THESIS

UTILIZING PLANT GENETIC RESOURCES FOR PRE-BREEDING OF WATER-EFFICIENT SORGHUM: GENETICS OF THE LIMITED TRANSPIRATION TRAIT

Submitted by

Gina Cerimele

Department of Soil and Crop Sciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2022

Master's Committee:

Advisor: Geoffrey Morris

Francesca Cotrufo
John McKay

ABSTRACT

UTILIZING PLANT GENETIC RESOURCES FOR PRE-BREEDING OF WATER-EFFICIENT SORGHUM: GENETICS OF THE LIMITED TRANSPIRATION TRAIT

Shifting precipitation patterns driven by the changing climate threaten productivity of dryland agricultural systems. Increasing the efficiency of water use by crops grown in dryland regions, such as sorghum (*Sorghum bicolor*), is a target for plant breeding to address this issue. c variants conferring efficient water use in sorghum may be found within collections of plant genetic resources (PGR). However, tropical sorghum PGR require adaptation to the target temperate environment to begin the pre-breeding trait discovery process. The landmark Sorghum Conversion Program unlocked diverse sorghum genetics for temperate breeding by adapting tropical African lines to temperate height and maturity standards.

In the U.S. Sorghum Belt, spanning South Dakota to central Texas, the limited transpiration (LT) trait could provide growers a 5% yield increase in water-limited conditions with high vapor pressure deficit (VPD) according to crop modeling. To transfer the LT trait into commercial breeding programs, an elite donor line must be developed. Characterizing the genetic architecture of LT informs markers and breeding strategy for development of an elite donor. To characterize the genetic architecture of LT, two biparental recombinant inbred line (RIL) mapping families were developed from crossing putative LT parents SC979 and BTx2752 by putative non-LT parent RTx430. For this study, the families were grown together as a mapping population in three locations (continental-humid eastern Kansas, semi-arid western Kansas, and semi-arid Colorado) in one year. The families were phenotyped for the LT trait using UAS-

collected thermal imaging and canopy temperature as a proxy. The families were initially designed with the goal of controlling phenotypic covariates of canopy temperature associated with height and flowering time, like neighbor-shading and artifactual temperature inflation related to panicle imaging.

To test whether the family design controlled for height and flowering time covariates, the populations were phenotyped for both traits. High broad-sense heritability (H^2) > 0.86 for all traits and families across locations indicates that the traits are not fixed. However, phenotypic distributions reveal that most lines are within an agronomically-relevant range that limits confounding covariates. Using DArTseq-LD genotyping data, GWAS analyses of height and flowering time reveal putatively significant marker-trait associations (MTA) with known loci underlying height and maturity in sorghum. These results collectively indicate that, while genetic variation for height and flowering exist in the LT mapping families, the resulting phenotypes are homogeneous enough to be suitable for LT genetic mapping.

To test hypotheses on the monogenic, oligogenic, or polygenic architecture of the LT trait, canopy temperature data collected by the UAS-thermal imaging missions was used. Non-zero H^2 of canopy temperature in most location-timepoints indicates genetic variation is present for LT in the population. Continuous phenotypic distributions imply a quantitative architecture. GWAS analyses revealed moderate marker-trait association peaks visible within timepoints and across locations, indicating oligogenic architecture of LT. Some of those peaks also colocalize with sorghum homologs of aquaporin genes in *Arabidopsis thaliana*, suggesting that aquaporin variation could be a molecular basis underlying the trait. These results provide a basis for marker-assisted selection in developing an LT donor line.

ACKNOWLEDGEMENTS

Completing an MSc program has been an illuminating and transformative process shaped by many wonderful people and experiences. Thank you to my family—Jim, Janet, Lindsey, and Wilma—for learning what sorghum is so you could enthusiastically tell your friends about my work. Your support, love, and pride has been so uplifting. Thank you to all my friends near and far for cheering me on during the highs and helping me through the lows. Especially thank you to my advisor, Dr. Geoff Morris, who took me on as a student under the most hectic circumstances but believed in me from day one, nonetheless. I know, as you say, "in my heart of hearts" that this lab is where I was meant to end up. Your commitment to rigorous science inspires me daily to use it to make the world a better place. Thank you to Dr. Pat Byrne for seeing potential in me for research and science communication at Colorado State University. A huge thank you to my committee members, Dr. Francesca Cotrufo and Dr. John McKay, for your investment in my research and asking all the hard questions to improve my work. Thank you to the team at Kansas State University—Terry, Sarah, Rob, Trevor, and Matt—who have supported parts of this thesis project from afar and shown me valuable skills in plant breeding. Finally, thank you always to all Morris Lab members for your willingness to collaborate and push me to be better. Most importantly, thank you for helping weed my fields in the hot July sun, I owe you one.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER II

CHAPTER III

CHAPTER I: SORGHUM CONVERSION PROGRAM UNLOCKS SORGHUM GENETIC DIVERSITY FOR TEMPERATE BREEIDNG

Genetic Diversity Sustains Crop Improvement

Utilization of genetic diversity is integral to crop improvement. Domestication, commercial crop breeding, and genetic drift cause a loss of genetic diversity within crop species, which in turn results in losing agronomically relevant traits from breeding lines (Jordan et al., 2011). Necessary genes for traits that could facilitate a higher or more stable yield, such as disease resistance and drought tolerance, may not be available in elite breeding lines used by commercial breeders (Papa et al., 2005). Additionally, this narrow genetic basis places crops at increased risk for total loss when faced with continually evolving and migrating stressors (Nevo et al., 1979). To aid in combating these issues, plant genetic resources (PGR) offer readily available pools of diversity that can be unlocked for use in crop improvement via pre-breeding, trait discovery, and knowledge of phylogenetic relationships (Zongo et al., 2005).

Pre-Breeding Bridges Germplasm and Crop Improvement

Development of conserved PGR is essential for their inclusion in crop improvement programs. Genetic resource conservation is facilitated both *in situ*, in the natural environment, and *ex situ*, by removal for collection at a genebank or associated facility (Gepts, 2006). PGR entering the genebanking system, broadly termed germplasm accessions, undergo characterization to classify various morphological, functional, and molecular attributes. Alongside the characterization data, detailed information on the collection site, origin, and taxonomy of the sample is cataloged (Ramanatha Rao & Hodgkin, 2002). This information is

integral for crop improvement next-users to identify candidate accessions that may hold the genetic diversity necessary to the advancement of their breeding programs (Gepts, 2006).

Once accessions have been identified, understanding the genetic basis of the trait or gene(s) of interest increases the feasibility of PGR use in crop improvement. Linkage drag with undesirable genes, complex polygenic architecture, gene pool incompatibility, and temperate-tropical adaptation are common roadblocks that may be overcome using pre-breeding when developing germplasm (Sharma et al., 2013). First, genes of interest are often found in unadapted, unimproved germplasm that may require adjustments to photoperiod sensitivity for production in the target range. Backcross introgression schemes using an adapted donor parent enable large-scale conversion of germplasm into usable lines (Allier et al., 2020; Cowling et al., 2009). Next, genes desirable for crop improvement may be in linkage disequilibrium with genes controlling for undesirable traits. An example is grain tannins in possible linkage with early-season chilling tolerance in Chinese sorghum germplasm, where tannins are unacceptable in U.S. hybrids. Using a repeated backcrossing and selection process, pre-breeders may be able to introgress only the desired genes into a donor breeding line (Jordan et al., 2011; Tanksley & Nelson, 1996). Next, while certain traits like some disease resistance may commonly be under qualitative control, many physiological traits are controlled quantitatively. The more complex the genetic architecture, the higher the likelihood that carefully constructed yield linkage groups in commercial breeding lines could be disrupted during recombination. This is unfavorable to commercial breeders, therefore qualitative traits are preferred for integration into programs (Allier et al., 2020). Finally, diverse genetic materials containing the genes of interest may belong to a secondary or tertiary gene pool of the target breeding lines. Crossing species from different gene pools often results in fertility issues, limited gene transfer, lethality, or complete

sterility (Harlan & Wet, 1971). Pre-breeding can employ radical techniques to make successful crosses and transfer target genes into donor lines. Together, all or some of these techniques bridge the gap between conserved unadapted germplasm and their feasible use in crop improvement.

Sorghum Conversion Program Unlocks Untapped Genetic Materials

The Sorghum Conversion Program (SCP) is a particularly notable example of utilizing PGR to develop collections of adapted germplasm readily available for crop improvement. Sorghum (*Sorghum bicolor* (L.) Moench) is the fifth most important cereal crop produced globally, serving as a staple to over 500 million people living in semi-arid regions (*FAOSTAT*, 2021). Prominent drought tolerance characteristics in sorghum facilitate its usefulness as a staple cereal in those semi-arid regions like Africa's Sahel (Mundia et al., 2019). In the United States, sorghum is primarily grown in the southern and central plains for animal feed, forage, and biofuel (Stamenković et al., 2020), in addition to the large export market primarily selling to China in recent years (Monk et al., 2014; Xing-Lin et al., 2017). Sorghum has a reported center of origin in northeastern Africa near the Sudan-Egypt border (Harlan & Stemler, 2011), but with evidence of other independent domestication events in West Africa and Ethiopia (Sagnard et al., 2011). Hence, the vast diversity of African sorghum offers potential sources of allelic diversity useful for sorghum breeding in global regions outside of Africa.

The SCP sought to adapt tropical African sorghums to temperate grain sorghum breeding specifications in the United States while retaining the wealth of allelic diversity present elsewhere in genomic regions not involved in the conversion (Stephens et al., 1967). First implemented in the 1960s as a partnership between the United States Department of Agriculture-Agricultural Research Service (USDA-ARS) and Texas A&M University-Texas Agricultural

Experiment Station (TAES), the SCP is still active today and continues to release adapted lines (R. Klein et al., 2015a, p. 40). To adapt these sorghums, they are converted from photoperiod-sensitive short-day, tall plants to photoperiod-insensitive, dwarfed plants able to be mechanically harvested in temperate latitudes (R. Klein et al., 2008; Stephens et al., 1967). Once adapted to these breeding standards, temperate sorghum pre-breeding programs are able to explore desirable alleles, genes, or traits of interest in the converted exotic germplasm and introgress them into elite breeding lines. From these introgressions, commercial breeders receive donor lines containing the target genes that are compatible with the elite commercial breeding material. These donor lines contain the genetic material of interest in a relevant elite background without undesirable linkage drag which disrupts constructed yield haplotypes. (Hao et al., 2021; K. Singh et al., 2019). Ultimately, adapting tropical germplasm to temperate breeding standards is necessary for broadening the genetic basis of grain sorghum in the US and ensuring continual advancement in temperate sorghum crop improvement.

