0.4208 | 0.8872 | 0.5035 | 0.5949 | 0.5942 |
| Family 5 | 2015 | drought | 0.6513 | 0.6308 | 0.9178 | 0.6836 | 0.6865 | 0.5947 |
| Family 5 | 2013 | heat | 0.7517 | 0.7696 | 0.7992 | 0.5032 | 0.2364 | 0.2426 |
| Family 5 | 2014 | heat | 0.8493 | 0.8549 | 0.9237 | 0.4728 | 0.6497 | 0.7067 |
| Family 5 | 2015 | heat | 0.5287 | - | 0.8069 | - | - | - |
| Family 8 | 2013 | drought | 0.5905 | 0.8105 | - | - | - | - |
| Family 8 | 2014 | drought | 0.4409 | 0.7367 | 0.9593 | 0.5194 | 0.3238 | 0.0793 |
| Family 8 | 2015 | drought | 0.6337 | 0.6717 | 0.9621 | 0.7526 | 0.8659 | 0.8176 |
| Family 8 | 2013 | heat | 0.6357 | 0.8721 | 0.9096 | 0.6584 | 0.7165 | 0.7336 |
| Family 8 | 2014 | heat | 0.8966 | 0.8496 | 0.9469 | 0.6593 | 0.7594 | 0.7709 |
| Family 8 | 2015 | heat | 0.5415 | - | 0.8203 | - | - | - |

- Trait not measured for particular trial.

**Table 3-5. Principal Component Analysis for phenotypic data. Components comprising more than 2% of total variation are displayed along with the percent contribution of CT and NDVI for each loading vector.**

| | | Percent of Variation Explained within Each Principal Component | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Population** | **Measurement** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| Family 5 | NDVI | 0.68 | 0.62 | 0.52 | 0.68 | 0.54 | 0.35 | 0.41 | 0.44 | 0.46 | 0.34 |
| Family 5 | IRT | 0.29 | 0.37 | 0.47 | 0.18 | 0.44 | 0.63 | 0.58 | 0.55 | 0.38 | 0.51 |
| Total Variance† | | 0.25 | 0.10 | 0.08 | 0.06 | 0.04 | 0.04 | 0.04 | 0.04 | 0.03 | 0.02 |
| | | | | | | | | | | | |
| Family 8 | NDVI | 0.69 | 0.79 | 0.49 | 0.50 | 0.48 | 0.40 | 0.42 | 0.44 | 0.30 | 0.42 |
| Family 8 | IRT | 0.30 | 0.19 | 0.38 | 0.50 | 0.48 | 0.57 | 0.54 | 0.53 | 0.47 | 0.50 |
| Total Variance† | | 0.26 | 0.11 | 0.07 | 0.06 | 0.05 | 0.04 | 0.04 | 0.03 | 0.03 | 0.02 |

†Total Variance is the amount of variation of the entire data set explained by the principal component.

**Table 3-6. Pearson correlation between estimated BLUEs for grain yield, plant height, days to heading, and senescence parameters. BLUEs estimated with 3 years and 2 environments. Family 8 correlations presented in the lower triangle, and Family 5 correlations displayed in the upper triangle.**

| | Traits | | | | | |
|---|---|---|---|---|---|---|
| | **Grain Yield** | **Plant Height** | **Days to Heading** | **$\Theta_1$** | **$\Theta_2$** | **$\Theta_3$** |
| Grain Yield | - | 0.46*** | -0.31*** | 0.26*** | 0.12 | -0.14 |
| Plant Height | 0.33*** | - | -0.23** | 0.22** | 0.04 | -0.01 |
| Days to Heading | -0.47*** | -0.16* | - | 0.19** | -0.45*** | 0.30*** |
| $\Theta_1$ | 0.22** | -0.23** | 0.05 | - | -0.60*** | 0.52*** |
| $\Theta_2$ | 0.41*** | 0.57*** | -0.39*** | -0.5*** | - | -0.95*** |
| $\Theta_3$ | -0.41*** | -0.54*** | 0.33*** | 0.43*** | -0.94*** | - |

*Model significant at the 0.05 level of probability
**Model significant at the 0.01 level of probability
*** Model significant at the <0.001 level of probability

**Table 3-7. Pearson correlation coefficients between measured parameters, heading date, grain yield, plant height, and senescence parameters for each trial environment for two populations.**

| Population | Year and Environment | Parameter | Grain Yield | Plant Height | Heading Date | $\Theta_1$ | $\Theta_2$ |
|---|---|---|---|---|---|---|---|
| Family 5 | Heat 2013 | Plant Height | 0.41*** | | | | |
| Family 5 | Heat 2013 | Heading Date | -0.11 | -0.27*** | | | |
| Family 5 | Heat 2013 | $\Theta_1$ | 0.64*** | 0.34*** | 0.01 | | |
| Family 5 | Heat 2013 | $\Theta_2$ | -0.15 | 0.05 | -0.03 | -0.60*** | |
| Family 5 | Heat 2013 | $\Theta_3$ | 0.07 | -0.13 | 0.01 | 0.49*** | -0.95*** |
| Family 5 | Heat 2014 | Plant Height | 0.55*** | | | | |
| Family 5 | Heat 2014 | Heading Date | -0.25* | -0.06 | | | |
| Family 5 | Heat 2014 | $\Theta_1$ | -0.06 | 0.06 | 0.39*** | | |
| Family 5 | Heat 2014 | $\Theta_2$ | 0.37*** | 0.09 | -0.50*** | -0.76*** | |
| Family 5 | Heat 2014 | $\Theta_3$ | -0.31** | -0.08 | 0.45*** | 0.72*** | -0.98*** |
| Family 5 | Drought 2014 | Plant Height | 0.42*** | | | | |
| Family 5 | Drought 2014 | Heading Date | -0.18* | -0.23** | | | |
| Family 5 | Drought 2014 | $\Theta_1$ | 0.13 | 0.14 | -0.24** | | |
| Family 5 | Drought 2014 | $\Theta_2$ | 0.11 | 0.13 | -0.06 | -0.78*** | |
| Family 5 | Drought 2014 | $\Theta_3$ | -0.02 | -0.06 | 0 | 0.78*** | -0.95*** |
| Family 5 | Drought 2015 | Plant Height | 0.40*** | | | | |
| Family 5 | Drought 2015 | Heading Date | -0.17* | 0.04 | | | |
| Family 5 | Drought 2015 | $\Theta_1$ | 0.42*** | 0.25** | 0.07 | | |
| Family 5 | Drought 2015 | $\Theta_2$ | -0.01 | 0.04 | -0.52*** | -0.38*** | |
| Family 5 | Drought 2015 | $\Theta_3$ | 0.08 | -0.02 | 0.29*** | 0.42*** | -0.95*** |
| Family 8 | Heat 2013 | Plant Height | 0.35*** | | | | |
| Family 8 | Heat 2013 | Heading Date | -0.13 | -0.14 | | | |
| Family 8 | Heat 2013 | $\Theta_1$ | 0.30*** | 0.02 | 0.22** | | |
| Family 8 | Heat 2013 | $\Theta_2$ | 0.15 | 0.43*** | -0.17* | -0.65*** | |
| Family 8 | Heat 2013 | $\Theta_3$ | -0.25** | -0.50*** | 0.19* | 0.51*** | -0.95*** |
| Family 8 | Heat 2013 | Plant Height | 0.67*** | | | | |
| Family 8 | Heat 2014 | Heading Date | -0.41*** | -0.12 | | | |
| Family 8 | Heat 2014 | $\Theta_1$ | -0.01 | 0.1 | 0.19* | | |
| Family 8 | Heat 2014 | $\Theta_2$ | 0.63*** | 0.34*** | -0.26** | -0.49*** | |
| Family 8 | Heat 2014 | $\Theta_3$ | -0.61*** | -0.33*** | 0.23** | 0.45*** | -0.97*** |
| Family 8 | Drought 2014 | Plant Height | 0.35*** | | | | |
| Family 8 | Drought 2014 | Heading Date | -0.1 | -0.05 | | | |
| Family 8 | Drought 2014 | $\Theta_1$ | 0.18* | -0.11 | -0.37*** | | |
| Family 8 | Drought 2014 | $\Theta_2$ | 0.20* | 0.43*** | 0.19* | -0.68*** | |
| Family 8 | Drought 2014 | $\Theta_3$ | -0.02 | -0.35*** | -0.12 | 0.48*** | -0.81*** |
| Family 8 | Drought 2014 | Plant Height | 0.20* | | | | |
| Family 8 | Drought 2015 | Heading Date | -0.46*** | 0.1 | | | |
| Family 8 | Drought 2015 | $\Theta_1$ | 0.07 | 0.04 | 0.39*** | | |
| Family 8 | Drought 2015 | $\Theta_2$ | 0.35*** | 0.13 | -0.74*** | -0.45*** | |
| Family 8 | Drought 2015 | $\Theta_3$ | -0.33*** | -0.1 | 0.70*** | 0.48*** | -0.98*** |

*Model significant at the 0.05 level of probability
**Model significant at the 0.01 level of probability
*** Model significant at the <0.001 level of probability

# Chapter 4 - Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding

A chapter to be submitted to The Plant Genome as:

Crain, J.L., Mondal, S., Rutkoski, J., Singh, R.P., and Poland, J. Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding.

## Abbreviations

BLUP, best linear unbiased predictor; CENEB, Campo Experimental Norman E. Borlaug; CIMMYT, International Maize and Wheat Improvement Center; CT, canopy temperature; EN, elastic net; EM, expectation maximization; EYT, Elite Yield Trials; GEBV, genomic estimated breeding value; GS, genomic selection; GBS, genotyping-by-sequencing; HTP, high-throughput phenotyping; $H^2$, Broad sense heritability; MAF, minor allele frequency; MAS, marker assisted selection; NDVI, normalized difference vegetation index; NGS, next generation sequencing; PLSR, partial lease squares regression; QTL, quantitative trait loci; GBLUP, genomic best linear unbiased predictor; RR-BLUP, random regression best linear unbiased prediction; SNP, single nucleotide polymorphism

## Abstract

Genomics and phenomics have promised to revolutionize the field of plant breeding. The integration of these two fields has just begun and is being driven by advances in next-generation sequencing and developments of field-based high-throughput phenotyping (HTP) platforms. To study how a large scale breeding programs may utilize genomics and phenomics, we evaluated three site years of elite yield trials of wheat (*Triticum aestivium*) grown at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico. In each experiment, 1,170 lines were evaluated in drought (2014 and 2015) and heat (2015). Normalized difference vegetation index and canopy temperature were collected on wheat lines via proximal sensing and tagged with precise GPS coordinates using a small, portable phenotyping platform called 'Phenocart'. Genotyping-by-sequencing (GBS) was used for marker discovery and genotyping. Using the 2,101 GBS markers along with over 1.1 million phenotypic observations, several genomic selection (GS) methods were evaluated. The models investigated how to best leverage genetic information along with HTP data for improved GS accuracy. Phenotypic data, whether used as a response in multivariate models or as a covariate in univariate models, consistently improved GS performance. All methods evaluated performed similar in predicting individual environments, but there were large differences within environments. On average, models incorporating phenotypic and genotypic data had a 12% increase in accuracy compared to genetic models alone. Continued advances in GS model performance as well as increasing data generating capabilities for both genomic and phenomic data will make these methods tractable for plant breeders to enhance the rate of genetic gain.

## Introduction

To meet the future global demand for food, researchers need to double crop production by 2050, which is equivalent to a 2.4% yield increase per year (Ray et al., 2013). Genetic gains in wheat yield, however, are currently estimated to be less than

1% per year (Reynolds et al., 2012; Ray et al., 2013); much lower than the 2.4% target. In order to increase genetic gains and maintain a stable food supply plant breeders and geneticist need to utilize contemporary methods to enhance classical breeding. Genomics has long promised to revolutionize plant breeding by characterizing germplasm and allowing individual loci to be dynamically manipulated for crop improvement (Beckmann and Soller, 1986).

Studies have shown that genomics and marker assisted selection (MAS) can be incorporated into breeding programs often resulting in a near 2-fold rate of genetic gain compared to traditional phenotypic selection (Eathington et al., 2007). Next generation sequencing (NGS) techniques have been developed that radically reduce the amount of time for marker discovery, one of the historical challenges of MAS (Xu and Crouch, 2008). Methods such as genotyping-by-sequencing (GBS) (Elshire et al., 2011), reduced representation libraries, and restriction site-associated DNA sequencing can be used to quickly discover thousands of markers and genotype individuals. Additionally, these approaches avoid ascertainment bias that can be present in high-throughput SNP arrays (Davey et al., 2011). Along with providing a high number of markers, NGS methods are also very cost effective for genomic profiling of thousands of breeding lines (Poland and Rife, 2012). In comparison to ever increasing cost associated with field trials (Bernardo and Yu, 2007; Heffner et al., 2010), NGS provides opportunities for crop breeders and geneticist to genotype entire breeding programs (Poland, 2015) and generate dense genomic information that can be used for plant improvement (Morrell et al., 2011).

One method utilizing this wealth of genomic data is genomic selection (GS) first proposed by Meuwissen et al. (2001). Genomic selection simultaneously estimates all marker effects to predict total genetic value. Genomic selection works by using a sufficient quantities of markers to cover the entire target genome so that each quantitative trait loci (QTL) is in linkage disequilibrium with a marker (Goddard and Hayes, 2007). By estimating all marker effects, variation can be captured that may not have identified above a significance threshold using traditional statistical approaches (Meuwissen and Goddard, 2001). Genomic selection has been shown to be extremely accurate in predicting estimated breeding value in relationship to true breeding values in simulation (Heffner et al., 2009), and GS is estimated to be twice as efficient as MAS in winter wheat (Heffner et al., 2010).

Genomic selection requires accurate phenotyping, which has long been the key to enhancing genetic gains through classical plant breeding. Phenotypic observations are collected on the training population that is then used by the GS model for predictions. The phenotypic information is used not to select individual plants *per se*, but to train prediction models and predict the performance of non-phenotyped individuals from their marker scores (Meuwissen et al., 2001). Thus phenotyping plays an essential role in the success of traditional phenotypic selection and genomic selection. However, the ability to assess phenotypes has lagged behind advancements in genomics capabilities (Campos et al., 2004; White et al., 2012; Cobb et al., 2013).