Intentions and Motivation

Grain sorghum has been grown in the United States in some capacity since the middle of the 19th century. Its inception was driven by farmer selections for early and dwarf mutant plants in fields of tropical Standard Yellow Milo and Guinea Kafir, varieties originally brought to the continent from Africa through the slave trade (Quinby, 1974). A few other temperate-adapted African lines, such as Hegari and PI 54484, were also introduced to U.S. production over time. Farmer selections within the same sets of limited germplasm alone could not significantly expand grain sorghum genetic diversity and yield (Quinby, 1974), and beneficial spontaneous mutations for photoperiod insensitivity were not abundant. The discovery of cytoplasmic-genetic male sterility (CMS) in sorghum in the early 1950s unlocked revolutionary new opportunities for

efficiently producing hybrids with heterosis. Hybrids offered promising potential that incentivized sorghum breeders to identify suitable diverse A/B and R-line pairs with good combining ability to use as parents. The available temperate-adapted US germplasm was limited in such pairings, causing breeders to turn their search to the geographic origins of sorghum. Under the Sorghum Conversion Program diverse tropical African varieties were to be sourced and converted to temperate breeding specifications.

The evolutionary history of sorghum in Africa produced a wealth of genetic diversity within tropical African landraces, crop wild relatives, and improved varieties (Ananda et al., 2020). Valuable alleles that improve yield and quality through increased heterosis, drought adaptation, disease and insect resistance, nutrition, and grain quality traits may only be found in tropical African germplasm that is unable to flower in temperate breeding environments. Spontaneous mutations in sorghum for photoperiod-insensitivity occurred in temperate regions of Africa, Asia, and the southern US, offering some viable lines able to be cultivated in higher latitudes (Quinby, 1974; Thurber et al., 2013a). Prior to the development of converted lines, U.S. breeding programs were limited to only that small pool of available temperate-adapted germplasm, unable to evaluate tropical materials in target environments.

The original Sorghum Conversion Program pursued introgression of elite haplotypes for early flowering and dwarf height into exotic African germplasm while leaving non-conversion haplotypes undisturbed (Stephens et al., 1967). The converted sorghums were then able to be used in trait screening and breeding in temperate latitudes based on pollen availability and flowering times better coinciding with that of existing US breeding lines. There was not a distinctly specified justification for most of the original candidate conversion lines. The accessions were selected with the goal of capturing an assortment of genetically diverse

germplasm that could potentially prove useful in sorghum improvement. Over time, the converted accessions could provide the necessary diversity to sorghum geneticists and breeders to fulfill trait exploration and elite breeding line development.

Timeline

The SCP began as a partnership between the USDA-ARS research station in Mayaguez, Puerto Rico and TAES in 1963. With plantings beginning in Mayaguez in 1963, one generation per year in each Puerto Rico and Lubbock, Texas was standard. The tropical Puerto Rican climate offered potential for two generations in some seasons. Expedited backcrossing and selection aided by the favorable Mayaguez environments allowed for the first partially converted SCP materials to be released in 1969. Distribution of the first 63 fully converted lines was completed the following year in 1970, (Rosenow, 1978) and the program released approximately 30-40 lines per year afterward through 1999. More than 700 fully converted lines in total were delivered by the SCP.

The SCP was reinstated in 2009, employing molecular tools to shorten conversion time (R. Klein et al., 2015b). Genomic selection is used to identify lines that display a higher proportion of the tropical African alleles after only a single backcross. Marker-assisted selection has also become the baseline standard to increase conversion efficiency. Many of the accessions chosen for conversion through the reinstated SCP are derived from the Ethiopian and Sudanese germplasm collections, based on the history of valuable alleles being found within them (R. Klein et al., 2015; R. Klein et al., 2013). The reinstated SCP, still active today, continues to release lines with prospective alleles for cytoplasmic male sterility and fertility restoration, disease and insect resistance, grain quality, and drought tolerance traits. The reinstated program

primarily focuses on partial conversions. Over 1500 total converted or partially-converted lines have been released through the programs to date.

Technical Approach

Facilitators of the SCP used a backcross and phenotypic selection approach in converting materials to early and short temperate breeding standards. Nearly all the African materials for conversion were crossed with U.S. breeding line BTx406 as a maturity and height donor. BTx406 contains the maturity genotypes *ma*$_1$, *Ma*$_2$, *Ma*$_3$, *Ma*$_4$, which produce earlier flowering phenotypes, and the height genotype *dw*$_2$, which produces shorter stature phenotypes. Pedigree analysis of maturity and height in BTx406 revealed that Early White Milo was the *ma*$_1$ donor and Double Dwarf Yellow Milo was the *dw*$_2$ donor (R. Klein et al., 2008; Quinby, 1967). In the F_2 generation, plants that displayed the appropriate shorter height and early flowering time but otherwise retained characteristics of the exotic parent were selected and then backcrossed to the exotic parent. For partial conversions, lines were released at the $BC_1F_{2:3}$ (R. Klein et al., 2015b). Full conversion involved advancement to the BC_4F_3 where additional selections for selfing and inbred development were carried out (C. W. Smith & Frederiksen, 2000).

Phenology in photoperiod-sensitive plants is related to the latitude of the growing environment and its influence on day length (Stinchcombe et al., 2004). The short-day nature of sorghum results in a critical photoperiod of approximately 13 hours or less. Outside of latitudes 0° to 20° N or S, this critical photoperiod is exceeded during the growing season, extending the number of days to flowering and maturity (Abdulai et al., 2012). Lack of late-season chilling tolerance in sorghum necessitates conversion to early, photoperiod insensitive lines to reach maturity before low temperatures kill off plants (S. P. Singh, 1985). In addition to latitudinal

positioning, phenology is influenced by temperature, circadian rhythms, developmental stage of the plant, and phytohormones (Casto et al., 2019).

The genetic control of flowering in sorghum has been studied using homologs in the model plant *Arabidopsis thaliana*. In *A. thaliana*, a long-day plant, the essential core flowering time gene *FLOWERING LOCUS T (FT)* is expressed as a response to light signals and circadian cues. *FT* is expressed under long-day conditions which initiates flowering (Corbesier et al., 2007). *FT* homologs are present in sorghum, as flowering time pathways are conserved across species to varying degrees. In sorghum, *SbCN8* and *SbCN12* are *FT*-like genes that regulate floral transition. They are expressed once day length falls within the critical short-day photoperiod. *SbCN8* and *SbCN12* are regulated by several genetic factors, including *CONSTANS* (Casto et al., 2019) and the six maturity loci, Ma_1-Ma_6, of which Ma_1 and Ma_6 control the largest effects (R. Klein et al., 2008). Lateness is dominant at each of the maturity loci.

The genetic control of height in sorghum has been of interest to U.S. breeding programs since the 1950s when mechanized harvest became a breeding priority (Quinby, 1974). Height is a result of the genes controlling the number of internodes and their respective lengths (Hadley, 1957). Height is primarily controlled by four additive loci in sorghum, Dw_1-Dw_4, with tallness being dominant at each (Hadley, 1957; Li et al., 2015). It is also influenced by genetic factors outside of the four dwarfing loci, clearly visualized with contrasting heights produced in different lines with the same height genotype (Quinby & Karper, 1953). Recessiveness at the loci causes brachytic dwarfing, where internode length is reduced independently of other stem or leaf traits to make the plant shorter (Hilley et al., 2017). While tall landraces in Africa are favored for stover in addition to grain, U.S. grain sorghum producers expect three-dwarf plants (recessive at three of the four height loci) no taller than approximately 150 centimeters.

Two large-effect maturity loci and one large-effect height locus important to temperate adaptation are mapped to the same linkage group (Lin et al., 1995). *Ma*$_6$, *Ma*$_1$, and *Dw*$_2$ were placed on the distal end of chromosome 6, where *Ma*$_6$ is mapped to a region between 0.5 and 0.8 Mb (Murphy et al., 2014), *Ma*$_1$ to ~40.3 Mb (Bowers et al., 2003; Thurber et al., 2013a), and *Dw*$_2$ to ~42.3 Mb (Hilley et al., 2017). This proximity places *ma*$_1$ and *dw*$_2$ in tight linkage, supported by the high co-inheritance of both haplotypes (Higgins et al., 2014; Lin et al., 1995; Morris et al., 2013). Retrospective analyses of the conversions have revealed that chromosome 6 is disproportionately identical-by-descent with Milo and Kafir ancestry present in the elite donor. The ancestry and inbreeding practices favored in U.S. sorghum improvement have created large haplotypes in genomic regions with high linkage disequilibrium. In some cases of converted lines, nearly the entire elite chromosome 6 was introgressed into the exotic recurrent parent as a single haplotype, failing to retain the exotic genome outside of the necessary conversion haplotypes (R. Klein et al., 2008; Thurber et al., 2013a). Failure to retain the exotic chromosome 6 diversity in converted lines creates a gap in allelic diversity on the linkage group and across the genome available for temperature breeding.

Breeding with Converted Lines Delivers Valuable Traits

Converted lines are continually included in screening panels for traits of interest, and several traits of agronomic importance have been derived from them to date. Resistance traits to biotic stresses, such as sorghum aphid and anthracnose, as well as adaptive traits to abiotic stresses, such as staygreen, were discovered in converted lines and successfully integrated into commercial breeding programs.

Sorghum Aphid Resistance

Phylogenetically, sorghum and sugarcane are closely related (Aono et al., 2021). In 2013, a sorghum-feeding biotype of the sugarcane aphid (*Melanahpis sacchari*), dubbed sorghum aphid (*Melanaphis sorghii*), moved into southeastern Texas and quickly expanded its range outward into the remainder of Texas, Louisiana, and Mississippi. After successfully overwintering in southern Texas, populations surged through much of the Sorghum Belt and had infested areas totaling 90 percent of U.S. sorghum acreage by 2015 (Medina et al., 2017). Meanwhile, sorghum aphid (SA) also appeared in Haiti in 2015 and caused catastrophic crop losses to the order of 80,000 metric tons. Haiti relies heavily on sorghum as a staple grain and forage, and the presence of SA threatened food security on a national level.

SA infestations present a variety of challenges in sorghum production. When large colonies proliferate on the underside of leaves and pass threshold levels, the sap-sucking feeding activity severely damages the stand. This stress contributes to substantial yield losses in pre-flowering infestations (Paudyal et al., 2020). Additionally, SA produces a sticky honeydew that is host to saprophytic fungi from which sooty mold arises and degrades crop quality. The honeydew in large quantities can also cause malfunctions in harvesting equipment (Mbulwe et al., 2016). These multi-modal effects of SA in Haiti and the U.S. Sorghum Belt launched plant breeders into action to develop SA-resistant lines.

The Chibas breeding program at Quisqueya University in Haiti was able to make remarkably quick progress on development of a SA-resistant sorghum variety using globally admixed germplasm and rapid intercrossing. The globally-rare resistance allele was originally identified in the exotic Ethiopian accession PI 276837 (Muleta et al., 2021) and is found at the *RMES1* locus between 2.6 and 2.8 Mb on chromosome 6 in a region consisting of 126kb (Wang

et al., 2013). Just two years after the devastation of the 2015 growing season, the Haitian breeding program was able to successfully select for the resistance allele in crosses using SA-resistant West African variety IRAT204 (PI 656031) and release SA-resistant Haitan variety Papèpichon (Muleta et al., 2021).

PI 276837 was converted as part of the Sorghum Conversion Program and registered as SC170 (PI 534157) in 1986 (Rosenow, 1986). This made it readily available for use in developing SA-resistant U.S. breeding lines alongside converted line SC110 (PI 533794), another resistance source registered in 1970. Popular U.S. resistant line RTx2783 (PI 656001) was developed using SC110 and remains heavily used in U.S. breeding efforts today (Armstrong et al., 2015).

The efforts of the Haitian breeding program and the ability to screen exotic African germplasm for SA-resistance alleles were imperative in responding rapidly to devastating SA infestations. In part due to their immediate accessibility, the converted lines that contained the crucial *RMES1* locus were critical in limiting further destructive infestations in U.S. sorghum production. While SA still exists in U.S. and Haitian sorghum fields, the development of resistant lines using exotic and converted material allows growers to minimize chemical insecticide inputs and maintain stable yields. Today, almost 90% of Haitian sorghum acreage is planted with the SA-resistant varieties, and the security of sorghum as a staple has returned to the region.

Anthracnose Resistance

Sorghum anthracnose, caused by the fungal pathogen *Colletotrichum sublineolum*, is a disease plaguing sorghum grown in warm, humid regions. Anthracnose is characterized by a variety of signs on all parts of the plant, but predominantly affects the stem and leaves.

Photosynthesis is inhibited by oval-shaped or circular lesions containing asexual fruiting bodies called acervuli on leaves (Cuevas & Prom, 2020). Infected stems supporting the panicle become internally discolored and blocked by the fungus, slowing or preventing the movement of water and nutrients through the xylem and phloem. This results in inhibited grain development, negatively affecting yield (Gwary et al., 2004). Yield losses due to anthracnose vary significantly and may exceed 60% in some cases based on factors like climate and management. Resistant varieties are a recommended component of integrated pest management for growers, but the highly variable genetic diversity of *C. sublineolum* populations poses a challenge for introducing comprehensive and lasting resistance to the pathogen (Cuevas et al., 2018). Another complication is the variable genetic architecture of anthracnose resistance in sorghum. Studies have reported single-gene resistance (Boora et al., 1998), multiple genes involved in resistance (Cuevas et al., 2018), and single gene resistance with multiple allelic forms. This impacts the feasibility of including some novel resistance sources in breeding. However, screening diverse sorghum germplasm, which may contain other novel anthracnose resistance alleles and genetic architectures, is a necessary endeavor to continue development of durably resistant varieties (Cruet-Burgos et al., 2020).

Both the Ethiopian and Sudanese core sorghum collections, in addition to the sorghum association panel (SAP), were heavily screened for anthracnose resistance genes. The Ethiopian core collection of 335 lines contained a majority of resistant accessions with 169 scoring as resistant in one study (Cuevas et al., 2018). These accessions include converted lines like SC155 (J Felderhoff et al., 2016; Patil et al., 2017), which have been successfully used in breeding anthracnose-resistant lines to the Texas and Georgia pathotypes (Erpelding, 2010). However,

many of the core collection accessions share the same resistance alleles and mechanism, which does not offer durable control of genetically diverse *C. sublineolum* populations.

Many of the Ethiopian accessions with high frequency of resistance alleles are of the durra type and contain the candidate resistance gene *Sobic.009G013300* on chromosome 9. *Sobic.009G013300* codes for pathogen recognition as an NBD-LLR (Cuevas et al., 2018). It has been proposed that, evolutionarily, the durra background may have generated the high frequency of resistance alleles at this locus. Ultimately, the Ethiopian core collection has provided valuable anthracnose-resistant breeding lines to US grain sorghum via the SCP but diversifying the genetic background of the resistance mechanism is necessary to provide durable resistance (Erpelding, 2010).

The Sudanese core collection was also screened for resistance genes. In one study, 318 Sudanese accessions were evaluated and less than a quarter were found to be resistant (55 accessions). In these 55 accessions, GWAS determined that resistance was primarily correlated with loci on chromosomes 2 and 5 (Patil et al., 2017), though each locus only explained between 2 and 7 percent of the observed variation (Cuevas & Prom, 2020). Consistent with the findings from the Ethiopian core collection, the durra type accessions contained nearly double the rare resistance alleles as the caudatum type accessions did. Notably, the Sudanese converted line SC748-5 derives its resistance from a locus on the distal region of chromosome 5 that codes for the *Cg1* gene (Perumal et al., 2009). SC748-5 is highly valued for breeding for the Texas pathotype (Burrell et al., 2015).

Accessions from different national core collections and geographic localities provide anthracnose resistance from a multitude of loci and genetic architectures. This illustrates the need to conserve accessions from many regions and employ their usage in breeding programs (Cuevas

et al., 2018). While fully understanding the genetic architecture of anthracnose resistance and the formation of pathotypes will make for more efficient germplasm screening, breeders are seeking resistance now. Diversifying U.S. breeding materials with converted lines from distinct African locales is a viable defense to aid in prevention of widespread anthracnose susceptibility in sorghum (Erpelding, 2010).

Stay-green

The exceptional drought tolerance present in sorghum is attributed to its domestication in the sub-Saharan Sahelian regions of Africa. Even so, as climate change shifts predictability of precipitation patterns and water availability, sorghum is affected by post-flowering drought (Vadez et al., 2013). The discovery of the stay-green trait in converted lines has provided promising potential as a safeguard against such conditions in rainfed systems. Stay-green lines are characterized by delayed leaf senescence post-anthesis independent of soil water content, which allows the sorghum to maximize grain growth rates and yield compared to senescent lines (Mahalakshmi & Bidinger, 2002).

The stay-green trait was first introduced in 1968. Most breeding for stay-green in US sorghum hybrids has origins from one African source, an Ethiopian durra landrace (Borrell et al., 2014). Known in industry as BTx642 or B35, this line was developed as a partially-converted derivative of SC35-6 (SC35-6 later underwent full conversion and was released in 1986). As a partially-converted line, B35 was selected at the BC_1 during the SCP height and maturity introgressions (Kassahun et al., 2010; Vadez et al., 2013). At this stage, the unadapted Ethiopian accession IS 12555 had been crossed to donor parent BTx406 and only one generation of backcross of selected progeny to IS 12555 had occurred. Since the introduction of B35 as a

breeding line, markers have been developed to aid in the selection of putative stay-green loci (Sukumaran et al., 2016).

At least four independent stay-green loci have been identified, *Stg1-Stg4*, though their independent roles in trait expression have not been fully determined (Kassahun et al., 2010). QTL studies have placed *Stg1* and *Stg*2 on chromosome 3, *Stg3* on chromosome 2, and *Stg4* on chromosome 5 (Sanchez et al., 2002). Additionally, analyses of the QTL effect sizes show that *Stg2* accounts for 30% of the phenotypic variability, *Stg1* accounts for 20%, *Stg3* accounts for 16%, and *Stg4* accounts for 10% (Xu et al., 2000).

There are a variety of physiological processes that contribute to delayed senescence that the loci may be involved in, including decreased root angle, decreased size of upper leaves, and prolonged chlorophyll retention (Thomas & Howarth, 2000). Only functional expressions of the stay green trait will ultimately contribute to stable yields in post-anthesis drought, meaning that photosynthetic activity must be maintained during grain fill (Borrell et al., 2014). Therefore, genotypes that simply remain green in post-flowering drought but do not conduct photosynthetic activity are non-functional stay-green and not relevant to sorghum crop improvement.

As the predictability of groundwater and precipitation in the U.S. Sorghum Belt shifts with the changing climate, the interest in integrating drought adaptive traits into breeding lines has greatly increased. While BTx642 has remained a favored line in breeding, identifying stay-green alleles in other converted germplasm is a necessary step to diversify the genetic basis of the trait in U.S. grain sorghum.

Unresolved Needs

While the SCP has provided valuable sources for traits that have positively-impacted US sorghum production, there are still gaps and issues to be addressed. The erasure of most or all of

the tropical background from chromosome 6 in converted lines (R. Klein et al., 2008) and the need to find new fertility restorers and male-sterile complements in the CMS seed production system (Kante et al., 2018) are both important considerations for the development and future of SCP lines.

The absence of marker-assisted breeding tools in the original SCP and the linkage disequilibrium between large-effect maturity and dwarfing loci on chromosome 6 resulted in the loss of up to 10% of the exotic diversity in many converted lines (R. Klein et al., 2008; Thurber et al., 2013a). Crucial alleles for traits of interest may exist in the chromosome 6 exotic background, and their unintentional exclusion from released converted lines hinders marker development trait architecture characterization (Thurber et al., 2013a). Gene flow between these adapted converted lines and the original tropical germplasm could recover some of the lost exotic haplotypes on chromosome 6. While the Reinstated Sorghum Conversion Program employs marker-assisted or genomic selection in all new conversions, revisiting registered materials from the original SCP could prove useful. Recovering the chromosome 6 exotic background in those materials could greatly broaden the variation for known and novel traits valued in temperate breeding.

In the 3-line CMS crossing scheme, heavy reliance on one cytoplasm can be costly if the lines developed with that cytoplasm all share susceptibility to a pathogen or abiotic stress (Murty, 1986). Sorghum hybrid production is heavily reliant on the A1 cytoplasm developed from Milo. The A2 cytoplasm, developed from Kafir, has received some attention, but the A3 and A4 cytoplasms are generally limited in breeding applications due to lack of elite restorer lines and low stability of restoration (Kante et al., 2018). This risk of susceptibility is perhaps best illustrated in the devastation of U.S. hybrid seed corn production by the fungal disease

southern corn leaf blight (SCLB) in the early 1970s. At the time, between 75 and 90% of hybrid seed corn in the U.S. was grown in the same cms-T cytoplasmic background, which suddenly became susceptible to SCLB via a spontaneous mutation in the pathogen. More than $1 billion USD of losses were incurred from destruction of 15% of the U.S. acreage (Bruns, 2017). The feasible use of alternate cytoplasms in sorghum hybrids could be addressed in part by screening for candidate A/B and R-line pairs in converted sorghum germplasm (Moran & Rooney, 2003; Murty, 1986).

Conclusions

The Sorghum Conversion Program embarked on a goal to significantly expand the genetic diversity of germplasm available for use in temperate sorghum breeding (Stephens et al., 1967). Converted germplasm has been widely adopted in breeding programs with use in the pedigrees of nearly all hybrids released since the 1970s (Smith et al., 2010). Converted germplasm has proven valuable in breeding for insect tolerance (Wang et al., 2013), disease resistance (Erpelding, 2010), drought tolerance (Xu et al., 2000), and other traits. The loss of diversity on much of chromosome 6 from conversions pre-dating marker-assisted breeding creates opportunity to recover those exotic haplotypes and investigate potentially adaptive loci (Thurber et al., 2013). As demonstrated by the SCP, target zone adaptation of PGR is imperative to using diverse genetic materials in crop improvement (Sharma et al., 2013). Broadening the genetic basis of our crops with the help of pre-breeding creates durable and productive agriculture systems.