Recently several field-based high-throughput phenotyping (HTP) platforms have been developed to alleviate the phenotyping bottleneck. Some of the platforms have been push carts (White and Conley, 2013; Crain et al., 2016), tractor mounted systems

(Busemeyer et al., 2013a; Andrade-Sanchez et al., 2014) and aerial vehicles (Liebisch et al., 2015). White et al. (2012) and Deery et al. (2014) provide detailed reviews of potential phenotyping platforms and the benefits and challenges associated with each system. While each phenotyping system is varied in its capabilities and cost, they all have the option to provide dense phenotypic data can be used to understand crop growth. Measurements provided by HTP systems have been shown to be highly correlated to manual measurements suggesting high accuracy relative to current approaches and many HTP platforms can measure multiple traits simultaneously (Andrade-Sanchez et al., 2014; Crain et al., 2016). For example, Busemeyer et al. (2013b) mapped temporal genetic dynamics for biomass accumulation using hyperspectral imaging and time of flight cameras.

The addition of HTP platforms provides opportunities to increase genomic selection through enhancing prediction models, but the question persists of how best to incorporate this new information. Currently, most GS models are single trait models, incorporating phenotypic information on only the target trait (Jia and Jannink, 2012). While single trait models have proven to be useful (Heffner et al., 2009, 2011), they do not take advantage of correlations between traits (Jia and Jannink, 2012). Often grain yield is predicted by GS models, but there are a myriad of physiological process that culminate in grain yield (Pask et al., 2012) and well documented cases of physiological phenotypes correlated to yield (Amani, 1996; Gutiérrez-Rodríguez et al., 2004; Babar and Reynolds, 2006). HTP methodologies have been very amenable to measuring traits such as spectral reflectance and canopy temperature (CT) that have potential for use as selection tools because of their high correlations to grain yield (Amani, 1996;

Gutiérrez-Rodríguez et al., 2004; Babar et al., 2006). Multiple-trait genomic selection has been proposed to leverage the shared information between correlated traits (Jia and Jannink, 2012). However, there are few examples in literature where multiple-trait or multivariate GS models have been deployed. Jia and Jannink (2012) showed that prediction accuracies could be increased significantly for traits with low heritability through multivariate GS models that include correlated traits to the trait of interest. Calus and Veerkamp (2011) also found that higher genomic prediction accuracy could be achieved by using multi-trait GS models. Their study found that multi-trait models performed better when there were high genetic correlations between traits resulting in up to a 0.14 increase in model accuracy, even with low genetic correlations they found a small increase in accuracy. While both examples show progress for the multi-trait model, it is difficult to ascertain how the models would perform in a large-scale wheat breeding program, particularly because of the large use of simulated data.

Utilizing the promise of genomic selection combined with HTP measurements, we evaluated how a large breeding program could incorporate GS models and high density HTP data. Specifically, we examined the following: 1) how best to incorporate dynamic HTP data into GS models through comparison of single and multi-trait GS models in abiotic stress environments based on model prediction accuracy, and 2) assess the potential for using high-throughput phenotyping platforms for selection decisions large breeding nurseries.

# Materials and Methods

## Field Trial Design and Management

All field trials were grown during the 2013-2014 and 2014-2015 growing seasons at the Campo Experimental Norman E. Borlaug (CENEB), Sonora, Mexico. The genetic materials were advanced lines in the Elite Yield Trials (EYT) of the CIMMYT bread wheat breeding program. Each year approximately 1100 lines are evaluated in replicated field trials under drought and heat stress conditions simulating mega environments, 4A and 5B respectively. Mega environments are considered substantial land areas that experience the common abiotic and biotic stresses (Braun et al., 1996), with over 13 million hectares globally represented by these two drought and heat stressed environments. Two experiments, or different years, of data were obtained for the drought experiment, and one year of data was collected for the heat environment. The drought environments were sown at optimal planting time, mid-November, and water was provided using drip irrigation with a season total of 180 mm. The heat experiment had a late planting date at the end of February that resulted in temperature stress throughout the life cycle, with irrigation applied sufficiently to prevent water stress.

The drought experiments were grown on a flat planting surface in plots that were 3.5 m x 1.6 m, while the heat environment was grown on raised 80cm beds, with similar plot dimensions. Because of the large number of experimental units, the total experimental area was greater than 2 ha in size for each environment. To manage spatial variability, each experiment area was divided into 39 individual trials (Figure 1). Each individual trial was comprised of 30 genotypes (2 check varieties and 28 entries) in an alpha lattice design with three replicates. For clarity within this study, experiment refers to year and environment combination (3 experiments, 2014 Drought, 2015 Drought, and 2015 Heat) and trials refer to subsets of plots within each experiment that contained three replicates of breeding lines arranged in alpha lattice designs.

### Phenotypic Data Collection

Days to heading and plant height were collected in all plots, with days to maturity only collected in the first replicate as the replications were highly correlated. Grain yield was determined at the end of the growing season using a small plot combine. In addition to hand measured phenotypic data, we used the Phenocart (Crain et al., 2016) to collect NDVI and CT from heading throughout senescence (Table 1). The Phenocart collects georeferenced NDVI and CT data, and after assigning data to plots resulted in an average of 34 measurements per plot for each sample date. To collect data from all of the plots required 5-7 hours depending on walking speed. Because of the time limitation, most often the entire experiment area was taken over the course of two consecutive days for a given time point during the growing season. To efficiently record data, the Phenocart was pushed along columns of plots rather than following individual trial layout (Figure 1). With the GreenSeeker being and active sensor, spectral readings of NDVI change little over the course of the day. Canopy temperature, however, was processed to compensate for ambient temperature changes resulting from the data collection pattern as noted in the data processing section.

### HTP Data Processing and Analysis

For each experiment we collected three to four sets of observations from heading through senescence. For each time of data collection, we processed the data as follows: 1) assigned data to its correct plot using methods similar to Crain *et al.* (2016). 2) calculated heritability for each individual trial for each measurement day, and 3) calculated genotype best linear unbiased predictors (BLUPs). The BLUPs utilize the field-design and help compensate for the large spatial variability as they present a difference from the mean of the trial for each genotype. This compensates for trials that had higher yield based on spatial location in the experiment.

Broad sense heritability ($H^2$) is the ratio of the genetic variance to the phenotypic variance and is often called repeatability (Piepho and Möhring, 2007). A high $H^2$ is indicative of higher precision, and has been related to higher predictive ability for ancillary traits correlated to grain yield (Crain et al. 2016). We computed trait heritability on a trial basis as (Fehr, 1987):

$$Equation\ 1: \quad H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{plot}^2}{r} + \frac{\sigma_{error}^2}{rs}}$$

where $\sigma_g^2$ is the genotypic variance, $\sigma_{plot}^2$ is the variance between plots, and $\sigma_{error}^2$ is the residual variance (within plot variance), r is the number of replications and s is the number of subsamples per plot. For traits measured only once, grain yield, plant height, and days to heading, the model simplifies to:

$$Equation\ 2: \quad H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{error}^2}{r}}$$

which removes the effect for subsampling.

NDVI is relatively static throughout the day, but CT is affected by ambient temperature so we assessed several methods with which to compensate for this potential problem. We tested spatial correction of the raw data by 1) normalized by direction of travel, 2) modeling the data with a time of observation covariate, and 3) modeling data with a date of heading covariate. For each of these correction approaches we assessed heritability of the corrected data and compared to the heritability of the raw recorded data. To calculate BLUPs for CT data to use in GS, we chose the model with CT data normalized by direction of travel. This method provided the highest heritability of CT data as this suggest that environmental variations had been minimized relative to the genotypic differences.

Best linear unbiased predictors for each genotype were calculated by fitting the following mixed model:

$$Equation\ 3\text{:}\ y_{ijkl} = \mu + r_j + b_{k(j)} + g_i + t_{l(ijk)} + \epsilon_{ijkl}$$

where $y_{ijkl}$ is the phenotype observation for the trait of interest, $\mu$ is the overall mean of the population, $g_i$ is the random genotype effect of entry *i* distributed as iid $g_i{\sim}N(0,\sigma_i^2)$, $r_j$ is the random effect for the *j* replication distributed as iid $r_j{\sim}N(0,\sigma_j^2)$, $b_{k(j)}$ is the random effect for the *k* block nested within replication distributed as iid $b_{k(j)}{\sim}N(0,\sigma_k^2)$, $t_{l(ijk)}$ is the *t* subsample nested within *i* genotype, *j* replicate, *k* block and is distributed as iid $t_{l(ijk)}{\sim}N(0,\sigma_l^2)$, and $\epsilon_{ijkl}$ is the residual distributed as iid $\epsilon_{ijkl}{\sim}N(0,\sigma_\epsilon^2)$. For traits measured without subsamples, the term $t_{l(ijk)}$ is removed from the model. Missing phenotypic data was imputed with an expectation maximization (EM) algorithm using the Amelia R package (Honaker et al., 2011). Pearson correlations between BLUPs for HTP traits and grain yield were calculated.

### Genotypic Information

Wheat lines were genotyped using genotyping-by-sequencing (GBS) following protocols by Poland *et al*. (2012). Single nucleotide polymorphisms (SNPs) were called using the TASSEL GBSv2 pipeline (Glaubitz et al., 2014) with Chinese Spring serving as the reference genome (Mayer et al., 2014). Initial SNP calling resulted in 2,079 individuals and 19,583 markers. Individuals with more than 85% missing data were excluded from the final data set. Filtering the genetic loci consisted of: 1) excluding markers with minor allele frequency (MAF) less than 0.05, 2) excluding markers with greater than 5% heterozygosity, 3) excluding markers with more than 30% missing data, 4) excluding markers in complete linkage disequilibrium with another marker. After filtering, the final genotypic marker set included 2,033 individuals and 2,101 markers. We imputed missing markers with an EM algorithm in the R package rrBLUP (Endelman, 2011). The EM imputation has been shown to be computationally efficient and provide increased GS accuracies over mean maker imputation (Rutkoski et al., 2013).

**Genomic Prediction**

We evaluated genomic prediction for grain yield using several statistical methods. A univariate-single trait model (uniGS), a model using only HTP traits as predictors (HTPr), a model with phenotypic covariates (GS+HTP), and a multi-trait model that included grain yield, NDVI and CT for responses (multiGS) were fit for each experiment individually. The GS models were evaluated using partial least squares regression (PLSR), elastic net (EN), and genomic BLUP (GBLUP). The R packages pls (Mevik et al., 2013), glmnet (Friedman et al., 2010), and asreml (Butler, 2009) were used to fit the PLSR, EN, and GBLUP methods respectively. A detailed review of methods for GS is given by Lorenz et al. (2011) and Heslot et al. (2012). The models were fit as a two-step process, first fitting BLUPs for each genotype and trait, and secondly fitting the GS method and model with the calculated BLUPs and genomic markers.

With each GS method and experiment, we fit four different model formulas. The general form that markers and predictors entered the models are: uniGS—A univariate formula for grain yield

$$Equation\ 4: y = \mu + \boldsymbol{Z}u + \varepsilon$$

where $y$ is grain yield and is predicted by μ overall mean and the deviation given from a random marker matrix $\boldsymbol{Z}u$, with $\varepsilon$ as an error term.

GS+HTP—A univariate model for grain yield that included covariate predictor traits.

$$Equation\ 5: y = \mu + \boldsymbol{X}\beta + \boldsymbol{Z}u + \varepsilon$$

where components are equal to Equation 4 with the incorporation of fixed-effect HTP predictor traits ($\boldsymbol{X}\beta$) of NDVI and CT for each measurement date.

MultiGS—A multivariate model predicting grain yield along with HTP traits NDVI and CT.

$$Equation\ 6: \boldsymbol{Y} = \mu + \boldsymbol{Z}u + \varepsilon$$

where ***Y*** is a matrix of grain yield and the HTP traits of NDVI and CT.

HTPr (HTP with multiple **r**egression)—A univariate model predicting grain yield with only the measured HTP traits.

$$Equation\ 7: y = \boldsymbol{X}\beta$$

where HTP traits are fixed effects. This model is similar to the GS+HTP with the removal of genetic marker information.

In addition to fitting individual days of predictor traits, BLUPs for NDVI and CT were averaged across entries as this reduced the computational time in fitting the multivariate models. To evaluate the accuracy of the GS models, we used a cross validation where the data from each experiment was divided by individual trials fit within the model. This division followed from the experimental design and with no filtering represented 39 trials and subsequently 39 cross validation folds. The most closely related lines (e.g. full-sibs) are arranged within the same trial, thus predicting individual trials presents a realistic application of GS when predicting into new sets of breeding lines. If a random sample of genotypes was used for the cross validation, it is likely genetically similar lines would be used to predict line performance possibly biasing the results. Thirty-eight of the divided data sets were used to train the prediction model, and from the trained model prediction were made on the final fold. The cross validation was repeated for each trial predicting all genotypes one time. For data sets where filtering was used the number of folds of cross validation represented the number of trials that were used in the particular model following the same leave one out strategy as presented above. The prediction accuracy was determined by Pearson correlation coefficient between the observed values and the cross-validated genomic estimated breeding value across all folds of cross validation. We did not divide the correlation coefficient by heritability, as that would introduce additional error from the heritability calculation (Heslot et al., 2012).

## Results and Discussion

### Phenotypic Data Collection and Processing

To collect high density phenotypic data, we utilized the Phenocart (Crain et al., 2016) to collect georeferenced NDVI and CT data. Each day of measurements for the experiment area resulted in ~100,000 data observations that were assigned to individual plots using geo-referencing. It took approximately 5-7 hours to cover the entire experiment area at teach time point during the season. Due to the large time difference in data collection as well as the fact that often the experiment area was covered over two consecutive days, we assessed how best to use CT as it is influenced by diurnal environmental conditions. Many of the observation days for CT had clear patterns that were not random and based upon the time of observation (Figure 2a). To mediate this, we examined a number of methods to account for the time variation. We evaluated models accounting for the time of observation, heading date of the genotypes as covariates as well as normalizing the data by column of travel (Figure 1). To assess these correction approaches, we used the average broad-sense heritability of all trials in the experiment. While using time of observation covariates and date of heading improved the models slightly, normalization of CT by column improved the average $H^2$ from to 0.34 to 0.55 (Figure 3) and removed much of the non-random patterns that were observed (Figures 2a and 2b). Also using the normalized data, we observed a negative correlation with grain yield which has been observed in other studies (Balota et al., 2008; Crain et al., 2016), further suggesting the data transformation was appropriate (Figure 4). We used NDVI measurements from each trial without correction. NDVI was collected with an active sensor, which limits effects due to environmental variation

(Solari et al., 2008), and the observed values changed little from one day to the next. The high $H^2$ values from the NDVI data also provide evidence that there was little variation over data collection time compared to CT (average $H^2$ over all trials 0.87 for NDVI compared to 0.34 for uncorrected CT). $H^2$ from each trial and observation date are reported in Supplemental Table S1.