REFERENCES

Abdulai, A. L., Kouressy, M., Vaksmann, M., Asch, F., Giese, M., & Holger, B. (2012). Latitude and Date of Sowing Influences Phenology of Photoperiod-Sensitive Sorghums. *Journal of Agronomy and Crop Science*, *198*(5), 340–348. https://doi.org/10.1111/j.1439-037X.2012.00523.x

Allier, A., Teyssèdre, S., Lehermeier, C., Moreau, L., & Charcosset, A. (2020). Optimized breeding strategies to harness genetic resources with different performance levels. *BMC Genomics*, *21*(1), 349. https://doi.org/10.1186/s12864-020-6756-0

Ananda, G. K. S., Myrans, H., Norton, S. L., Gleadow, R., Furtado, A., & Henry, R. J. (2020). Wild Sorghum as a Promising Resource for Crop Improvement. *Frontiers in Plant Science*, *11*, 1108. https://doi.org/10.3389/fpls.2020.01108

Aono, A. H., Pimenta, R. J. G., Garcia, A. L. B., Correr, F. H., Hosaka, G. K., Carrasco, M. M., Cardoso-Silva, C. B., Mancini, M. C., Sforça, D. A., dos Santos, L. B., Nagai, J. S., Pinto, L. R., Landell, M. G. de A., Carneiro, M. S., Balsalobre, T. W., Quiles, M. G., Pereira, W. A., Margarido, G. R. A., & de Souza, A. P. (2021). The Wild Sugarcane and Sorghum Kinomes: Insights Into Expansion, Diversification, and Expression Patterns. *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/articles/10.3389/fpls.2021.668623

Armstrong, J. S., Rooney, W. L., Peterson, G. C., Villenueva, R. T., Brewer, M. J., & Sekula-Ortiz, D. (2015). Sugarcane Aphid (Hemiptera: Aphididae): Host Range and Sorghum Resistance Including Cross-Resistance From Greenbug Sources. *Journal of Economic Entomology*, *108*(2), 576–582. https://doi.org/10.1093/jee/tou065

Boora, K. S., Frederiksen, R., & Magill, C. (1998). DNA-Based Markers for a Recessive Gene Conferring Anthracnose Resistance in Sorghum. *Crop Science*, *38*(6), cropsci1998.0011183X003800060048x. https://doi.org/10.2135/cropsci1998.0011183X003800060048x

Borrell, A. K., Mullet, J. E., George-Jaeggli, B., van Oosterom, E. J., Hammer, G. L., Klein, P. E., & Jordan, D. R. (2014). Drought adaptation of stay-green sorghum is associated with canopy development, leaf anatomy, root growth, and water uptake. *Journal of Experimental Botany*, *65*(21), 6251–6263. https://doi.org/10.1093/jxb/eru232

Bowers, J. E., Abbey, C., Anderson, S., Chang, C., Draye, X., Hoppe, A. H., Jessup, R., Lemke, C., Lennington, J., Li, Z., Lin, Y.-R., Liu, S.-C., Luo, L., Marler, B. S., Ming, R., Mitchell, S. E., Qiang, D., Reischmann, K., Schulze, S. R., … Paterson, A. H. (2003). A high-density genetic recombination map of sequence-tagged sites for sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. *Genetics*, *165*(1), 367–386.

Bruns, H. A. (2017). Southern Corn Leaf Blight: A Story Worth Retelling. *Agronomy Journal*, *109*(4), 1218–1224. https://doi.org/10.2134/agronj2017.01.0006

Burrell, A. M., Sharma, A., Patil, N. Y., Collins, S. D., Anderson, W. F., Rooney, W. L., & Klein, P. E. (2015). Sequencing of an Anthracnose-Resistant Sorghum Genotype and Mapping of a Major QTL Reveal Strong Candidate Genes for Anthracnose Resistance. *Crop Science*, *55*(2), 790–799. https://doi.org/10.2135/cropsci2014.06.0430

Casto, A. L., Mattison, A. J., Olson, S. N., Thakran, M., Rooney, W. L., & Mullet, J. E. (2019). Maturity2, a novel regulator of flowering time in Sorghum bicolor, increases expression of SbPRR37 and SbCO in long days delaying flowering. *PLoS ONE*, *14*(4), e0212154.

https://doi.org/10.1371/journal.pone.0212154

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., & Coupland, G. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science (New York, N.Y.)*, *316*(5827), 1030–1033. https://doi.org/10.1126/science.1141752

Cowling, W. A., Buirchell, B. J., Falk, D. E., Cowling, W. A., Buirchell, B. J., & Falk, D. E. (2009). A model for incorporating novel alleles from the primary gene pool into elite crop breeding programs while reselecting major genes for domestication or adaptation. *Crop and Pasture Science*, *60*(10), 1009–1015. https://doi.org/10.1071/CP08223

Cruet-Burgos, C. M., Cuevas, H. E., Prom, L. K., Knoll, J. E., Stutts, L. R., & Vermerris, W. (2020). Genomic Dissection of Anthracnose (Colletotrichum sublineolum) Resistance Response in Sorghum Differential Line SC112-14. *G3: Genes|Genomes|Genetics*, *10*(4), 1403–1412. https://doi.org/10.1534/g3.120.401121

Cuevas, H. E., & Prom, L. K. (2020). Evaluation of genetic diversity, agronomic traits, and anthracnose resistance in the NPGS Sudan Sorghum Core collection. *BMC Genomics*, *21*(1), 88. https://doi.org/10.1186/s12864-020-6489-0

Cuevas, H. E., Prom, L. K., Cooper, E. A., Knoll, J. E., & Ni, X. (2018). Genome-Wide Association Mapping of Anthracnose (Colletotrichum sublineolum) Resistance in the U.S. Sorghum Association Panel. *The Plant Genome*, *11*(2), 170099. https://doi.org/10.3835/plantgenome2017.11.0099

Erpelding, J. E. (2010). Anthracnose Resistance in Sorghum Breeding Lines Developed from Ethiopian Germplasm. *Plant Health Progress*, *11*(1), 3. https://doi.org/10.1094/PHP-2010-1123-02-RS

FAOSTAT. (n.d.). Retrieved December 11, 2021, from https://www.fao.org/faostat/en/#data

Gepts, P. (2006). Plant Genetic Resources Conservation and Utilization: The Accomplishments and Future of a Societal Insurance Policy. *Crop Science*, *46*(5), 2278–2292. https://doi.org/10.2135/cropsci2006.03.0169gas

Gwary, D. M., Rabo, T. D., & Anaso, A. B. (2004). The development of anthracnose symptoms on sorghum genotypes in the Nigerian savanna / Die Entwicklung von Anthraknosesymptomen an Hirse-Genotypen in der nigerianischen Savanne. *Zeitschrift Für Pflanzenkrankheiten Und Pflanzenschutz / Journal of Plant Diseases and Protection*, *111*(1), 96–103.

Hadley, H. H. (1957). An Analysis of Variation in Height in Sorghum1. *Agronomy Journal*, *49*(3), 144–147. https://doi.org/10.2134/agronj1957.00021962004900030010x

Hao, H., Li, Z., Leng, C., Lu, C., Luo, H., Liu, Y., Wu, X., Liu, Z., Shang, L., & Jing, H.-C. (2021). Sorghum breeding in the genomic era: Opportunities and challenges. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 1–26. https://doi.org/10.1007/s00122-021-03789-z

Harlan, J. R., & Stemler, A. (2011). The Races of Sorghum in Africa. In *The Races of Sorghum in Africa* (pp. 465–478). De Gruyter Mouton. https://doi.org/10.1515/9783110806373.465

Harlan, J. R., & Wet, J. M. J. de. (1971). Toward a Rational Classification of Cultivated Plants. *TAXON*, *20*(4), 509–517. https://doi.org/10.2307/1218252

Higgins, R. H., Thurber, C. S., Assaranurak, I., & Brown, P. J. (2014). Multiparental mapping of plant height and flowering time QTL in partially isogenic sorghum families. *G3 (Bethesda, Md.)*, *4*(9), 1593–1602. https://doi.org/10.1534/g3.114.013318

Hilley, J. L., Weers, B. D., Truong, S. K., McCormick, R. F., Mattison, A. J., McKinley, B. A., Morishige, D. T., & Mullet, J. E. (2017). Sorghum Dw2 Encodes a Protein Kinase Regulator of Stem Internode Length. *Scientific Reports*, *7*(1), 4616. https://doi.org/10.1038/s41598-017-04609-5

J Felderhoff, T., M McIntyre, L., Saballos, A., & Vermerris, W. (2016). Using Genotyping by Sequencing to Map Two Novel Anthracnose Resistance Loci in Sorghum bicolor. *G3 (Bethesda, Md.)*, *6*(7), 1935–1946. https://doi.org/10.1534/g3.116.030510

Jordan, D. R., Mace, E. S., Cruickshank, A. W., Hunt, C. H., & Henzell, R. G. (2011). Exploring and Exploiting Genetic Variation from Unadapted Sorghum Germplasm in a Breeding Program. *Crop Science*, *51*(4), 1444–1457. https://doi.org/10.2135/cropsci2010.06.0326

Kante, M., Rattunde, H. F. W., Nébié, B., Weltzien, E., Haussmann, B. I. G., & Leiser, W. L. (2018). QTL mapping and validation of fertility restoration in West African sorghum A1 cytoplasm and identification of a potential causative mutation for Rf2. *Theoretical and Applied Genetics*, *131*(11), 2397–2412. https://doi.org/10.1007/s00122-018-3161-z

Kassahun, B., Bidinger, F. R., Hash, C. T., & Kuruvinashetti, M. S. (2010). Stay-green expression in early generation sorghum [Sorghum bicolor (L.) Moench] QTL introgression lines. *Euphytica*, *172*(3), 351–362. https://doi.org/10.1007/s10681-009-0108-0

Klein, R., Miller, F., Bean, S., & Klein, P. (2015a). Registration of 40 Converted Germplasm Sources from the Reinstated Sorghum Conversion Program. *Journal of Plant Registrations*, *10*. https://doi.org/10.3198/jpr2015.05.0034crg

Klein, R., Miller, F., Bean, S., & Klein, P. (2015b). Registration of 40 Converted Germplasm Sources from the Reinstated Sorghum Conversion Program. *Journal of Plant*

Registrations, *10*. https://doi.org/10.3198/jpr2015.05.0034crg

Klein, R., Mullet, J., Jordan, D., Miller, F., Rooney, W., Menz, M., Franks, C. D., & Klein, P. (2008). The Effect of Tropical Sorghum Conversion and Inbred Development on Genome Diversity as Revealed by High-Resolution Genotyping. *Crop Science*, *48*. https://doi.org/10.2135/cropsci2007.06.0319tpg

Klein, R. R., Miller, F. R., Klein, P. E., & Burke, J. J. (2013). Registration of Partially Converted Germplasm from 44 Accessions of the USDA-ARS Ethiopian and Sudanese Sorghum Collections. *Journal of Plant Registrations*, *7*(3), 368–372. https://doi.org/10.3198/jpr2012.08.0025crgs

Li, X., Li, X., Fridman, E., Tesso, T. T., & Yu, J. (2015). Dissecting repulsion linkage in the dwarfing gene Dw3 region for sorghum plant height provides insights into heterosis. *Proceedings of the National Academy of Sciences*, *112*(38), 11823–11828. https://doi.org/10.1073/pnas.1509229112

Lin, Y. R., Schertz, K. F., & Paterson, A. H. (1995). Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. *Genetics*, *141*(1), 391–411. https://doi.org/10.1093/genetics/141.1.391

Mahalakshmi, V., & Bidinger, F. R. (2002). Evaluation of stay-green sorghum germplasm lines at ICRISAT. *Crop Science*, *42*(3), 965–974.

Mbulwe, L., Peterson, G. C., Scott-Armstrong, J., & Rooney, W. L. (2016). Registration of Sorghum Germplasm Tx3408 and Tx3409 with Tolerance to Sugarcane Aphid [Melanaphis sacchari (Zehntner)]. *Journal of Plant Registrations*, *10*(1), 51–56. https://doi.org/10.3198/jpr2015.04.0025crg

Medina, R. F., Armstrong, S. J., & Harrison, K. (2017). Genetic population structure of

sugarcane aphid, Melanaphis sacchari, in sorghum, sugarcane, and Johnsongrass in the continental USA. *Entomologia Experimentalis et Applicata*, *162*(3), 358–365. https://doi.org/10.1111/eea.12547

Monk, R., Franks, C., & Dahlberg, J. (2014). Sorghum. In *Yield Gains in Major U.S. Field Crops* (pp. 293–310). John Wiley & Sons, Ltd. https://doi.org/10.2135/cssaspecpub33.c11

Moran, J., & Rooney, W. (2003). Effect of Cytoplasm on the Agronomic Performance of Grain Sorghum Hybrids. *Crop Science - CROP SCI*, *43*. https://doi.org/10.2135/cropsci2003.0777

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., & Kresovich, S. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences*, *110*(2), 453–458. https://doi.org/10.1073/pnas.1215985110

Muleta, K. T., Felderhoff, T., Winans, N., Walstead, R., Charles, J. R., Armstrong, J. S., Mamidi, S., Plott, C., Vogel, J. P., Lemaux, P. G., Mockler, T. C., Grimwood, J., Schmutz, J., Pressoir, G., & Morris, G. P. (2021). *The recent evolutionary rescue of a staple crop depended on over half a century of global germplasm exchange* (p. 2021.05.11.443651). https://doi.org/10.1101/2021.05.11.443651

Mundia, C. W., Secchi, S., Akamani, K., & Wang, G. (2019). A Regional Comparison of Factors Affecting Global Sorghum Production: The Case of North America, Asia and Africa's Sahel. *Sustainability*, *11*(7), 2135. https://doi.org/10.3390/su11072135

Murphy, R. L., Morishige, D. T., Brady, J. A., Rooney, W. L., Yang, S., Klein, P. E., & Mullet, J. E. (2014). Ghd7 (Ma6) Represses Sorghum Flowering in Long Days: Ghd7 Alleles

Enhance Biomass Accumulation and Grain Production. *The Plant Genome*, *7*(2), plantgenome2013.11.0040. https://doi.org/10.3835/plantgenome2013.11.0040

Murty, U. R. (1986). MILO AND NON-MILO SOURCES OF CYTOPLASMS IN SORGHUM BICOLOR (L) MOENCH. II. FERTILITY RESTORERS AND STERILITY MAINTAINERS ON NON-MILO CYTOPLASMS. *Cereal Research Communications*, *14*(2), 191–196.

Nevo, E., Zohary, D., Brown, A. H. D., & Haber, M. (1979). Genetic Diversity and Environmental Associations of Wild Barley, Hordeum spontaneum, in Israel. *Evolution*, *33*(3), 815–833. https://doi.org/10.2307/2407648

Papa, R., Acosta, J., Delgado-Salinas, A., & Gepts, P. (2005). A genome-wide analysis of differentiation between wild and domesticated Phaseolus vulgaris from Mesoamerica. *Theoretical and Applied Genetics*, *111*(6), 1147–1158. https://doi.org/10.1007/s00122-005-0045-9

Patil, N. Y., Klein, R. R., Williams, C. L., Collins, S. D., Knoll, J. E., Burrell, A. M., Anderson, W. F., Rooney, W. L., & Klein, P. E. (2017). Quantitative Trait Loci Associated with Anthracnose Resistance in Sorghum. *Crop Science*, *57*(2), 877–890. https://doi.org/10.2135/cropsci2016.09.0793

Paudyal, S., Armstrong, J. S., Giles, K. L., Hoback, W., Aiken, R., & Payton, M. E. (2020). Differential responses of sorghum genotypes to sugarcane aphid feeding. *Planta*, *252*(1), 14. https://doi.org/10.1007/s00425-020-03419-w

Perumal, R., Menz, M., Mehta, P., Katilé, S., Gutierrez-Rojas, L., Klein, R., Klein, P., Prom, L., Schlueter, J., Rooney, W., & Magill, C. (2009). Molecular mapping of Cg1, a gene for resistance to anthracnose (Colletotrichum sublineolum) in sorghum. *Euphytica*, *165*.

https://doi.org/10.1007/s10681-008-9791-5

Quinby, J. R. (1967). The Maturity Genes of Sorghum. In A. G. Norman (Ed.), *Advances in Agronomy* (Vol. 19, pp. 267–305). Academic Press. https://doi.org/10.1016/S0065-2113(08)60737-3

Quinby, J. R. (1974). *Sorghum improvement and the genetics of growth*. Texas A&M University Press.

Quinby, J. R., & Karper, R. E. (1953). Inheritance of height in sorghum. *Inheritance of Height in Sorghum.* http://www.cabdirect.org/cabdirect/abstract/19541603750

Ramanatha Rao, V., & Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture*, *68*(1), 1–19. https://doi.org/10.1023/A:1013359015812

Rosenow, D. T. (1978). Stalk rot resistance breeding in Texas. *Sorghum Diseases: A World Review-Proc. International Workshop on Sorghum Diseases, Hyderabad, India*, 306–314.

Rosenow, D. T. (1986). Released converted lines: Release of 240 converted exotic lines from the world sorghum collection. *Sorghum Newsletter*, *29*, 97.

Sagnard, F., Deu, M., Dembélé, D., Leblois, R., Touré, L., Diakité, M., Calatayud, C., Vaksmann, M., Bouchet, S., Mallé, Y., Togola, S., & Traoré, P. C. S. (2011). Genetic diversity, structure, gene flow and evolutionary relationships within the Sorghum bicolor wild–weedy–crop complex in a western African region. *Theoretical and Applied Genetics*, *123*(7), 1231. https://doi.org/10.1007/s00122-011-1662-0

Sanchez, A. C., Subudhi, P. K., Rosenow, D. T., & Nguyen, H. T. (2002). Mapping QTLs associated with drought resistance in sorghum (Sorghum bicolor L. Moench). *Plant*

Molecular Biology, *48*(5), 713–726. https://doi.org/10.1023/A:1014894130270

Sharma, S., Upadhyaya, H. D., Varshney, R. K., & Gowda, C. L. L. (2013). Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Frontiers in Plant Science*, *4*. https://doi.org/10.3389/fpls.2013.00309

Singh, K., Kumar, S., Subramani, R., Singh, M., & Gupta, K. (2019). Plant genetic resources management and pre-breeding in genomics era. *Indian Journal of Genetics and Plant Breeding (The)*, *79*. https://doi.org/10.31742/IJGPB.79S.1.1

Singh, S. P. (1985). Sources of cold tolerance in grain sorghum. *Canadian Journal of Plant Science = Revue Canadienne de Phytotechnie*. https://scholar.google.com/scholar_lookup?title=Sources+of+cold+tolerance+in+grain+sorghum&author=Singh%2C+S.P.&publication_year=1985

Smith, C. W., & Frederiksen, R. A. (2000). *Sorghum: Origin, History, Technology, and Production*. John Wiley & Sons.

Smith, S., Primomo, V., Monk, R., Nelson, B., Jones, E., & Porter, K. (2010). Genetic Diversity of Widely Used U.S. Sorghum Hybrids 1980–2008. *Crop Science*, *50*(5), 1664–1673. https://doi.org/10.2135/cropsci2009.10.0619

Stamenković, O. S., Siliveru, K., Veljković, V. B., Banković-Ilić, I. B., Tasić, M. B., Ciampitti, I. A., Đalović, I. G., Mitrović, P. M., Sikora, V. Š., & Prasad, P. V. V. (2020). Production of biofuels from sorghum. *Renewable and Sustainable Energy Reviews*, *124*, 109769. https://doi.org/10.1016/j.rser.2020.109769

Stephens, J. C., Miller, F. R., & Rosenow, D. T. (1967). Conversion of Alien Sorghums to Early Combine Genotypes1. *Crop Science*, *7*(4), cropsci1967.0011183X000700040036x. https://doi.org/10.2135/cropsci1967.0011183X000700040036x

Stinchcombe, J. R., Weinig, C., Ungerer, M., Olsen, K. M., Mays, C., Halldorsdottir, S. S., Purugganan, M. D., & Schmitt, J. (2004). A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA. *Proceedings of the National Academy of Sciences*, *101*(13), 4712–4717. https://doi.org/10.1073/pnas.0306401101

Sukumaran, S., Li, X., Li, X., Zhu, C., Bai, G., Perumal, R., Tuinstra, M. R., Prasad, P. V. V., Mitchell, S. E., Tesso, T. T., & Yu, J. (2016). QTL Mapping for Grain Yield, Flowering Time, and Stay-Green Traits in Sorghum with Genotyping-by-Sequencing Markers. *Crop Science*, *56*(4), 1429–1442. https://doi.org/10.2135/cropsci2015.02.0097

Tanksley, S. D., & Nelson, J. C. (1996). Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *92*(2), 191–203. https://doi.org/10.1007/BF00223376

Thomas, H., & Howarth, C. J. (2000). Five ways to stay green. *Journal of Experimental Botany*, *51 Spec No*, 329–337. https://doi.org/10.1093/jexbot/51.suppl_1.329

Thurber, C. S., Ma, J. M., Higgins, R. H., & Brown, P. J. (2013a). Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biology*, *14*(6), R68. https://doi.org/10.1186/gb-2013-14-6-r68

Thurber, C. S., Ma, J. M., Higgins, R. H., & Brown, P. J. (2013b). Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biology*, *14*(6), R68. https://doi.org/10.1186/gb-2013-14-6-r68

Vadez, V., Deshpande, S., Kholova, J., Ramu, P., & Hash, C. T. (2013). *Molecular Breeding for Stay-Green: Progress and Challenges in Sorghum*. 125–141.

https://doi.org/10.1002/9781118728482.ch8

Wang, F., Zhao, S., Han, Y., Yutao, S., Dong, Z., Gao, Y., Zhang, K., Liu, X., Li, D., Chang, J., & Wang, D. (2013). Efficient and fine mapping of RMES1 conferring resistance to sorghum aphid Melanaphis sacchari. *Molecular Breeding*, *31*. https://doi.org/10.1007/s11032-012-9832-6

Xing-Lin, H., De-Liang, W., Wu-Jiu, Z., & Shi-Ru, J. (2017). The production of the Chinese baijiu from sorghum and other cereals. *Journal of the Institute of Brewing*, *123*. https://doi.org/10.1002/jib.450

Xu, W., Subudhi, P. K., Crasta, O. R., Rosenow, D. T., Mullet, J. E., & Nguyen, H. T. (2000). Molecular mapping of QTLs conferring stay-green in grain sorghum (Sorghum bicolor L. Moench). *Genome*, *43*(3), 461–469.