Based on these data, it appears that sensor readings can be standardized to better reflect the true biological information that is present in the field. While we have used a normalization technique, other sensors may require a different type of calibration. Nevertheless, this finding is useful for scientists trying to collect dynamic measurements over large field trials.

## Genomic Prediction

We evaluated several different GS models to identify the best approaches to leverage genomic information and HTP measurements. Because genomic selection can enhance plant breeding by reducing the time per breeding cycle (Meuwissen et al., 2001; Heffner et al., 2009) and HTP data should increase the accuracy of phenotypic selection and training models for GS prediction (Cobb et al., 2013), the combination of GS and HTP should translate to higher genetic gains in the field. We used three distinct models (PLS, EN, GBLUP) for genomic prediction and assessed their accuracy using cross validation with accuracy reported as the Pearson correlation coefficient between BLUP for grain yield and the predicted value for grain yield.

The $H^2$ values varied widely among traits and days of observations, so we explored the effect of filtering data based on $H^2$ values in the Heat 2015 experiment. If an individual trial had $H^2$ below 0.25 for CT or NDVI the entire trial was removed from the data set, resulting in a data set of 22 of the original 39 trials with complete phenotypic observations.  We also removed CT Observations dates 1 and 4 as the majority of trials had low $H^2$ for these dates.  While this reduced the size of the data set, we wanted to evaluate the effect that only choosing high heritability data would have on the predictions.  To evaluate if the changes in prediction accuracy came from filtering for higher heritability or from decreasing the training population size, we fit the same models using a random selection of 22 trials.  Overall, we found that filtering based on $H^2$ values made little changes in the prediction accuracy (Figure 5).  Across all methods, there was not a clear trend that filtering resulted in higher performing models.  Based on our finding of limited effect for filtering phenomic data on the model performance, we fit all data for each experiment.

**Genomic Prediction Accuracy**

Across the three environments there were large differences in prediction accuracy among and within environments (Figure 6).  Overall, environments that had low heritability for grain yield (Drought 2015) had worse prediction performance than environments with higher heritability for grain yield (Drought 2014 and Heat 2015). Across environments, models tended to perform similarly (Table 2)  While the PLSR and EN models were computationally efficient, the GBLUP methods required much more computation time, thus we fit each model with all of the HTP traits (6 to 8 HTP observation time points) and the same models with the average of the HTP traits.  Using the averaged HTP traits resulted in predictions that were not statistically different to models using individual HTP traits (Table 3), except for a few of the HTPr models.

The highest prediction accuracy came from models using only HTP traits as predictors (HTPr) models (Figure 6). Across environments, the models displayed several trends. Partial least squares regression models usually had the lowest prediction accuracy. With the exception of PLSR HTPr models, the addition of HTP traits had no significant impact on prediction accuracy within these models. Elastic net models tended to have the highest performance with the HTPr and GS+HTP models. Using the HTP data for covariates performed almost as well as EN HTPr models. The GBLUP models performed the best when using multiGS compared to other models. The multiGS GBLUP model achieved accuracies almost comparable to HTPr trait only models (Figure 6). This is significant as the GBLUP model using the covariance matrix between traits aids in predictive power almost comparable to utilizing HTP traits alone.

The observed trend that HTPr generally outperformed the other models (UniGS, GS+HTP, and multiGS) (Figure 6) may be due to the assumptions that each model inherently makes and how the models utilize the data. For example, partial least squares regression is a model that extracts latent variables and is useful when there are a number of variables and their relationships are not clearly understood (Tobias, 1995). Latent variables are extracted as linear combinations from the original data sets (both predictors and response) and the latent variables are chosen to maximize the predicted response (Lorenz et al., 2011). Within our data set, we have variables which are expected to have relationships to yield from previous research (Amani, 1996; Gutiérrez-Rodríguez et al., 2004; Babar and Reynolds, 2006), and the GS method like PLSR may be identifying these known associations. Elastic net penalizes the models when the number of predictors (marker set) is much larger than the number of observations (genotypes). It functions by both selecting variables (markers are dropped from the model) as well as imposing shrinkage on the variables that remain in the model (Zou, 2005). Using HTP variables, would allow the EN method to single out a few predictors and drop other terms (markers) while still maintaining or improving model performance over the uniGS models. In both GS+HTP and HTPr models, the EN models could effectively select variables related to HTP that are predicting yield and result in the nearly identical performance between the two models (Figure 6). The GBLUP methods are a popular statistical technique to perform GS using random regression best unbiased prediction (RR-BLUP) (Lorenz et al., 2011). In RR-BLUP, the typical least squares estimates are shrunken toward zero using a penalty term that reduces collinearity between predictors but keeps all predictors (Lorenz et al., 2011). Keeping all predictors could account for the difference in EN and GBLUP model performance for the HTP+GS methods (Figure 6). The implicit assumptions that each method imposes on the data also assert practical limits on the utility of the methods as elaborated in the model assessment section.

Compared to the standard GS models, the addition of HTP traits had a large increase in efficiency. Across all models that utilized genomic information (GS+HTP and multiGS), adding HTP traits resulted in 12% average increase compared to marker (uniGS) selection alone. While the HTPr models did not rely on genetic information, they averaged the highest increase of 1.5 fold greater than GS. Regardless of how a breeding program chooses to use HTP information, this approaches seems very effective in increasing selection accuracy.

## Model Assessment

Genomic selection has the potential to revolutionize the practical aspects of plant breeding programs (Desta and Ortiz, 2014). Determining how best to utilize this information in concert is needed to allow breeders to maximize resources and achieve maximum genetic gains. In this study, we have examined several possibilities to incorporate HTP information into genomic selection models, based on our future vision that breeding programs will have genomic profiles available for all breeding lines and that high-throughput phenotyping will become routine. Our findings show that HTP traits consistently improve model performance over univariate GS models alone. Elastic net and HTPr models were often the highest in terms of prediction accuracy. Due to computational constraints we only fit the mulitGS model with HTP data that had been averaged across all days; however, its prediction was equivalent to HTPr models. Application of dimensional reduction techniques such as matrix decomposition (Mrode, 2014) may allow for faster model fitting and for the addition of more traits given that the number of effects (variance and covariance parameters) rise linearly with the addition of more traits and observation time points. As phenotyping programs expand the number of traits measured and the frequency of measurements making efficient computational use of the data will be an active area of research. Currently, the models selected by a breeding program may depend on the use of the predictions. For a breeding program HTPr models may work well for decisions in advancing lines to further generations while providing fast computation. The application of HTPr models is limited to field trials as the model provides no information about the underlying genetic architecture and can only be assessed on the same environment that the line *per se* is evaluated. However, if a breeder wanted to select lines to add in the crossing block GBLUP methods may be more appropriate in representing the potential breeding value based on genomic composition.

## Conclusions

Sustaining food production into the future will be challenging with expected population increases and limited availability of land and resources (Hawkesford et al., 2013). To meet the expected increase in demand, a marked increase in genetic gain is needed. We investigated several GS methods and the how best to model the phenotypic information. Utilizing HTP traits improved model performance, thus enabling more accurate selection of superior breeding lines from larger populations. The advances in genotyping combined with efficient phenotyping platforms will continue to push the integration of genomics and phenomics for breeding and genetics. By utilizing both genetic information and phenotypic data, it is expected that breeders will be able to better connect genotype to phenotype while more efficiently identifying and selecting superior higher yielding crop varieties.

## Acknowledgements

This work was supported through the National Science Foundation - Plant Genome Research Program (IOS-1238187) and the US Agency for International Development (USAID) Feed the Future Innovation Lab for Applied Wheat Genomics (Cooperative Agreement No. AID-OAA-A-13-00051). Field trials at CIMMYT, Mexico were supported by the Durable Rust Resistance in Wheat project funded by Bill and Melinda Gates Foundation (BMGF) and the British Department of Foreign aid for International Development (DFID), Cereal Systems Initiative for South Asia Project funded by BMGF and USAID, and Grain Research and Development Corporation (GRDC), Australia. J. Crain is supported through a fellowship from the Monsanto Beachell-Borlaug International Scholars program.

## References

Amani, I. 1996. Canopy temperature depression association with yield of irrigated spring wheat cultivars in a hot climate. J. Agron. Crop Sci. 129: 119–129.

Andrade-Sanchez, P., M.A. Gore, J.T. Heun, K.R. Thorp, A.E. Carmo-Silva, A.N. French, M.E. Salvucci, J.W. White, and E.A. Carmo-Silva. 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Funct. Plant Biol. 41(2007): 68–79.

Babar, M. a., M. Van Ginkel, a. R. Klatt, B. Prasad, and M.P. Reynolds. 2006. The potential of using spectral reflectance indices to estimate yield in wheat grown under reduced irrigation. Euphytica 150(1-2): 155–172.

Babar, M., and M. Reynolds. 2006. Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. Crop Sci. 46(3): 1046–1057.

Balota, M., W. a. W. Payne, S.S.R. Evett, and T.R. Peters. 2008. Morphological and physiological traits associated with canopy temperature depression in three closely related wheat lines. Crop Sci. 48(5): 1897–1910.

Beckmann, J.S., and M. Soller. 1986. Restriction fragment length polymorphism and genetic improvement of agricultural species. Euphytica 35(1): 111–124.

Bernardo, R., and J. Yu. 2007. Prospects for Genomewide Selection for Quantitative Traits in Maize. Crop Sci. 47(3): 1082.

Braun, H.-J., S. Rajaram, and M. Ginkel. 1996. CIMMYT's approach to breeding for wide adaptation. Euphytica 92(1-2): 175–183.

Busemeyer, L., D. Mentrup, K. Möller, E. Wunder, K. Alheit, V. Hahn, H.P. Maurer, J.C. Reif, T. Würschum, J. Müller, F. Rahe, and A. Ruckelshausen. 2013a. BreedVision--a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors (Basel). 13(3): 2830–47.

Busemeyer, L., A. Ruckelshausen, K. Möller, A.E. Melchinger, K. V Alheit, H.P. Maurer, V. Hahn, E. a Weissmann, J.C. Reif, and T. Würschum. 2013b. Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. Sci. Rep. 3: 2442.

Butler, D. 2009. asreml: asreml() fits the linear mixed model.

Calus, M.P., and R.F. Veerkamp. 2011. Accuracy of multi-trait genomic selection using different methods. Genet. Sel. Evol. 43(1): 26.

Campos, H., M. Cooper, J.E. Habben, G.O. Edmeades, and J.R. Schussler. 2004. Improving drought tolerance in maize: a view from industry. F. Crop. Res. 90(1): 19–34.

Cobb, J.N., G. DeClerck, A. Greenberg, R. Clark, and S. McCouch. 2013. Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126(4): 867–887.

Crain, J.L., Y. Wei, J. Barker III, S.M. Thompson, P.D. Alderman, M. Reynolds, N. Zhang, and J. Poland. 2016. Development and deployment of a portable field phenotyping platform. Crop Sci.

Davey, J., P. Hohenlohe, and P. Etter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. … 12(7): 499–510.

Deery, D., J. Jimenez-Berni, H. Jones, X. Sirault, and R. Furbank. 2014. Proximal Remote Sensing Buggies and Potential Applications for Field-Based Phenotyping.

Desta, Z.A., and R. Ortiz. 2014. Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci. 19(9): 592–601.

Eathington, S.R., T.M. Crosbie, M.D. Edwards, R.S. Reiter, and J.K. Bull. 2007. Molecular markers in a commercial breeding program. Crop Sci. 47(SUPPL. DEC.).

Elshire, R.J., J.C. Glaubitz, Q. Sun, J. a Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6(5): e19379.

Endelman, J.B. 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome J. 4(3): 250.

Fehr, W.R. 1987. Principles of cultivar development: theory and technique. New York.

Friedman, J., T. Hastie, and R. Tibshirani. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 33(1): 1–22.

Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, and E.S. Buckler. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS One 9(2): e90346.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. J. Anim. Breed. Genet. 124(6): 323–30.

Gutiérrez-Rodríguez, M., M.P. Reynolds, J.A. Escalante-Estrada, and M.T. Rodríguez-González. 2004. Association between canopy reflectance indices and yield and physiological traits in bread wheat under drought and well-irrigated conditions. Aust. J. Agric. Res. 55(11): 1139–1147.

Hawkesford, M.J., J.L. Araus, R. Park, D. Calderini, D. Miralles, T. Shen, J. Zhang, and M. a. J. Parry. 2013. Prospects of Doubling Global Wheat Yields. Food Energy Secur. 2: 34–48.

Heffner, E., J. Jannink, and M. Sorrells. 2011. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Genome: 65–75.

Heffner, E.L., A.J. Lorenz, J.L. Jannink, and M.E. Sorrells. 2010. Plant breeding with Genomic selection: Gain per unit time and cost. Crop Sci. 50(5): 1681–1690.

Heffner, E.L., M.E. Sorrells, and J.-L. Jannink. 2009. Genomic Selection for Crop Improvement. Crop Sci. 49(1): 1.

Heslot, N., H.-P. Yang, M.E. Sorrells, and J.-L. Jannink. 2012. Genomic selection in plant breeding: A comparison of models. Crop Sci. 52(1): 146–160.

Honaker, J., G. King, and M. Blackwell. 2011. {Amelia II}: A Program for Missing Data. J. Stat. Softw. 45(7): 1–47.

Jia, Y., and J.-L. Jannink. 2012. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics 192(4): 1513–22.

Liebisch, F., N. Kirchgessner, D. Schneider, A. Walter, and A. Hund. 2015. Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. Plant Methods 11(1).

Lorenz, A., S. Chao, and F. Asoro. 2011. 2 Genomic Selection in Plant Breeding: Knowledge and Prospects. Adv. … 110.

Mayer, K.F.X., J. Rogers, J. Doležel, C. Pozniak, K. Eversole, C. Feuillet, B. Gill, B. Friebe, A.J. Lukaszewski, P. Sourdille, and others. 2014. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science (80-. ). 345(6194): 1251788.

Meuwissen, T.H., and M.E. Goddard. 2001. Prediction of identity by descent probabilities from marker-haplotypes. Genet. Sel. Evol. 33(6): 605–34.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4): 1819–1829.

Mevik, B.-H., R. Wehrens, and K.H. Liland. 2013. pls: Partial Least Squares and Principal Component regression.

Morrell, P.L., E.S. Buckler, and J. Ross-Ibarra. 2011. Crop genomics: advances and applications. Nat. Rev. Genet. 13(2): 85–96.

Mrode, R.A. 2014. Linear Models for the Prediction of Animal Breeding Values: 3rd Edition.

Pask, A., J. Pietragalla, D. Mullan, and M. Reynolds. 2012. Physiological breeding II a field guide to wheat phenotyping. Mexico City, Mexico.

Piepho, H.-P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177(3): 1881–8.

Poland, J. 2015. Breeding-assisted genomics. Curr. Opin. Plant Biol. 24: 119–124.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.-L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2): e32253.

Poland, J. a., and T.W. Rife. 2012. Genotyping-by-Sequencing for Plant Breeding and Genetics. Plant Genome J. 5(3): 92.

Ray, D.K., N.D. Mueller, P.C. West, and J. a Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One 8(6): e66428.

Reynolds, M., J. Foulkes, R. Furbank, S. Griffiths, J. King, E. Murchie, M. Parry, and G. Slafer. 2012. Achieving yield gains in wheat. Plant, Cell Environ. 35(10): 1799–1823.

Rutkoski, J.E., J. Poland, J.-L. Jannink, and M.E. Sorrells. 2013. Imputation of unordered markers and the impact on genomic selection accuracy. G3 (Bethesda). 3(3): 427–39.

Solari, F., J. Shanahan, R. Ferguson, J. Schepers, and A. Gitelson. 2008. Active sensor reflectance measurements of corn nitrogen status and yield potential. Agron. J. 100(3): 571–579.

Tobias, R.D. 1995. An Introduction to Partial Least Squares Regression. Proc. Ann. SAS Users Gr. Int. Conf. 20th: 1250–1257.

White, J., P. Andrade-Sanchez, M.A. Gore, K.F. Bronson, T.A. Coffelt, M.M. Conley, K.A. Feldmann, A.N. French, J.T. Heun, D.J. Hunsaker, M.A. Jenks, B.A. Kimball, R.L. Roth, R.J. Strand, K.R. Thorp, G.W. Wall, and G. Wang. 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133: 101–112.

White, J.W., and M.M. Conley. 2013. A Flexible, Low-Cost Cart for Proximal Sensing. Crop Sci. 53(4): 1646–1649.

Xu, Y., and J.H. Crouch. 2008. Marker-Assisted Selection in Plant Breeding: From Publications to Practice. Crop Sci. 48(2): 391.

Zou, H. 2005. Regularization and Variable Selection via the Elastic Net. : 301–320.

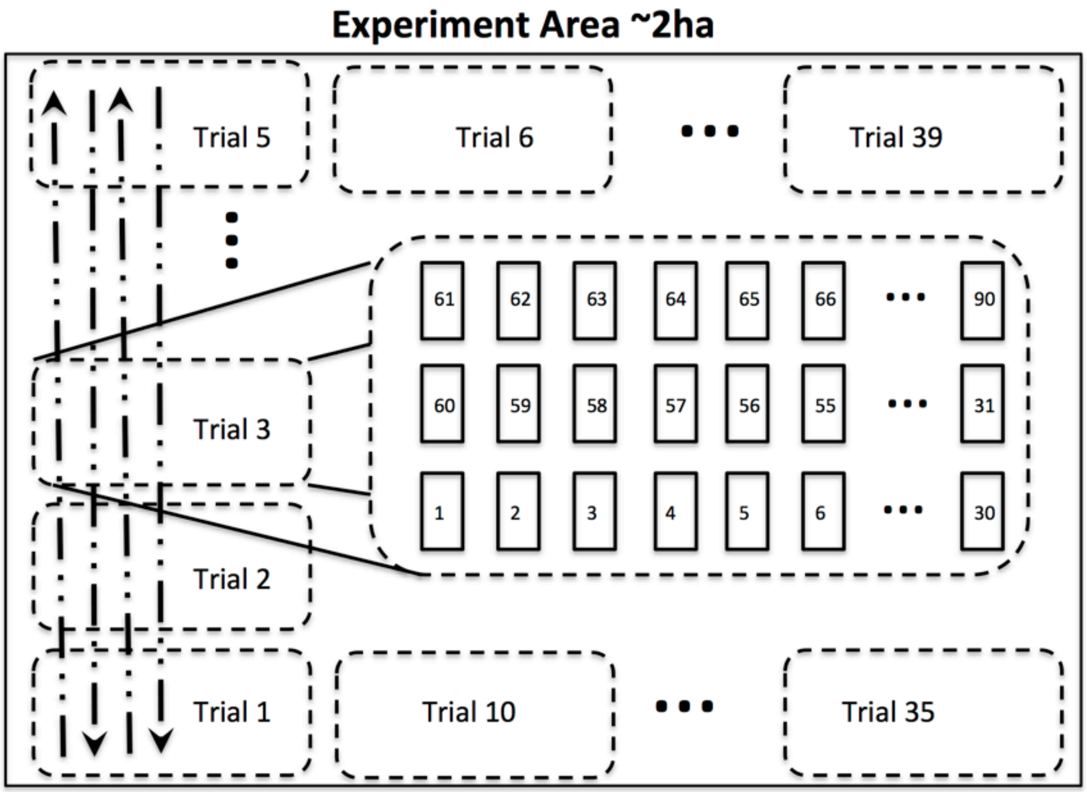


**Figure 4-1. Example field layout and design of each experiment area (2014 and 2015 Drought, and 2015 Heat). Each experiment area was approximately 2 ha in size which was subdivided into 39 individual trials that had 30 unique entries in an alpha lattice design. Broken lines with arrows in the first 5 trials represent the pattern of data collection, with CT data normalized based on the column pass of all trials.**

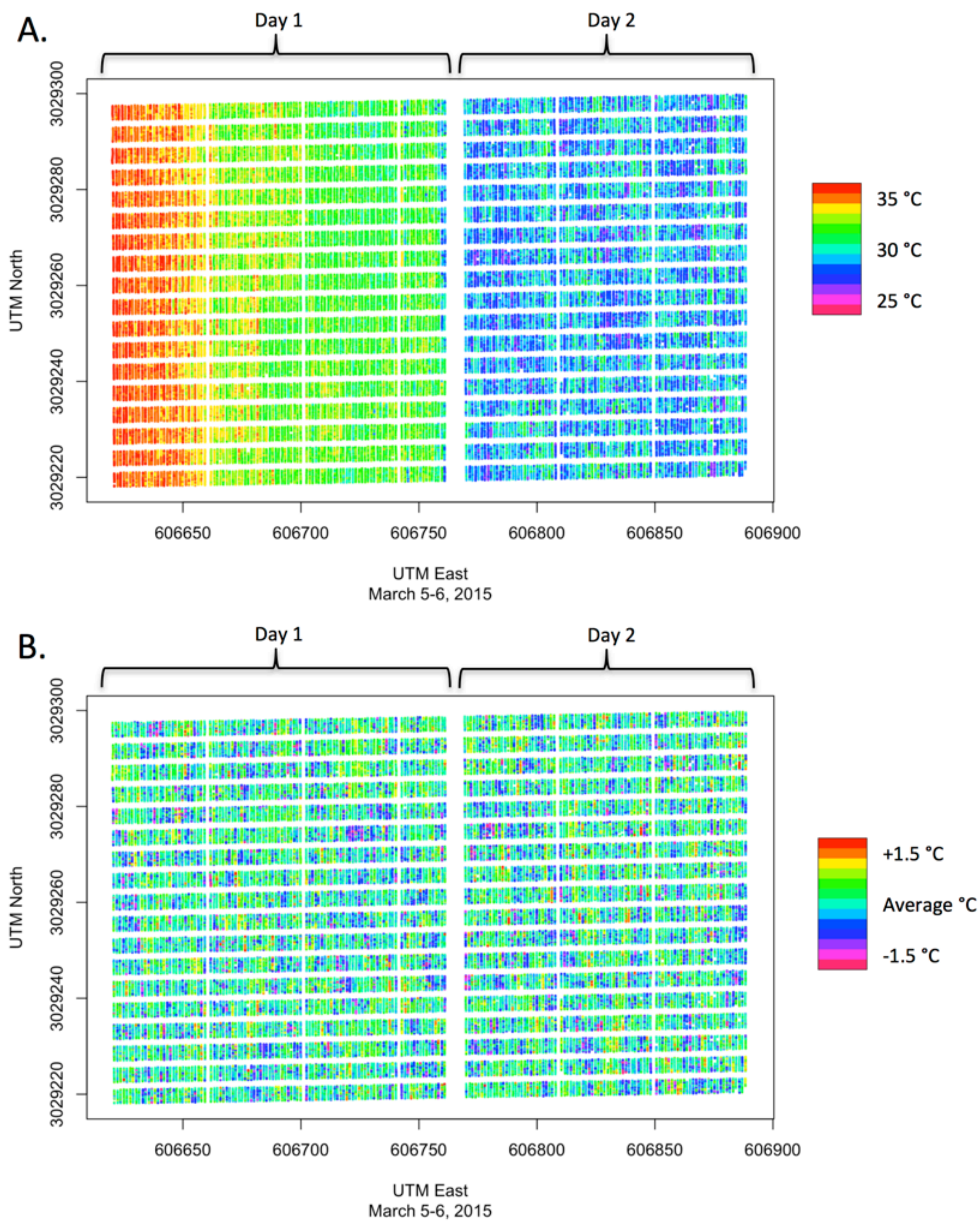


**Figure 4-2. Canopy temperature (CT) data assigned to plot that is color coded by value of 3,510 plots. Panel A shows the uncorrected CT data that was obtained collecting data column by column (Figure 1). The data collection was over two days, and there are clear differences in the right and left portions showing environmental differences due to day as well as daily changes. Day 1 shows non random measurements with an increasing temperature progression from left to right. Panel B is the temperature differences normalized by column. The inter and intra daily gradients have been effectively removed, resulting in gradients from top to bottom that are related to irrigation patterns. Along with the minor spatial patterns, there is also differences among genotypes.**

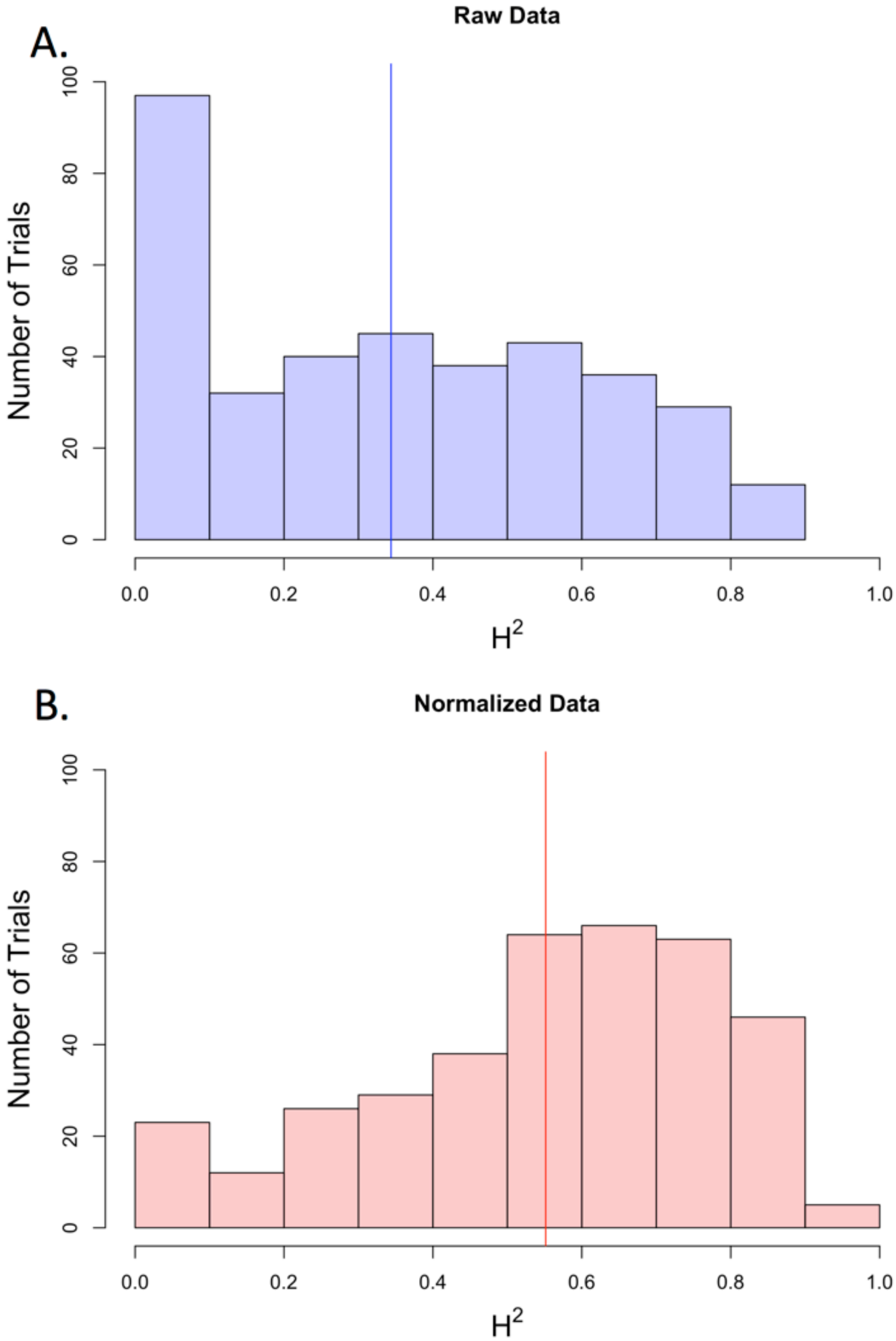


**Figure 4-3. Broad sense heritability (*$H^2$*) for canopy temperature (CT) for each individual trial. Unprocessed data is shown in blue with a vertical line at 0.34 representing the mean, Panel A. Normalizing the data resulted in much higher *$H^2$* values with a mean of 0.55 shown by the red vertical line, Panel B.**