Zongo, J., Gouyon, P.-H., Sarr, A., & Sandmeier, M. (2005). Genetic Diversity and Phylogenic Relations Among Sahelian Sorghum Accessions. *Genetic Resources and Crop Evolution*, *52*, 869–878. https://doi.org/10.1007/s10722-003-6091-8

CHAPTER II: DEVELOPMENT AND CHARACTERIZATION OF TWO BIPARENTAL MAPPING FAMILIES FOR GENETIC ANALYSES OF WATER-USE DYNAMICS IN SORGHUM

INTRODUCTION

Precipitation patterns are shifting due to climate change, placing dryland cropping systems at risk of drought timing (Dai et al., 2018). Rainfed systems are reliant on precipitation only for water inputs, placing breeding for drought-resilient crops at the forefront of managing drought stresses. Sorghum (*Sorghum bicolor* (L.) Moench) is a drought-resilient African crop valued for animal feed and bioenergy in the United States (Stamenković et al., 2020), with genetic diversity untapped for drought tolerance traits. Plant genetic resources (PGR) offer potential sources of allelic diversity for traits that influence characteristics of plant water-use, and in turn facilitate more stable yields under water-limited conditions (Maxted et al., 2016). Diverse genetic materials, however, present unique considerations for drought adaptive trait discovery.

Feasibility of using diverse PGR for trait discovery is impacted by differences in target zone adaptation, potential gene pool incompatibilities, undesirable linkage disequilibrium, and phenotypic covariates of trait mapping (Sharma et al., 2013). Pre-breeding is used to overcome some of these challenges. The Sorghum Conversion Program (SCP) was a pre-breeding program that delivered hundreds of diverse tropical African lines adapted to U.S. height and maturity breeding specifications while retaining the exotic genome outside of the conversion regions (Klein et al., 2008; Stephens et al., 1967). The semi-dwarf, photoperiod-insensitive lines produced by the SCP allow diverse sorghum genetic resources to be evaluated in temperate target environments for novel and existing traits of interest.

Limited transpiration (LT) is a water-use trait in sorghum being explored for trait technology development (Sadok et al., 2021). As observed in other crops like maize (Gholipoor et al., 2013) and peanuts (Devi et al., 2010), sorghum lines with the LT trait are expected to limit their transpiration rate when vapor pressure deficit (VPD) is high in the environment, reducing plant water demand during very dry conditions (Sinclair et al., 2005) (Figure 2.1). This response is modeled to produce a potential five percent yield increase in water-limited environments by conserving soil water during the vegetative growth stage to be available during grain fill (Raymundo et al., in preparation). A breakpoint in transpiration rate of sorghum genotypes was first observed in growth chamber experiments (Gholipoor et al., 2010; Shekoofa et al., 2014), but the trait requires genetic characterization under relevant field conditions to develop molecular markers and establish a trait introgression strategy for pre-breeding.

This field-based trait mapping study in a well-watered environment uses canopy temperature as an LT phenotype such that differences in canopy temperature reflect differences in transpiration rate (Inoue et al., 1990; G. J. Rebetzke et al., 2012). Lines with the LT trait are predicted to produce a hotter canopy temperature due to the reduction in transpiration cooling effect than those without the trait. The temporal variability in canopy temperature prompts the need for a fast, high-throughput approach to collect thermal data (Turner et al., 2012). Unoccupied aircraft system (UAS)-based aerial thermal imaging at the canopy closure stage (but prior to anthesis) during the seasonal and daily periods of highest VPD provides the capacity to collect whole-field data in minutes, reducing temporal variability (Hou et al., 2021; Pignon et al., 2021; G. J. Rebetzke et al., 2012). However, aerial thermal imaging for water-use trait studies is easily confounded by phenotypic covariates of agronomic background.

Trait mapping combines genotype data and phenotype data to elucidate regions of the genome that may be underlying the trait of interest (Otto & Jones, 2000). To map water-use traits, single nucleotide polymorphism (SNP) genotype data can be combined with phenotype data such as stomatal conductance or canopy temperature. Agronomic traits like plant height and flowering time are critical to assess when selecting parental germplasm for water-use trait mapping. Height and flowering time can affect how the study plants experience managed drought stress, claim environmental resources, and produce phenotypic covariates (G. Rebetzke et al., 2008; G. J. Rebetzke et al., 2008). A water-use trait mapping population composed of germplasm with substantially different heights can result in plots of genetically-taller plants dominating the light and soil water resources while genetically-shorter plants are shaded and outcompeted for available soil water (G. J. Rebetzke et al., 2008) (Figure 2.2A). Substantial differences in flowering time can result in plots of early-flowering plants experiencing less drought stress than late-flowering plants, given stored soil water diminishes throughout the season (Faye et al., 2022; Raymundo et al., 2021). Additionally, differences in flowering time create covariates that confound phenotyping using physiological proxies. For example, canopy temperatures are falsely inflated in flowering plots as the panicle does not contribute significantly to gas exchange and the transpiration cooling effect of the canopy is masked (Chang et al., 2020; Girma & Krieg, 1992) (Figure 2.2B). Therefore, establishing uniform height and flowering time across a mapping population is important for water-use trait mapping (Deery et al., 2016).

In U.S. sorghum, height variation is controlled by four main additive loci, *Dw*$_1$-*Dw*$_4$, with tallness dominant and maturity variation is controlled by six main loci, *Ma*$_1$-*Ma*$_6$, with lateness dominant (Murphy et al., 2014; Quinby, 1974). Lines dominant at all four major height loci will

typically grow upwards of 16 feet. Lines dominant at all major maturity loci will typically flower later than 90 days after emergence (Quinby, 1974). In the SCP, height and maturity haplotypes from elite US breeding line BTx406 were targeted for introgression into nearly all exotic lines (Quinby, 1967). Theoretically, this approach using a common donor parent should result in isogenic height and flowering time adaptation across converted germplasm. However, the original SCP predates development of marker-assisted selection tools, and instead used phenotypic selection in backcross generations. Phenotypic selection is unreliable in ensuring that the same genetic haplotypes are selected in each line or generation, or that those haplotypes are identical by descent (Klein et al., 2008; Thurber et al., 2013). Selecting parental germplasm for a water-use trait mapping study therefore requires careful consideration of height and flowering time of candidate lines.

Previously-developed sorghum mapping populations did not account for these specific agronomic trait covariates, and are therefore less suitable for LT and other water-use trait mapping (Bouchet et al., 2017; Boyles et al., 2017; Perumal et al., 2021). Testing the hypotheses that the mapping families are (1) ideal, (2) acceptable, or (3) unacceptable for UAS-based thermal imaging across the mapping families allows us to establish the degree to which the mapping families control for associated phenotypic covariates. In addition to characterizing the influence of potential covariates on the thermal data obtained from LT mapping populations, this furthers our knowledge on the power and utility of the families in LT and other water-use genetic studies.

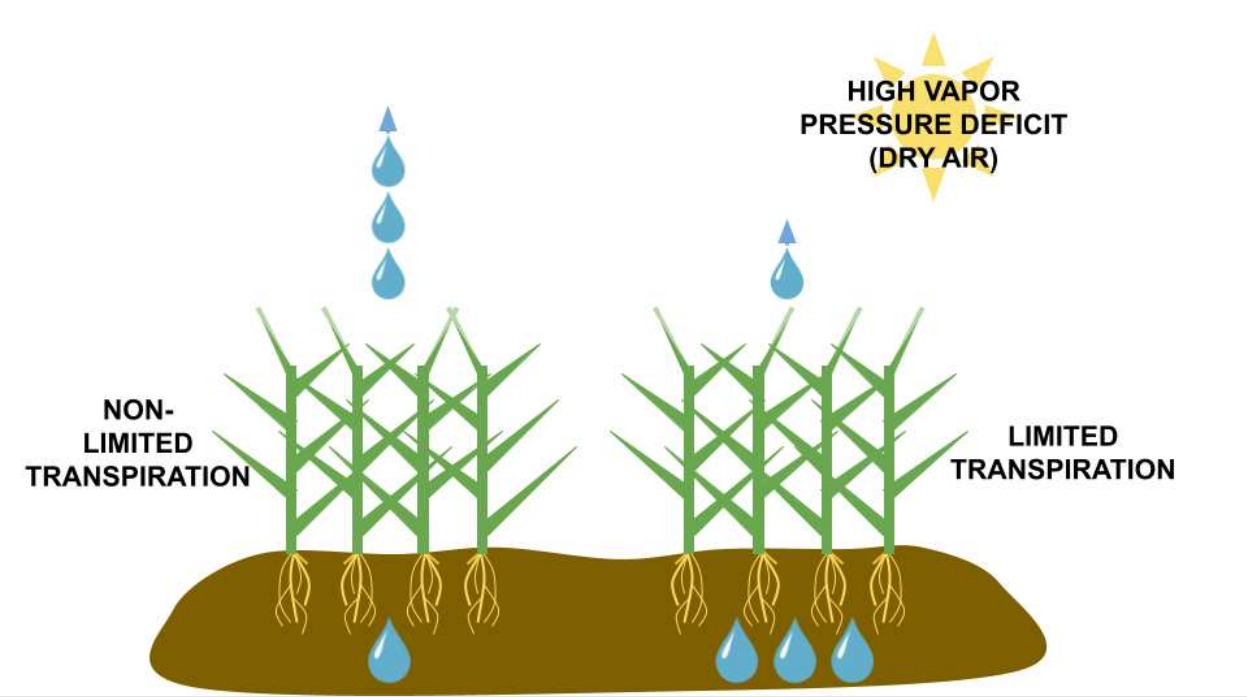


Figure 2.1. Conceptual illustration of a snapshot of the limited transpiration trait in sorghum. While experiencing high vapor pressure deficit during the vegetative stage, LT genotypes will reduce transpiration rate (water drops above plants) and conserve soil water. Non-LT genotypes will continue to transpire and deplete soil water resources at a constant rate. Note, the illustration of water availability differences represents the expected expression of the trait in a production environment. In the experimental LT phenotyping system, all genotypes are well-watered to circumvent confounding effect of earlier water use differences.