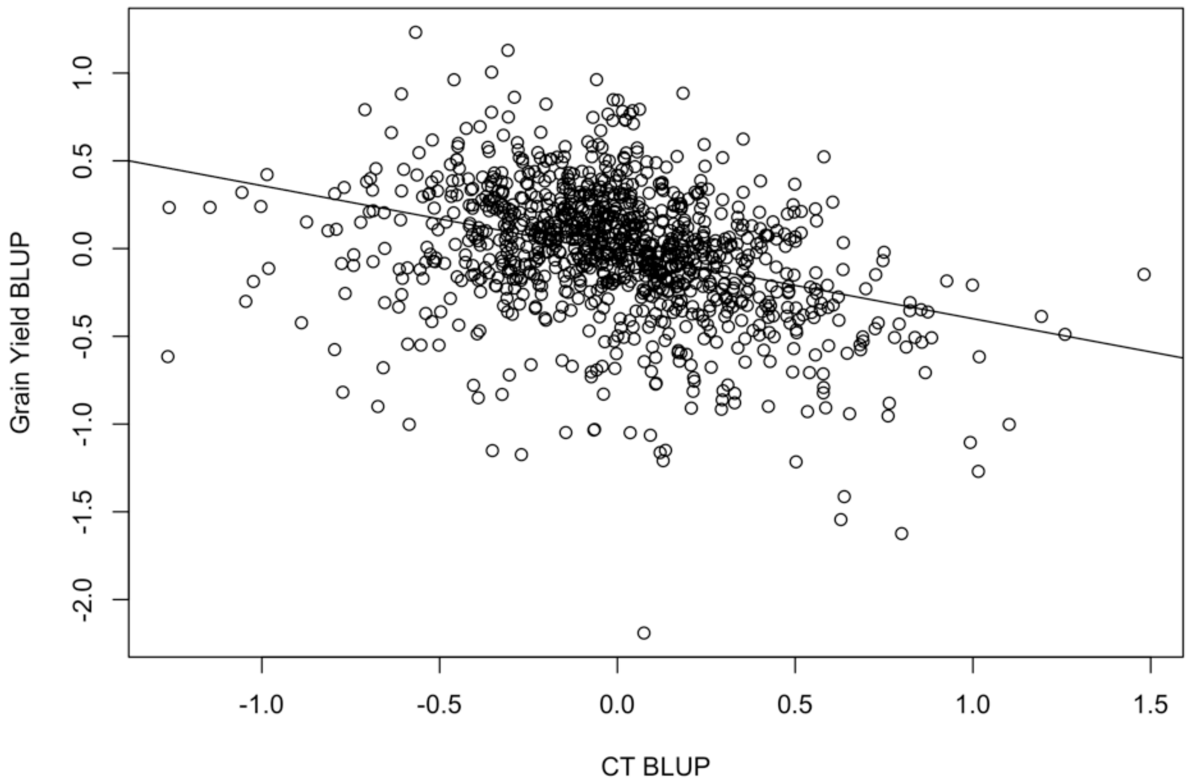


**Figure 4-4. Pearson relationship between best linear unbiased predictors (BLUPs) for canopy temperature (CT) and grain yield using BLUPs for CT that had been normalized by column of data collection. Trend line is line of best fit (*r* = -0.35).**

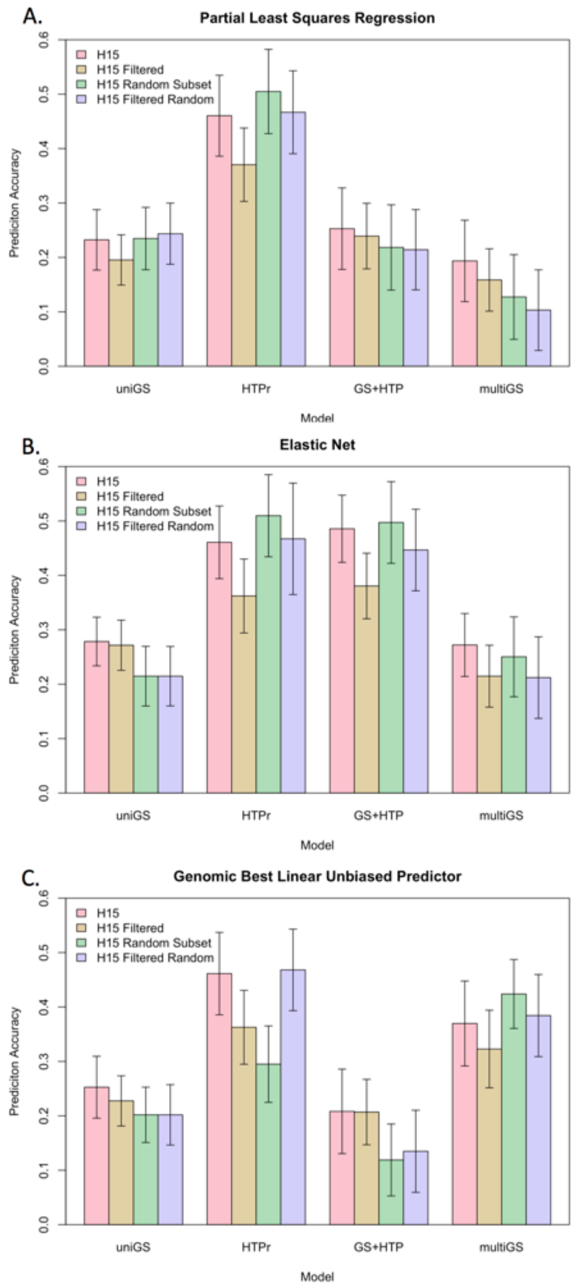


**Figure 4-5. Performance of genomic selection (GS) methods and models filtering on heritability of high-throughput phenotyping (HTP) traits. Each panel shows a particular GS method, partial least squares regression (PLSR), elastic net (EN), and best linear unbiased predictor, panels A, B, and C respectively. The models were fit to four sets of data from the Heat 2015 trial, all data (H15), a filtered data set based on heritability of HTP traits (H15 Filtered), and a random subset that was filtered and unfiltered that matched the size of the H15 Filtered data set (H15 Filtered Subset and H15 Random Subset respectively). uniGS is genomic selection only with no phenotypic data, GS+HTP uses HTP as covariates, multiGS uses HTP data as responses, and HTPr is prediction only with HTP data. Error bars represent 95% confidence intervals.**

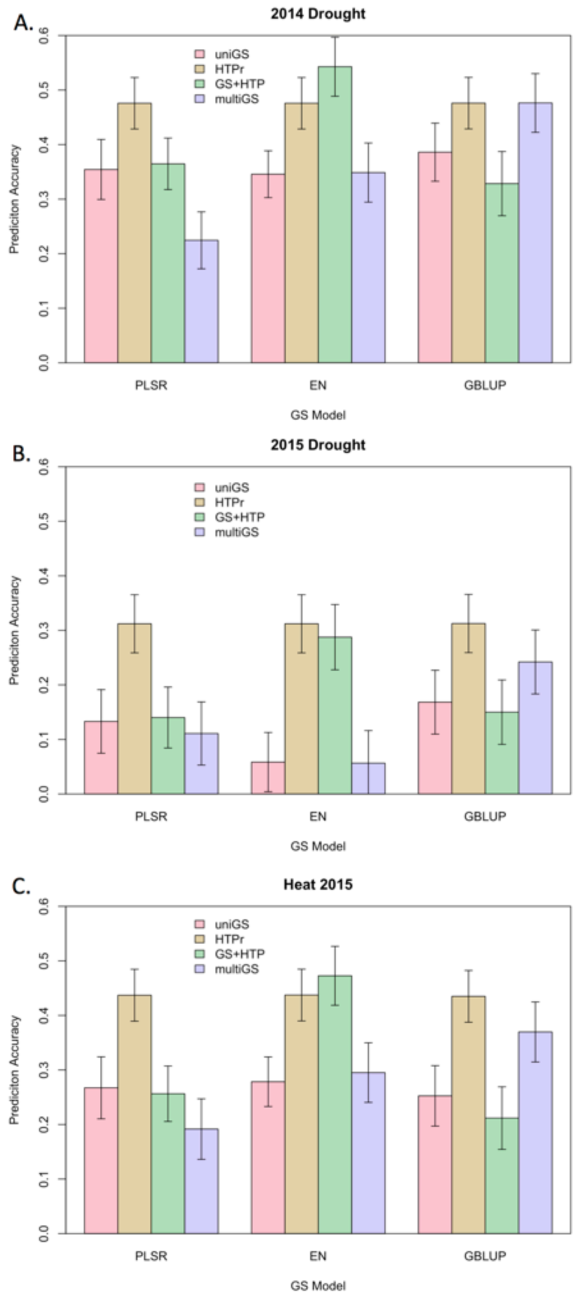


**Figure 4-6. Performance of genomic selection (GS) across three experiments. Each panel represents one experiment area with the three different GS methods and four GS models. Methods are grouped on the x-axis by type of statistical method, BLUP best linear unbiased predictor, EN elastic net, and PLSR is partial least squares regression. Four model formulations are only genetic data (uniGS), genetic data and phenotypic data (GS+HTP), multi response for grain yield and HTP (multiGS), and phenotypic traits only (HTPr). Model accuracy are given on the y-axis in terms of the correlation coefficient between best linear unbiased predictor for grain yield and the genomic estimated breeding value from model prediction with error bars representing 95% confidence interval.**

**Table 4-1. Days of phenotypic observation, Ciudad Obregon, Mexico, for normalized difference vegetation index (NDVI) and canopy temperature (CT) for each experiment along with the total number of collected data points and average number of data per plot. The average number is sum of NDVI and CT data combined.**

| Experiment | Type of observation | 1st Observation | 2nd Observation | 3rd Observation | 4th Observation | Total Observations |
|---|---|---|---|---|---|---|
| 2014-Drought | Date | Feb. 15-16 | Feb. 22-23 | Feb. 28-Mar. 1 | NA | |
| | Number of data | 109545 | 107517 | 134955 | | 352017 |
| | Average data per plot | 31 | 37 | 38 | | |
| 2015-Drought | Date | Feb. 20-21 | Feb. 26 | Mar. 5-6 | NA | |
| | Number of data | 110699 | 142795 | 140490 | | 393984 |
| | Average data per plot | 37 | 41 | 40 | | |
| 2015-Heat | Date | Apr. 23-24 | Apr. 30 | May 7 | May 18-19 | |
| | Number of data | 95223 | 97857 | 98732 | 101838 | 393650 |
| | Average data per plot | 27 | 28 | 28 | 29 | |

**Table 4-2. Genomic prediction and 95% confidence intervals (CI) for prediction accuracies for three different genomic selection (GS) methods in three experiments. Accuracy is given as Pearson correlation between observed best linear unbiased predictor (BLUP) for grain yield and genomic estimated breeding value (GEBV) for grain yield. uniGS is a univariate model with only genetic markers, GS+HTP is a univariate model for grain yield with high-throughput phenotyping traits as covariates, multiGS is a multi response GS model with both grain yield and high throughput-phenotyping traits as responses, HTPr is a univariate model for grain yield predicted only with high-throughput phenotyping traits. MultiGS model for BLUP is not fit based on computational constraints.**

| | | 2015 Heat | | 2015 Drought | | 2014 Drought | | | |
|---|---|---|---|---|---|---|---|---|---|
| Method | Model | *r* | 95% CI | *r* | 95% CI | *r* | 95% CI | Avg. Accuracy | Increase from HTP traits |
| Partial Least Squares Regression | uniGS | 0.24 | 0.18-0.3 | 0.14 | 0.08-0.21 | 0.34 | 0.29-0.4 | 0.24 | 1 |
| | GS+HTP | 0.27 | 0.21-0.33 | 0.14 | 0.08-0.2 | 0.35 | 0.3-0.41 | 0.25 | 1.04 |
| | multiGS | 0.19 | 0.14-0.25 | 0.12 | 0.06-0.18 | 0.23 | 0.17-0.29 | 0.18 | 0.75 |
| | HTPr | 0.46 | 0.41-0.51 | 0.43 | 0.38-0.48 | 0.52 | 0.47-0.56 | 0.47 | 1.96 |
| Elastic Net | uniGS | 0.28 | 0.22-0.33 | 0.06 | 0-0.12 | 0.35 | 0.29-0.4 | 0.23 | 1 |
| | GS+HTP | 0.49 | 0.44-0.53 | 0.34 | 0.29-0.39 | 0.55 | 0.51-0.59 | 0.46 | 2 |
| | multiGS | 0.27 | 0.22-0.33 | 0.11 | 0.05-0.16 | 0.35 | 0.3-0.41 | 0.24 | 1.04 |
| | HTPr | 0.46 | 0.41-0.51 | 0.43 | 0.38-0.48 | 0.52 | 0.47-0.56 | 0.47 | 2.04 |
| GBLUP | uniGS | 0.25 | 0.2-0.31 | 0.17 | 0.11-0.23 | 0.39 | 0.33-0.44 | 0.27 | 1 |
| | GS+HTP | 0.21 | 0.15-0.27 | 0.11 | 0.05-0.17 | 0.32 | 0.27-0.38 | 0.21 | 0.78 |
| | multiGS | NA | NA | NA | NA | NA | NA | NA | NA |
| | HTPr | 0.46 | 0.41-0.51 | 0.43 | 0.38-0.48 | 0.52 | 0.47-0.56 | 0.47 | 1.74 |