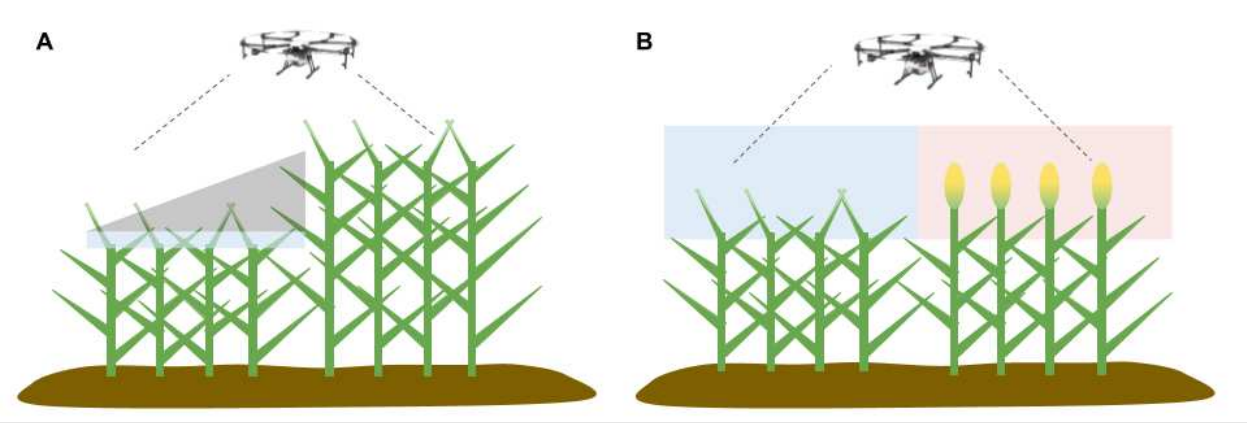


Figure 2.2. Conceptual illustration of phenotypic covariates produced by non-uniform height and flowering time in a genetic mapping population phenotype using UAS thermal imaging. (A) Plots with genetically-taller plant height will shade neighboring plots, artifactually reducing canopy temperatures. (B) Plots with genetically-early flowering (yellow ovals) will show inflated canopy temperatures. The thermal imaging will capture the panicle foremost over the canopy, where the lack of gas exchange raises the captured temperature.

MATERIALS AND METHODS

Design of Mapping Families

The limited transpiration sorghum mapping population includes two families developed from parental lines BTx2752 (PI 656018), SC979 (PI 576428), and RTx430 (PI 655996). The two families were created from RTx430 × SC979 and RTx430 × BTx2752 crosses. SC979 and BTx2752 were selected as putative LT parents based on their agronomic characteristics being uniform, meeting U.S. breeding standards, performance in previous LT growth chamber studies (Gholipoor et al., 2010; Shekoofa et al., 2014), and preliminary stomatal conductance evaluations. RTx430 was selected as a putative non-LT parent for its agronomic background uniform with the putative LT parents and its widespread industry usage as a pollinator parent within the cytoplasmic male sterility (CMS) hybrid breeding system (Sahoo et al., 2010). The timeline and development of the families is detailed below.

Parent lines for LT mapping families were aiming to capture genetic variation for LT while establishing a uniform, agronomically relevant background suitable for water use phenotyping and commercial sorghum breeding. Evaluation of candidate parent lines began in 2018, where approximately 40 lines were grown and measured for agronomic characteristics like plant height, days to flowering, and tillering (Raymundo et al. in prep). These candidate lines were chosen based on inclusion in prior LT studies assessing transpiration rate breakpoints in growth chambers or field conditions (Choudhary et al., 2013; Gholipoor et al., 2010; Riar et al., 2015; Shekoofa et al., 2014). Any lines not close to target U.S. specifications for these agronomic characteristics were removed from further consideration. Twelve lines were advanced to preliminary studies in 2019 and 2020. Those studies continued evaluating relevance of agronomic background while also testing stomatal conductance and canopy temperature of the lines as proxies of transpiration rate under well-watered conditions.

Analyses of the preliminary studies showed consistency with a lower stomatal conductance and higher canopy temperature responses as well as suitable agronomic background in lines SC979 and BTx2752. Therefore, these lines were chosen as putative LT parents for mapping family development. SC979 is a fully-converted Ethiopian durra-type inbred line delivered by the SCP. BTx2752 is a male inbred line released by the Texas A&M University sorghum breeding program. The putative non-LT parent RTx430 was selected for its consistent higher stomatal conductance and lower canopy temperature in addition to widespread adoption as an excellent fertility restorer in industry.

Mapping Family Development

The goal of the LT mapping families was recombinant inbred lines (RILs) at a generation suitable for trait mapping with uniform plant height and flowering time across the families. The

initial crosses and summer generation advancements were carried out at Kansas State University in Manhattan, KS and winter nursery advancements were carried out in Puerto Vallarta, Jalisco, Mexico. All generations were advanced by selfing.

Initial crossing of RTx430 × SC979 was completed in 2014 using plastic bag sterilization for the female parent. This cross, done prior to LT parent evaluation, was made as a preliminary interest from the prior LT literature. In 2015, the lines were advanced from F_1 to F_2. Lines were advanced to F_3 in the 2017-2018 winter nursery, to F_4 in summer 2018, to F_5 in summer 2020, and to F_6 in 2020-2021 winter nursery. Initial crossing of RTx430 × BTx2752 was completed in 2018 for inclusion in a separate diversity panel under development. Lines were advanced to the F_2 in 2018-2019 winter nursery, to F_3 in summer 2019, to F_4 in summer 2020, and to F_5 in 2020-2021 winter nursery. The crosses were officially chosen for use in LT mapping after evaluation of the 2020 preliminary studies.

While mapping family development aims to reduce biased selection and optimize recombinant diversity, some selection pressures were imposed during line advancement to address concerns with water-use trait mapping. First, any lines not within the target height and flowering time specifications were dropped from the families. The goal was to unify the canopies and control for associated phenotypic covariates to increase the quality of thermal image data. Next, some progeny of the RTx430 × SC979 cross produce an unusual “lazy” plant phenotype that do not grow upright (Dong et al., 2013). Any “lazy” lines were dropped from the family. Next, the original RTx430 × BTx2752 cross was made using nuclear male-sterile BTx2752. The progeny segregated for sterility, and sterility was selected against on the basis of being unable to self and advance those lines. Partial sterility was also selected against due to poor seed set, making it difficult to advance those lines without sufficient bulk seed. Finally, there was

selection in each generation to advance plants flowering 1-2 days later than others within a line. This later flowering could offer a slightly extended window for data collection in the targeted period of hot, dry conditions before flowering would confound canopy temperatures.

RIL generations powerful for QTL mapping include F_4 to F_7 (Takuno et al., 2012). The generation of the RILs grown in the mapping population varies from F_4 to F_6 in 2021. Generations are spread due to timing of initial crosses, or poor germination or seed set requiring a line to be left out from a season's generation advancement. The sterility in some RTx430 × BTx2752 progeny ultimately resulted in some unrecoverable lines and therefore fewer total RILs in that family than RILs in the RTx430 × SC979 family. The mapping populations grown in 2021 and 2022 do not include all RILs developed from either family.

Field Design and Management

The mapping population was phenotyped in three locations for one growing season (Figure 2.3). The locations include east-central Colorado (Fort Collins, CO, 104.9948993°W 40.6468871°N; semi-arid climate), western Kansas (Colby, KS, 101.0668376°W 39.3914063°N; semi-arid climate), and eastern Kansas (Manhattan, KS, 96.6366494°W 39.1386254°N; humid-continental climate). The field designs and management for each location are described below.

North-Central Colorado

The LT mapping population grown in Colorado in 2021 contained the three inbred parental lines (SC979, BTx2752, RTx430), four commercial checks (ADV G2275, DKS54-00, 84G62, 86Y89), and 160 RILs as entries. Of the RILs, 84 were entries advanced from the RTx430 × SC979 crosses and 76 were entries advanced from the RTx430 × BTx2752 crosses. The entries were grown in a randomized complete block design (RCBD) containing four randomized complete blocks. Within each randomized complete block, each RIL was replicated

once (4 reps total), each inbred parental line was replicated four times (16 reps total), and each commercial check was replicated between 1 and 6 times.

The entire field area of the mapping population covered 2.6 acres, composed of 25 ranges across 136 rows. Each entry was grown in a four-row plot approximately 3 meters long with 0.75 meter-wide rows. Seed was sown at a density of 80 seeds per row per plot based on the well-watered irrigation of the study. A border of commercial hybrid Pioneer 86Y89 surrounded the entire mapping population in rows 1-4 and 133-136, and in ranges 1 and 25.

Field preparation for the study site began with subsoiling and tilling three months prior to planting the study in mid-June. Mulching and cultivation were completed two months prior to planting. Inputs included an application of 46-0-0 urea applied at a rate of 325 lbs/acre two months prior to planting, followed by herbicide applications for weed control (Medal II EC at 1 pint/A, Roundup PowerMAX at 18 oz/A, Induce at 2.5 fl oz/A) at one month before planting. The full study was planted on June 13th using a SRES Standard Research Planter (Seed Research Equipment Solutions; South Hutchinson, KS). Low soil moisture caused initial poor germination, and thereafter the field was irrigated using a linear move irrigation system every four days for two weeks and germination improved. The study was designated as well-watered, therefore the field was irrigated between 1.00” and 1.50” on the same morning each week depending on temperature, VPD, and soil moisture. Irrigation ceased in late September as environmental stresses lessened and plants began to senescence. The field was harvested by individual plots in early December to obtain yield data.

Western Kansas

The LT mapping population grown in western Kansas in 2021 contained the three inbred parental lines (SC979, BTx2752, RTx430), three commercial checks (ADV G2275, DKS54-00,

84G62), and 160 RILs as entries. Of the RILs, 84 were entries advanced from the RTx430 × SC979 crosses and 76 were entries advanced from the RTx430 × BTx2752 crosses. The entries were grown in a RCBD containing four randomized complete blocks. Within each randomized complete block, each RIL was replicated four times (16 reps total), each inbred parental line was replicated at least eight times (at least 24 reps total), and each commercial check was replicated at least four times (at least 16 reps total).

Due to field space and irrigation availability, the mapping population was grown in one 3.5-acre quadrant of a center-pivot irrigated field. The curved wedge-shape of the quadrant led the randomized complete blocks to be grown across ten rectangular sections (3 sections each containing 60 plots, 7 sections each containing 72 plots) within the quadrant. Each entry was grown in a four-row plot approximately 3 meters long with 0.75-meter-wide rows. Seed was sown at a density of 80 seeds per row based on the well-watered irrigation of the study. A border of commercial hybrid Pioneer 86Y89 was used to fill any remaining space in the quadrant around the 10 field sections.