**Table 4-3. Genomic prediction and 95% confidence intervals (CI) for prediction accuracies for three different genomic selection (GS) methods in three experiments. Accuracy is given as Pearson correlation between observed best linear unbiased predictor (BLUP) for grain yield and genomic estimated breeding value (GEBV) for grain yield. uniGS is a univariate model with only genetic markers, GS+HTP is a univariate model for grain yield with high-throughput phenotyping traits as covariates, multiGS is a multi response GS model with both grain yield and high throughput-phenotyping traits as responses, HTPr is a univariate model for grain yield predicted only with high-throughput phenotyping traits. All methods and models were fit with average NDVI and CT values from all dates of observations.**

| | | 2015 Heat | | 2015 Drought | | 2014 Drought | | | Increase |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Avg. | from HTP |
| Method | Model | *r* | 95% CI | *r* | 95% CI | *r* | 95% CI | Accuracy | traits |
| Partial Least Squares Regression | uniGS | 0.27 | 0.21-0.32 | 0.13 | 0.07-0.19 | 0.35 | 0.3-0.41 | 0.25 | 1 |
| | GS+HTP | 0.26 | 0.2-0.31 | 0.14 | 0.08-0.2 | 0.36 | 0.31-0.42 | 0.25 | 1 |
| | multiGS | 0.19 | 0.13-0.25 | 0.11 | 0.05-0.17 | 0.22 | 0.16-0.28 | 0.17 | 0.68 |
| | HTPr | 0.44 | 0.39-0.48 | 0.31 | 0.26-0.37 | 0.48 | 0.43-0.52 | 0.41 | 1.64 |
| Elastic Net | uniGS | 0.28 | 0.22-0.33 | 0.06 | 0-0.12 | 0.35 | 0.29-0.4 | 0.23 | 1 |
| | GS+HTP | 0.47 | 0.42-0.52 | 0.29 | 0.23-0.34 | 0.54 | 0.5-0.59 | 0.43 | 1.87 |
| | multiGS | 0.3 | 0.24-0.35 | 0.06 | 0-0.12 | 0.35 | 0.29-0.4 | 0.24 | 1.04 |
| | HTPr | 0.44 | 0.39-0.48 | 0.31 | 0.26-0.37 | 0.48 | 0.43-0.52 | 0.41 | 1.78 |
| GBLUP | uniGS | 0.25 | 0.2-0.31 | 0.17 | 0.11-0.23 | 0.39 | 0.33-0.44 | 0.27 | 1 |
| | GS+HTP | 0.21 | 0.15-0.27 | 0.15 | 0.09-0.21 | 0.33 | 0.27-0.38 | 0.23 | 0.85 |
| | multiGS | 0.37 | 0.32-0.42 | 0.24 | 0.18-0.30 | 0.48 | 0.43-0.52 | 0.36 | 1.33 |
| | HTPr | 0.43 | 0.38-0.48 | 0.31 | 0.26-0.37 | 0.48 | 0.43-0.52 | 0.41 | 1.52 |

# Chapter 5 - Conclusions

Plant breeding is a cornerstone to providing sufficient food, fiber, and fuel to meet the world's demand. Utilizing new technologies to gain efficiencies is imperative to meeting this ever-increasing demand. The fields of genomics and phenomics have advanced at an outstanding pace, and the integration of these two disciplines is in its infancy. When work began on developing the Phenocart (late 2012), there were just a handful of reviews (Tester and Langridge, 2010; Houle et al., 2010; White et al., 2012) and few available platforms (Montes et al., 2011; Busemeyer et al., 2013), additionally much of what was available would be cost prohibitive for developing countries. Today, phenotyping platforms are more common, including the Phenocart, tractor mounted and pulled platforms (Busemeyer et al., 2013; Andrade-Sanchez et al., 2014) to zeppelins (Liebisch et al., 2015), but data collection, management, analysis, and data curation remain challenging. This dissertation addresses many concerns in connecting the genotype to the phenotype starting at phenotyping platform development, to deployment and utilization, throughout the entire data cycle resulting in enhanced genomic predictions.

While enhancing genomic selection models and increasing the rate of genetic gain has been the ideal, this work has produced a number of notable findings. An affordable and high portable phenotyping platform has been developed. The software controlling the system is freely available allowing anyone to clone a platform for his or her use. Additionally, the software is being rewritten in python, a freely available scripting language. This will further allow new users to build their own platforms as well as modify the platforms with any sensors that they desire. Along with platform development, there has also been extensive work in developing efficient methods to handle the data. One method of georeferencing plots was presented within this dissertation, but solving problems of locating plots and assigning data has led to further database and algorithm development. Another data method was that of normalizing large amounts of canopy temperature data to reduce environmental noise while still maintaining biological relevance. Finally, utilizing the Phenocart allowed users to generate much more data than is normally captured in a breeding program. Much of this data was correlated directly to yield, but was also modeled temporally through mathematical functions describing plant growth. The parameters within these functions can also be utilized for genomic selection and association mapping.

The pinnacle of this work integrated both the phenomic data from the Phenocart with next generation sequencing data into genomic selection models. Utilizing phenotypic data resulted in genomic selection models that were on average 12% more efficient than models on using genotypic data. This synergy will allow for a faster dissection of the genotype to phenotype problem and most importantly allow plant breeders and geneticists to meet the demands of a growing population.

# Chapter 6 - References

This dissertation has been compiled with three main parts, Chapters 2-4. While those chapters have been formatted as independent, self-contained articles, other references were used for the introduction and conclusion. The references included in this section complete those dependent sections.


Andrade-Sanchez, P., M.A. Gore, J.T. Heun, K.R. Thorp, A.E. Carmo-Silva, A.N. French, M.E. Salvucci, J.W. White, and E.A. Carmo-Silva. 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Funct. Plant Biol. 41(2007): 68–79.

Arruda, M.P., P. Brown, G. Brown-Guedira, A.M. Krill, C. Thurber, K.R. Merrill, B.J. Foresman, and F.L. Kolb. 2016. Genome-Wide Association Mapping of Fusarium Head Blight Resistance in Wheat using Genotyping-by-Sequencing. Plant Genome.

Babcok, E.B. 1940. The chronology of hope wheat. J. Hered. 31(3): 132–133.

Beckmann, J.S., and M. Soller. 1986. Restriction fragment length polymorphism and genetic improvement of agricultural species. Euphytica 35(1): 111–124.

Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: Learning from the last 20 years. Crop Sci. 48(5): 1649–1664.

Borlaug, N. 2007. Sixty-two years of fighting hunger: personal recollections. Euphytica 157(3): 287–297.

Busemeyer, L., D. Mentrup, K. Möller, E. Wunder, K. Alheit, V. Hahn, H.P. Maurer, J.C. Reif, T. Würschum, J. Müller, F. Rahe, and A. Ruckelshausen. 2013. BreedVision--a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors (Basel). 13(3): 2830–47.

Campos, H., M. Cooper, J.E. Habben, G.O. Edmeades, and J.R. Schussler. 2004. Improving drought tolerance in maize: a view from industry. F. Crop. Res. 90(1): 19–34.

Cobb, J.N., G. DeClerck, A. Greenberg, R. Clark, and S. McCouch. 2013. Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126(4): 867–887.

Crossa, J., Y. Beyene, S. Kassa, P. Pérez, J.M. Hickey, C. Chen, G. de los Campos, J. Burgueño, V.S. Windhausen, E. Buckler, J.L. Jannink, M. A. Lopez Cruz, and R. Babu. 2013. Genomic prediction in maize breeding populations with genotyping-by-sequencing. G3 (Bethesda). 3(11): 1903–26.

Ebert, J. 2014. Faith, Hope & Charity. Cap. J.

Ehrlich, P.R. 1968. The Population Bomb. Ballantine Books, New York.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6(5): e19379.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. J. Anim. Breed. Genet. 124(6): 323–30.

Granier, C., and D. Vile. 2014. Phenotyping and beyond: Modelling the relationships between traits. Curr. Opin. Plant Biol. 18(1): 96–102.

Houle, D., D.R. Govindaraju, and S. Omholt. 2010. Phenomics: the next challenge. Nat. Rev. Genet. 11(12): 855–866.

James, C. 2010. Global status of Commercialized biotech / GM Crops : 2010. ISAAA Br. 42(44): 2009–2009.

Kapila, U. 2009. India's Economic Development Since 1947 (U KAPILA, Ed.).

de Kruif, P. 1928. Hunger Fighthers. Harcourt, Brace and Company, New York.

Lange, K., J.C. Papp, J.S. Sinsheimer, and E.M. Sobel. 2014. Next Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data. Annu. Rev. Stat. its Appl. 1(1): 279–300.

Liebisch, F., N. Kirchgessner, D. Schneider, A. Walter, and A. Hund. 2015. Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. Plant Methods 11(1).

Luikart, G., P.R. England, D. Tallmon, S. Jordan, and P. Taberlet. 2003. The power and promise of population genomics: from genotyping to genome typing. Nat. Rev. Genet. 4(12): 981–994.

Mardis, E.R. 2008. Next-generation DNA sequencing methods. Annu. Rev. Genomics Hum. Genet. 9: 387–402.

Montes, J.M., F. Technow, and B.S. Dhillon. 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. F. Crop. Res. 121(2): 268–273.

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich, and E.S. Buckler. 2009. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21(8): 2194–202.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.-L. Jannink. 2012a. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2): e32253.

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu, Y. Manes, S. Dreisigacker, J. Crossa, H. Sánchez-Villeda, M. Sorrells, and J.L. Jannink. 2012b. Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. Plant Genome J. 5(3): 103.

Poland, J. A., and T.W. Rife. 2012. Genotyping-by-Sequencing for Plant Breeding and Genetics. Plant Genome J. 5(3): 92.

Quinn, K.M. 2009. Norman E. Borlaug - Extended Biograph. Available at http://www.worldfoodprize.org/en/dr_norman_e_borlaug/extended_biography/ (verified 6 February 2016).

Ray, D.K., N.D. Mueller, P.C. West, and J. A. Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One 8(6): e66428.

Tester, M., and P. Langridge. 2010. Breeding technologies to increase crop production in a changing world. Science 327(5967): 818–22.

United Nations, Department of Economic and Social Affairs, P.D. 2015. World Popoulation Prospects: The 2015 Revision, Key Findings and Advance Tables.

Wetterstrand, K. 2016. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP).

White, J., P. Andrade-Sanchez, M.A. Gore, K.F. Bronson, T.A. Coffelt, M.M. Conley, K.A. Feldmann, A.N. French, J.T. Heun, D.J. Hunsaker, M.A. Jenks, B.A. Kimball, R.L. Roth, R.J. Strand, K.R. Thorp, G.W. Wall, and G. Wang. 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133: 101–112.

White, J.W., and M.M. Conley. 2013. A Flexible, Low-Cost Cart for Proximal Sensing. Crop Sci. 53(4): 1646–1649.

Xu, Y., and J.H. Crouch. 2008. Marker-Assisted Selection in Plant Breeding: From Publications to Practice. Crop Sci. 48(2): 391.

# Appendix A - Copyright Permission

This appendix includes copyright permission for the published article Development and Deployment of a Portable Field Phenotyping Platform, Chapter 2.

Copyright Clearance Center

**Note:** Copyright.com supplies permissions but not the copyrighted content itself.

1 PAYMENT | 2 REVIEW | 3 CONFIRMATION

## Step 3: Order Confirmation

**Thank you for your order!** A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us at +1.855.239.3415 Toll Free, M-F between 3:00 AM and 6:00 PM (Eastern), or write to us at info@copyright.com. This is not an invoice.

**Confirmation Number: 11553234**
**Order Date: 04/05/2016**

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.

### Payment Information

Jared Crain
jcrain@ksu.edu
+1 (785)4775872
Payment Method: n/a

### Order Details

### Crop science

| | |
|---|---|
| **Order detail ID:** | 69679918 |
| **Order License Id:** | 3842670263322 |
| **Article Title:** | Development and Deployment of a Portable Field Phenotyping Platform |
| **Author(s):** | Crain, Jared L. ; et al |
| **DOI:** | 10.2135/CROPSCI2015.05.0290 |
| **Date:** | Jan 01, 2016 |
| **ISSN:** | 0011-183X |
| **Publication Type:** | Journal |
| **Volume:** | 0 |
| **Issue:** | 0 |
| **Start page:** | 0 |
| **Publisher:** | CROP SCIENCE SOCIETY OF AMERICA. |
| **Author/Editor:** | CROP SCIENCE SOCIETY OF AMERICA |

**Permission Status:** ✔ **Granted**

| | |
|---|---|
| **Permission type:** | Republish or display content |
| **Type of use:** | republish in a thesis/dissertation |
| **Licensed copyright line** | -01-Copyright � by the Crop Science Society of America, Inc. |
| **Requestor type** | Author of requested content |
| **Format** | Electronic |
| **Portion** | chapter/article |
| **Rights for** | Main product |
| **Creation of copies for the disabled** | no |
| **With minor editing privileges** | no |
| **For distribution to** | Worldwide |

# Appendix B - Supplementary Materials Chapter 2

This appendix includes the supplementary figures and tables for the chapter Development and Deployment of a Portable Field Phenotyping Platform.

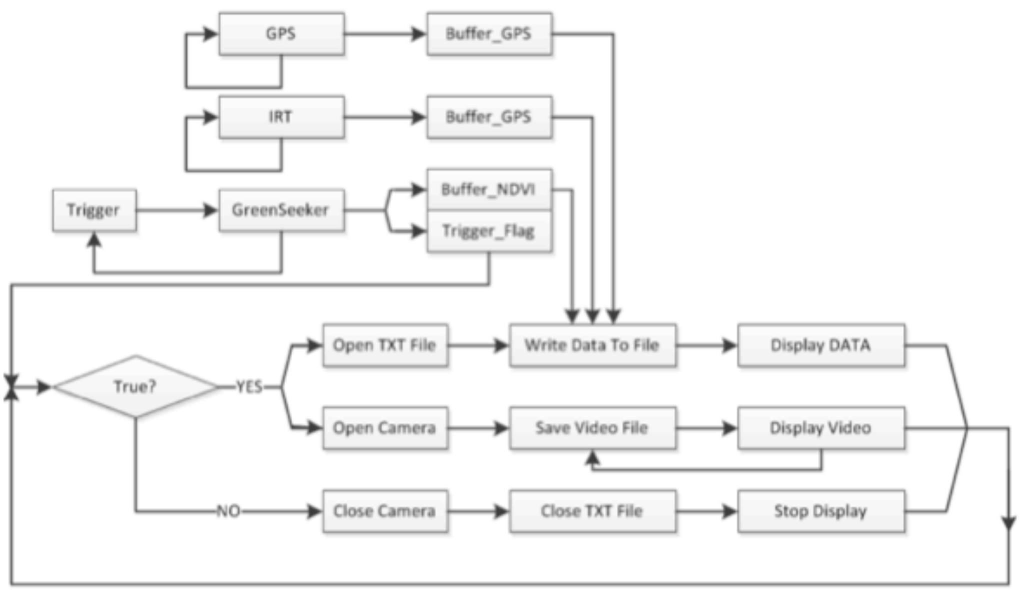


**Figure B-1. Block diagram of the Phenocart software, showing each sensor module and main module that records data from each sensor.**

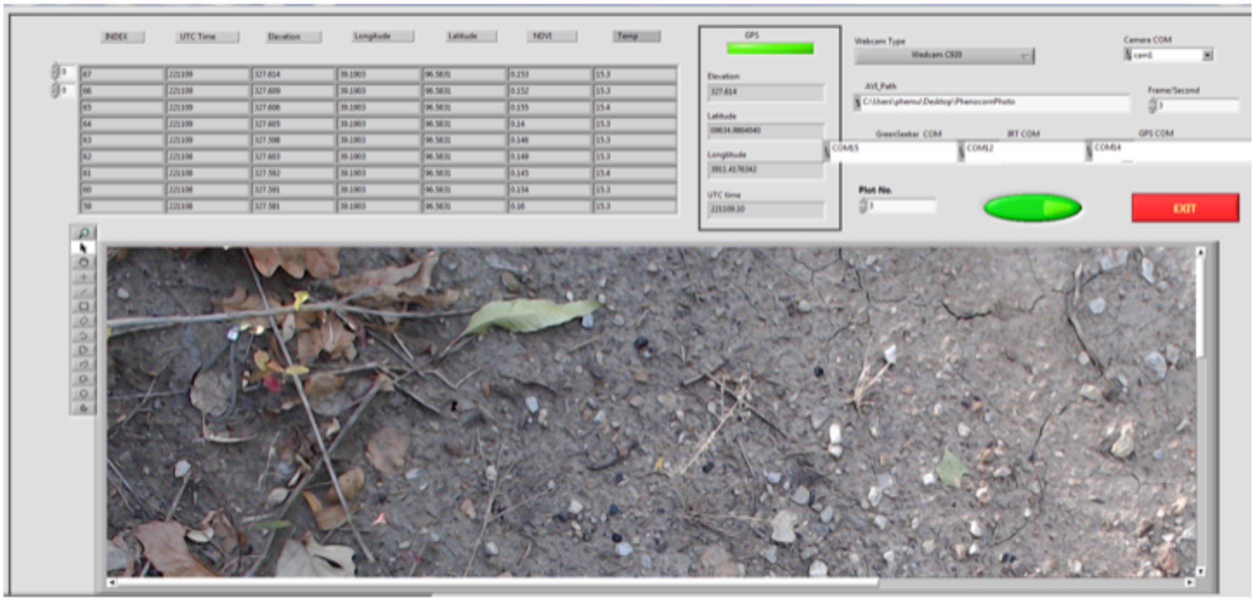


**Figure B-2. Graphical user interface that displays data in real time.**

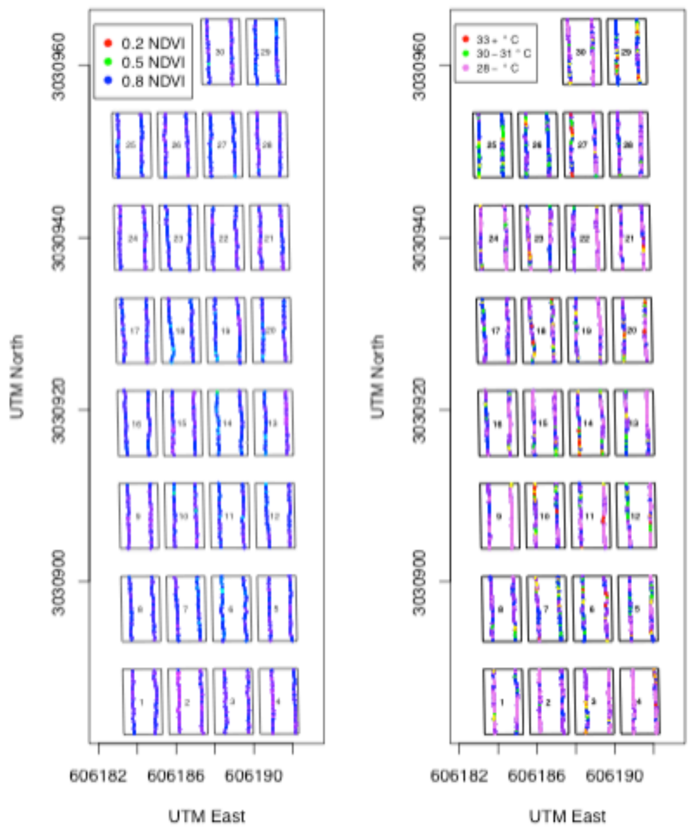


**Figure B-3. A. NDVI color coded by value for vegetative wheat. Low value NDVI is red in color, high value NDVI is blue to purple. B. Canopy temperature color coded by value. Hotter temperatures are in red, with cooler temperatures light in color. Data recorded March 29, 2013, Ciudad, Obregon, Sonora, Mexico.**

**Table B-1. Phenocart Components and Technical Specifications.**

†Represents approximate cost, may be different

| Instrument | Company | Field of View | Output | Output Type | Other Specifications | Sensor Cost† |
|---|---|---|---|---|---|---|
| Aspire 4830 | Acer Inc., San Jose, CA | | | | Windows 7, 2.4 Ghz process, 8 GB RAM, 64 bit | $750 |
| GreenSeeker | Trimble, CA | 1cmX60 cm | 10 Hz/ 38400 baud | RS-232 | RS-232 to USB converter used | $4000 |
| SXBlue III-L | Geneq, Montreal, Canada | | 10 Hz / 19200 baud | USB | Omnistar G2 for 95% accuracy <= 10cm | $6695<br>$1400 Omnistar G2 yearly subscription |
| CT Series Thermometer | MicorEpsilon, Raliegh, NC | 1:2 | 10 Hz/ 9600 baud | analog | Analog to Digital (AD) converter, Temperature resolution 0.1 °C, System accuracy: ±1 °C | $355<br>$75 AD converter |
| C920 Logitech | | | ~3 Hz | USB | HD 1080p | $100 |

# Appendix C - Additional Material for

**Figure C-1. Original Color image as captured by the Phenocart, March 6, 2014, Ciudad Obregon, Mexico.**

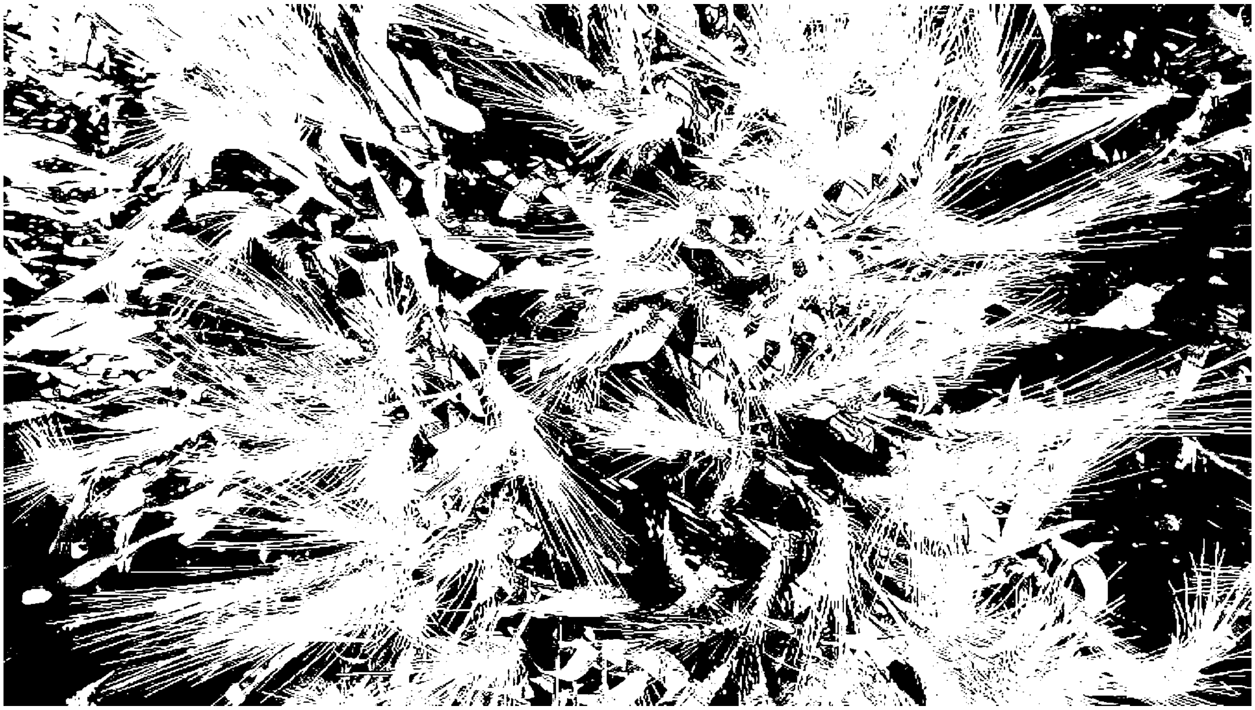

**Figure C-2. Image using color threshold. Pixels that have green HUE between 55-110 have been converted to white. All other pixels are black allowing for a percent green pixels to be calculated. March 6, 2014, Ciudad Obregon, Mexico.**

## IRT Sensor Root Mean Square Error

Root mean square error was calculated for relationship between the handheld infrared thermometer and grain yield, along with Phenocart IRT observations and grain yield. Across five different time points, there were no significant differences in the root mean square error between the two sensor as determined by a t-test (t=0.413, p=0.694).

**Table C-1. Root mean square error values for relationship between measured canopy temperature with Phenocart and Handheld IRT and grain yield for five sampling days.**

| Date | Phenocart RMSE | Handheld RMSE |
|---|---|---|
| 3/29/13 | 27.4 | 25.3 |
| 4/1/13 | 25.8 | 26.2 |
| 4/4/13 | 26.7 | 18.7 |
| 5/10/13 | 22.8 | 26.8 |
| 5/14/13 | 25.1 | 27.2 |

# Appendix D - Supplementary Material Chapter 4

This appendix includes the supplementary figures and tables for the chapter Utilizing High-Throughput Phenotyping for Enhanced Genomic Prediction Accuracy.

**Table D-1. Broad sense heritability for grain yield and HTP traits of canopy temperature (CT) and normalized difference vegetation index (NDVI) for measurement date in each trial.**

| Trial | Grain Yield | NDVI 1 | CT 1 | NDVI 2 | CT 2 | NDVI 3 | CT 3 |
|---|---|---|---|---|---|---|---|
| 2014 Drought Trial 1 | 0.82 | 0.95 | 0.43 | 0.9 | 0 | 0.89 | 0.41 |
| 2014 Drought Trial 2 | 0.78 | 0.91 | 0.57 | 0.91 | 0 | 0.9 | 0.36 |
| 2014 Drought Trial 3 | 0.85 | 0.84 | 0.43 | 0.88 | 0.21 | 0.83 | 0.38 |
| 2014 Drought Trial 4 | 0.8 | 0.96 | 0.31 | 0.96 | 0.6 | 0.96 | 0.77 |
| 2014 Drought Trial 5 | 0.82 | 0.87 | 0.27 | 0.88 | 0.35 | 0.89 | 0.59 |
| 2014 Drought Trial 6 | 0.84 | 0.92 | 0.46 | 0.93 | 0.28 | 0.93 | 0.53 |
| 2014 Drought Trial 7 | 0.96 | 0.92 | 0.53 | 0.94 | 0 | 0.94 | 0 |
| 2014 Drought Trial 8 | 0.98 | 0.9 | 0.02 | 0.96 | 0.3 | 0.95 | 0.69 |
| 2014 Drought Trial 9 | 0.95 | 0.84 | 0 | 0.85 | 0 | 0.91 | 0.33 |
| 2014 Drought Trial 10 | 0.95 | 0.8 | NA | 0.91 | 0.43 | 0.95 | 0.21 |
| 2014 Drought Trial 11 | 0.85 | 0.75 | NA | 0.86 | 0.1 | 0.85 | 0.21 |
| 2014 Drought Trial 12 | 0.78 | 0.84 | NA | 0.88 | 0.31 | 0.93 | 0.45 |
| 2014 Drought Trial 13 | 0.97 | 0.94 | 0.34 | 0.96 | 0.61 | 0.96 | 0.29 |
| 2014 Drought Trial 14 | 0.96 | 0.95 | 0 | 0.96 | 0 | 0.96 | 0 |
| 2014 Drought Trial 15 | 0.96 | 0.91 | 0 | 0.94 | 0.6 | 0.96 | 0.14 |
| 2014 Drought Trial 16 | 0.98 | 0.88 | 0.53 | 0.94 | 0.55 | 0.95 | 0.52 |
| 2014 Drought Trial 17 | 0.96 | 0.87 | 0.22 | 0.92 | 0.07 | 0.9 | 0.58 |
| 2014 Drought Trial 18 | 0.94 | 0.9 | 0.49 | 0.83 | 0.35 | 0.9 | 0.44 |
| 2014 Drought Trial 19 | 0.93 | 0.9 | 0.24 | 0.93 | 0.47 | 0.95 | 0.23 |
| 2014 Drought Trial 20 | 0.78 | 0.96 | 0.45 | 0.97 | 0.61 | 0.98 | 0.26 |
| 2014 Drought Trial 21 | 0.95 | 0.92 | 0.4 | 0.95 | 0.64 | 0.96 | 0.5 |
| 2014 Drought Trial 22 | 0.97 | 0.92 | 0.76 | 0.95 | 0.14 | 0.95 | 0.63 |
| 2014 Drought Trial 23 | 0.95 | 0.9 | 0.21 | 0.94 | 0.55 | 0.95 | 0.45 |
| 2014 Drought Trial 24 | 0.9 | 0.8 | 0.26 | 0.85 | 0.47 | 0.89 | 0.34 |
| 2014 Drought Trial 25 | 0.92 | 0.96 | 0.51 | 0.96 | 0.72 | 0.97 | 0.13 |
| 2014 Drought Trial 26 | 0.98 | 0.97 | 0.68 | 0.98 | 0.75 | 0.97 | 0.43 |
| 2014 Drought Trial 27 | 0.93 | 0.89 | 0.41 | 0.92 | 0.5 | 0.94 | 0.19 |
| 2014 Drought Trial 28 | 0.94 | 0.9 | 0.39 | 0.93 | 0.53 | 0.94 | 0.07 |
| 2014 Drought Trial 29 | 0.94 | 0.88 | 0.2 | 0.93 | 0.6 | 0.94 | 0.55 |
| 2014 Drought Trial 30 | 0.95 | 0.9 | 0.63 | 0.83 | 0 | 0.85 | 0.16 |
| 2014 Drought Trial 31 | 0.98 | 0.93 | 0.56 | NA | NA | 0.95 | 0.05 |
| 2014 Drought Trial 32 | 0.99 | 0.94 | 0.31 | NA | NA | 0.94 | 0.53 |
| 2014 Drought Trial 33 | 0.97 | 0.92 | 0.19 | NA | NA | 0.91 | 0.6 |
| 2014 Drought Trial 34 | 0.93 | 0.95 | 0.64 | NA | NA | 0.96 | 0.36 |
| 2014 Drought Trial 35 | 0.93 | 0.94 | 0.64 | NA | NA | 0.96 | 0.55 |
| 2014 Drought Trial 36 | 0.96 | 0.95 | 0.62 | NA | NA | 0.92 | 0.39 |
| 2014 Drought Trial 37 | 0.97 | 0.78 | 0.57 | NA | 0.82 | 0.82 | 0.43 |
| 2014 Drought Trial 38 | 0.99 | 0.89 | 0.63 | NA | 0.39 | 0.93 | 0.43 |
| 2014 Drought Trial 39 | 0.95 | 0.87 | 0.4 | NA | 0.67 | 0.95 | 0.62 |

Supplemental Table Continued

| Trial | Grain Yield | NDVI 1 | CT 1 | NDVI 2 | CT 2 | NDVI 3 | CT 3 |
|---|---|---|---|---|---|---|---|
| 2015 Drought Trial 1 | 0 | 0.4 | 0.54 | 0.65 | 0.29 | 0.89 | 0.59 |
| 2015 Drought Trial 2 | 0.12 | 0.79 | 0.61 | 0.87 | 0.63 | 0.94 | 0.78 |
| 2015 Drought Trial 3 | 0.02 | 0.84 | 0.72 | 0.93 | 0.61 | 0.97 | 0.73 |
| 2015 Drought Trial 4 | 0.74 | 0.97 | 0.56 | 0.97 | 0.73 | 0.99 | 0.76 |
| 2015 Drought Trial 5 | 0.83 | 0.97 | 0 | 0.96 | 0.41 | 0.98 | 0.69 |
| 2015 Drought Trial 6 | 0.73 | 0.92 | 0.2 | 0.94 | 0.55 | 0.96 | 0.66 |
| 2015 Drought Trial 7 | 0.7 | 0.94 | 0.5 | 0.96 | 0.71 | 0.97 | 0.58 |
| 2015 Drought Trial 8 | 0.82 | 0.97 | 0.24 | 0.98 | 0.59 | 0.99 | 0.54 |
| 2015 Drought Trial 9 | 0.66 | 0.95 | 0.51 | 0.97 | 0.4 | 0.97 | 0.45 |
| 2015 Drought Trial 10 | 0.63 | 0.88 | 0 | 0.95 | 0.38 | 0.96 | 0.63 |
| 2015 Drought Trial 11 | 0 | 0.97 | 0.24 | 0.98 | 0.49 | 0.98 | 0.82 |
| 2015 Drought Trial 12 | 0.04 | 0.96 | 0 | 0.97 | 0.76 | 0.99 | 0.76 |
| 2015 Drought Trial 13 | 0.25 | NA | NA | 0.97 | 0.47 | 0.97 | 0.47 |
| 2015 Drought Trial 14 | 0.71 | NA | NA | 0.98 | 0.73 | 0.98 | 0.83 |
| 2015 Drought Trial 15 | 0.61 | NA | NA | 0.99 | 0.67 | 0.99 | 0.81 |
| 2015 Drought Trial 16 | 0.69 | NA | NA | 0.93 | 0.68 | 0.98 | 0.41 |
| 2015 Drought Trial 17 | 0.47 | NA | NA | 0.97 | 0.68 | 0.99 | 0.77 |
| 2015 Drought Trial 18 | 0.71 | NA | NA | 0.93 | 0.12 | 0.98 | 0.79 |
| 2015 Drought Trial 19 | 0.46 | 0.96 | 0.55 | 0.95 | 0.59 | 0.98 | 0.68 |
| 2015 Drought Trial 20 | 0.74 | 0.9 | 0 | 0.95 | 0.54 | 0.95 | 0.39 |
| 2015 Drought Trial 21 | 0.82 | 0.94 | 0.6 | 0.96 | 0.77 | 0.98 | 0.68 |
| 2015 Drought Trial 22 | 0.81 | 0.71 | 0.17 | 0.63 | 0.23 | 0.86 | 0.84 |
| 2015 Drought Trial 23 | 0.74 | 0.84 | 0.21 | 0.93 | 0.65 | 0.97 | 0.87 |
| 2015 Drought Trial 24 | 0.48 | 0.47 | 0.17 | 0.79 | 0.13 | 0.92 | 0.85 |
| 2015 Drought Trial 25 | 0.5 | 0.95 | 0.3 | 0.92 | 0.59 | 0.98 | 0.85 |
| 2015 Drought Trial 26 | 0.75 | 0.93 | 0.47 | 0.95 | 0.71 | 0.99 | 0.85 |
| 2015 Drought Trial 27 | 0.87 | 0.89 | 0.09 | 0.93 | 0.5 | 0.98 | 0.79 |
| 2015 Drought Trial 28 | 0.91 | 0.88 | 0.72 | 0.92 | 0.4 | 0.97 | 0.85 |
| 2015 Drought Trial 29 | 0.74 | 0.8 | 0.67 | 0.86 | 0.4 | 0.94 | 0.9 |
| 2015 Drought Trial 30 | 0.66 | 0.79 | 0 | 0.81 | 0.43 | 0.9 | 0.67 |
| 2015 Drought Trial 31 | 0.6 | 0.74 | 0.6 | 0.79 | 0.46 | 0.92 | 0.77 |
| 2015 Drought Trial 32 | 0.88 | 0.95 | 0.75 | 0.93 | 0.74 | 0.96 | 0.69 |
| 2015 Drought Trial 33 | 0.91 | 0.97 | 0.73 | 0.99 | 0.59 | 0.99 | 0.85 |
| 2015 Drought Trial 34 | 0.87 | 0.98 | 0.5 | 0.99 | 0.32 | 0.99 | 0.82 |
| 2015 Drought Trial 35 | 0.75 | 0.97 | 0.83 | 0.98 | 0.62 | 0.99 | 0.79 |
| 2015 Drought Trial 36 | 0.05 | 0.96 | 0.39 | 0.97 | 0.12 | 0.98 | 0.7 |
| 2015 Drought Trial 37 | 0.9 | NA | NA | 0.98 | 0.27 | 0.98 | 0.8 |
| 2015 Drought Trial 38 | 0.9 | NA | NA | 0.98 | 0.67 | 0.99 | 0.83 |
| 2015 Drought Trial 39 | 0.85 | NA | NA | 0.98 | 0.42 | 0.99 | 0.75 |

Supplemental Table Continued

| Trial | Grain Yield | NDVI 1 | CT 1 | NDVI 2 | CT 2 | NDVI 3 | CT 3 | NDVI 4 | CT 4 |
|---|---|---|---|---|---|---|---|---|---|
| 2015 Heat Trial 1 | 0.53 | 0.82 | 0.44 | 0.7 | 0.1 | 0.81 | 0.74 | 0.93 | 0.84 |
| 2015 Heat Trial 2 | 0.93 | 0.91 | 0.77 | 0.83 | 0.27 | 0.84 | 0.54 | 0.92 | 0.86 |
| 2015 Heat Trial 3 | 0.72 | 0.9 | 0.85 | 0.8 | 0.6 | 0.83 | 0.66 | 0.88 | 0.82 |
| 2015 Heat Trial 4 | 0.72 | 0.81 | 0.74 | 0.75 | 0.61 | 0.78 | 0.49 | 0.88 | 0.69 |
| 2015 Heat Trial 5 | 0.73 | 0.86 | 0.51 | 0.83 | 0.42 | 0.79 | 0.2 | 0.88 | 0.74 |
| 2015 Heat Trial 6 | 0.83 | 0.53 | 0.54 | 0.51 | 0.27 | 0.78 | 0.4 | 0.86 | 0.63 |
| 2015 Heat Trial 7 | 0.84 | 0.42 | 0.27 | 0.61 | 0.69 | 0.67 | 0.26 | 0.94 | 0.64 |
| 2015 Heat Trial 8 | 0.89 | 0.64 | 0.54 | 0.56 | 0.44 | 0.72 | 0.78 | 0.92 | 0.51 |
| 2015 Heat Trial 9 | 0.87 | 0.75 | 0.86 | 0.65 | 0.82 | 0.79 | 0.64 | 0.96 | 0.85 |
| 2015 Heat Trial 10 | 0.64 | 0.56 | 0.63 | 0.56 | 0.67 | 0.76 | 0.56 | 0.95 | 0.74 |
| 2015 Heat Trial 11 | 0.7 | 0.59 | 0.72 | 0.66 | 0.82 | 0.72 | 0.63 | 0.92 | 0.62 |
| 2015 Heat Trial 12 | 0.84 | 0.63 | 0.63 | 0.66 | 0.66 | 0.85 | 0.62 | 0.93 | 0.65 |
| 2015 Heat Trial 13 | 0.91 | 0.65 | 0.71 | 0.68 | 0.75 | 0.85 | 0.81 | 0.92 | 0.9 |
| 2015 Heat Trial 14 | 0.9 | 0.6 | 0.74 | 0.91 | 0.86 | 0.96 | 0.81 | 0.96 | 0.86 |
| 2015 Heat Trial 15 | 0.95 | 0.87 | 0.86 | 0.82 | 0.73 | 0.96 | 0.81 | 0.98 | 0.92 |
| 2015 Heat Trial 16 | 0.73 | 0.79 | 0.8 | 0.82 | 0.77 | 0.89 | 0.75 | 0.96 | 0.92 |
| 2015 Heat Trial 17 | 0.87 | 0.43 | 0.59 | 0.65 | 0.42 | 0.9 | 0.57 | 0.96 | 0.87 |
| 2015 Heat Trial 18 | 0.91 | 0.63 | 0.56 | 0.71 | 0.58 | 0.82 | 0.62 | 0.96 | 0.77 |
| 2015 Heat Trial 19 | 0.94 | 0.72 | 0.57 | 0.71 | 0.72 | 0.92 | 0.77 | 0.97 | 0.77 |
| 2015 Heat Trial 20 | 0.92 | 0.67 | 0.74 | 0.79 | 0.83 | 0.88 | 0.84 | 0.97 | 0.82 |
| 2015 Heat Trial 21 | 0.93 | 0.46 | 0.7 | 0.86 | 0.81 | 0.93 | 0.87 | 0.98 | 0.86 |
| 2015 Heat Trial 22 | 0.87 | 0.43 | 0.39 | 0.72 | 0.73 | 0.92 | 0.6 | 0.98 | 0.78 |
| 2015 Heat Trial 23 | 0.97 | 0.79 | 0.53 | 0.93 | 0.79 | 0.94 | 0.59 | 0.97 | 0.79 |
| 2015 Heat Trial 24 | 0.87 | 0.55 | 0.71 | 0.84 | 0.78 | 0.87 | 0.76 | 0.96 | 0.89 |
| 2015 Heat Trial 25 | 0.72 | 0.66 | 0.52 | 0.81 | 0.53 | 0.86 | 0.56 | 0.93 | 0.78 |
| 2015 Heat Trial 26 | 0.79 | 0.66 | 0.63 | 0.76 | 0.46 | 0.78 | 0.56 | 0.9 | 0.61 |
| 2015 Heat Trial 27 | 0.78 | 0.76 | 0.77 | 0.88 | 0.61 | 0.91 | 0.57 | 0.97 | 0.87 |
| 2015 Heat Trial 28 | 0.86 | 0.72 | 0.67 | 0.67 | 0.7 | 0.88 | 0.63 | 0.95 | 0.78 |
| 2015 Heat Trial 29 | 0.81 | 0.61 | 0.47 | 0.56 | 0.39 | 0.81 | 0.57 | 0.97 | 0.78 |
| 2015 Heat Trial 30 | 0.88 | 0.43 | 0.57 | 0.72 | 0.32 | 0.79 | 0.69 | 0.95 | 0.87 |
| 2015 Heat Trial 31 | 0.89 | 0.68 | 0.51 | 0.9 | 0.76 | 0.95 | 0.66 | 0.97 | 0.86 |
| 2015 Heat Trial 32 | 0.88 | 0.71 | 0.65 | 0.71 | 0.6 | 0.9 | 0.6 | 0.97 | 0.82 |
| 2015 Heat Trial 33 | 0.84 | 0.79 | 0.67 | 0.84 | 0.23 | 0.9 | 0.59 | 0.96 | 0.8 |
| 2015 Heat Trial 34 | 0.84 | 0.79 | 0.61 | 0.91 | 0.5 | 0.94 | 0.77 | 0.97 | 0.8 |
| 2015 Heat Trial 35 | 0.89 | 0.88 | 0.63 | 0.94 | 0.77 | 0.92 | 0.75 | 0.93 | 0.86 |
| 2015 Heat Trial 36 | 0.85 | 0.65 | 0.67 | 0.75 | 0.59 | 0.84 | 0.47 | 0.86 | 0.75 |
| 2015 Heat Trial 37 | 0.88 | 0.79 | 0.57 | 0.88 | 0.31 | 0.9 | 0.68 | 0.91 | 0.8 |
| 2015 Heat Trial 38 | 0.7 | 0.77 | 0.56 | 0.89 | 0.33 | 0.92 | 0.68 | 0.97 | 0.89 |
| 2015 Heat Trial 39 | 0.7 | 0.58 | 0.62 | 0.86 | 0.71 | 0.93 | 0.73 | 0.96 | 0.9 |