Field preparation for the study site began with disking two months prior to planting, and completing field sweeps one week prior to planting. Inputs included an application of 200 lbs 32-0-0 nitrogen applied at a rate of 65 gal/Acre a week prior to planting, followed by herbicide applications for weed control (Atrazine 4L at 32 oz/Acre, Brawl II at 32 oz/Acre, Detonate at 8oz/Acre, Buccaneer 5 Extra at 27 oz/Acre) 3 days prior to planting. The full study was planted on June 17th and irrigated with 0.75” the following day using a center-pivot irrigation system. Irrigation resumed the following month on July 29th with applications of 1.25” every 3-4 days until the final application on August 17th. The field was combined in mid-November to obtain yield data on a per-plot basis.

Eastern Kansas

The LT mapping population grown in eastern Kansas in 2021 contained the three inbred parental lines (SC979, BTx2752, RTx430), three commercial checks (ADV G2275, DKS54-00, 84G62), and 131 RILs as entries. Of the RILs, 66 were entries advanced from the RTx430 × SC979 crosses and 65 were entries advanced from the RTx430 × BTx2752 crosses. The entries were grown in a RCBD containing three randomized complete blocks. Within each randomized complete block, each RIL was replicated four times (12 reps total), each inbred parental line was replicated eight times (24 reps total), and each commercial check was replicated four times (12 reps total).

The entire field area of the mapping population covered 1.6 acres, composed of 24 ranges across 84 rows (including border). Each entry was grown in a four-row plot approximately 3 meters long with 0.75-meter-wide rows. Seed was sown at a density of 80 seeds per row per plot based on the well-watered irrigation of the study. A border of commercial hybrid Pioneer 86Y89 surrounded the entire mapping population in rows 1-4 and 81-84, and in ranges 1 and 24.

Field preparation for the study site began with disking three weeks prior to planting. Inputs included a broadcast fertilizer application of UAN (120-0-0) at a rate of 40 gal/Acre applied directly after disking, followed by field cultivation. Pre-emergent herbicide applications for weed control (Explorer at 3 oz/Acre, Brawl II at 22 oz/Acre, Atrazine 4L at 20 oz/Acre) were completed two weeks prior to planting. The full study was planted on June 23rd and supplemented with regular irrigation to ensure no drought stress to the population. The field was combined in November to obtain yield data on a per-plot basis.

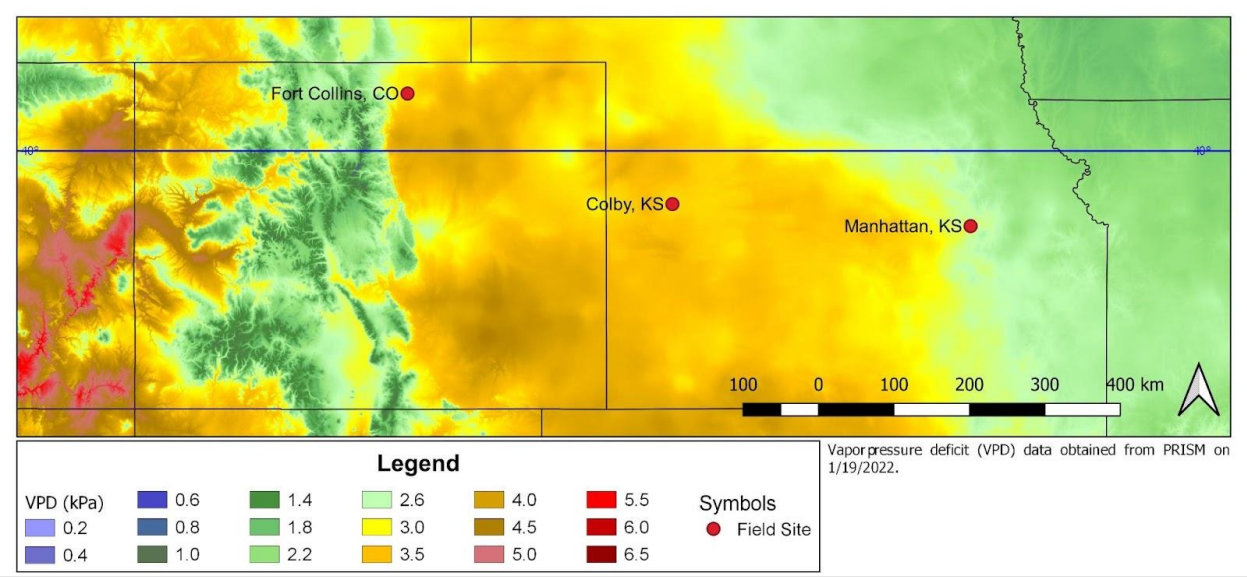


Figure 2.3. Field phenotyping locations for 2021 mapping populations with 30-year average maximum vapor pressure deficit (VPD) data for the LT mapping data collection window (the month of August): Colorado (Fort Collins, CO), western Kansas (Colby, KS), and eastern Kansas (Manhattan, KS).

Height and Flowering Time Phenotyping

The height of each plot was measured at least two weeks after all plots in the study completed anthesis. In this study, height was measured as the distance from the soil surface to the top of the panicle. Each measurement was taken as an approximate average (centimeters) of the center two rows of the four-row plot. Two individuals each independently measured each plot in each location. The height measurements were recorded using a wireless QR code scanner (REALINN) and 240 cm measuring stick with individual 1 cm QR codes corresponding to the height. For each plot, the QR code on the plot ID hang tag was scanned first, then the height of the plot was determined and the corresponding QR code on the measuring stick was scanned.

The flowering time of each plot was determined based on the number of days after emergence to flowering. A plot was marked flowering when fifty percent of the plants in the center two rows of the four-row plot had at least fifty percent of the panicle flowering. The field was surveyed twice a week during the boot stage to identify plots approaching flowering. Once

plots began flowering, the entire field was evaluated every other day and each plot was assessed independently. Flowering time measurements were collected on the Colorado and eastern Kansas field sites.

Phenotypic Analysis

Height and flowering time distributions were used to evaluate phenotypic ranges within and across the mapping families. For plant height, the two independent measurements for each plot were averaged in each location and used in analysis. Height and flowering time data were normalized using standard scaling in each location. Each entry was characterized into a 'Family' group: RTx430 × SC979 RILs and parents (SC979 and RTx430), RTx430 × BTx2752 RILs and parents (BTx2752 and RTx430), and checks.

The family groups for height and flowering time in each location were independently checked for normality using the Shapiro-Wilk test. The nonparametric Kruskal-Wallis H and pairwise Wilcoxon rank sum tests were used to evaluate if the groups were significantly different. Phenotype values were plotted in violin plots to visualize the distribution of phenotypes. Phenotype values were plotted alongside guidelines of industry standards for hybrids in the target environment to assess agronomic relevance of the mapping germplasm.

Estimation of Broad-Sense Heritability

Broad-sense heritability (H^2) for height and flowering time were estimated for both families using variance components (σ^2). The H^2 formula divides total genetic variance (σ^2_G) by phenotypic variance (σ^2_P), producing a proportion from 0 to 1. The lmer() function from the lme4 package (Bates et al., 2015) in R (R Core Development Team, 2020) was used to estimate variance components. Height variance component calculation included multi-location replicated data. Terms in the lmer() model included genotype and location. Flowering time variance

components were calculated using multi-location replicated data. Terms in the lmer() model included genotype and location. All effects were treated as random.

$$H^2 = \sigma^2_G/\sigma^2_P$$

The H^2 for height and flowering time in each family was estimated using the calculated variance components and the Cullis method due to the unbalanced number of genotypes included in each location (Schmidt et al., 2019). The Cullis method leverages the mean variance of pairwise differences of genotypic best linear unbiased predictors (BLUPs) to account for unbalanced data using the formula:

$$H^2_{\text{Cullis}} = 1 - \frac{\bar{v}^{\text{BLUP}}_{\Delta}}{2\sigma^2_g}$$

Where σ^2_g is the genotypic variance and $\bar{v}^{BLUP}_{\Delta}$ is the mean variance of the difference of two genotypic best linear unbiased predictions (BLUPs).

Genotyping and GWAS Analysis

Tissue samples for DNA extractions were collected from seed increase plots in 2021-2022 winter nursery. Extraction samples were sent to Diversity Arrays Technology (DArT; Canberra, Australia) for genotyping with DArTseq-LD (low density) SNP calling. Marker data were filtered to remove monomorphic SNPs, duplicate SNPs, non-biallelic SNPs, SNPs with a call rate < 0.5, individuals with a call rate < 0.5, and SNPs with a RepAvg < 0.9 (index of reproducibility) to produce a set of 2,738 SNPs for 366 lines. Imputation was completed using BEAGLE (Browning et al., 2021).

BLUPs for the average height and flowering times for each line across all locations were calculated for use as phenotype inputs in a genome-wide association study (GWAS). Location and genotype were included in the model, and all effects were treated as random. The genotype

data was filtered to only include lines with phenotypes. The Genome Association and Prediction Integrated Tool (GAPIT) package (Lipka et al., 2012) for R was used to conduct GWAS. A general linear model (GLM) model was used (Price et al., 2006), with no principal components due to limited population structure resulting from high relatedness of the lines. The GLM model treats all individuals as a single group, increasing detection power without population structure in a population of RILs. A minor allele frequency (MAF) filter of 5% was set to help reduce false positives. A nominal threshold for significant SNPs was determined using the formula "-log10(0.01/effective number of SNPs)" according to the Bonferroni method (Duggal et al., 2008).

RESULTS

H^2 Elucidates Fixation of Genes Underlying Height and Maturity

To test the hypothesis that the mapping families control for phenotypic covariates caused by heterogeneity in height and flowering time when grown in a mapping population, the genetic architecture of height and maturity variation in the families was characterized (Table 2.1). Under the hypothesis that the parental germplasm used in family development is isogenic for major-effect height and maturity QTL, it is predicted that the broad-sense heritability (H^2) of these traits would be close to zero. Conversely, the hypothesis that the germplasm contains genetic variation for those major effect height and maturity loci would be supported by moderate to high H^2.

Using one year of height data with three locations, the H^2 of height was calculated to be 0.86 for the RTx430 × BTx2752 family and 0.93 for the RTx430 × SC979 family. The H^2 in both families is relatively high, indicating genetic variation explains most of the variance in height phenotypes (Table 2.2).

Using one year of flowering time data in two locations, the H^2 of maturity was calculated to be 0.93 for the RTx430 × BTx2752 family and 0.89 for the RTx430 × SC979 family. The relatively high H^2 in both families represents a substantial genetic contribution to the phenotypic variance for flowering time (Table 2.1).

Table 2.1. Broad-sense heritability (H^2) estimated using the Cullis method of height and flowering time in each family across all locations in 2021. Genotype and location used as model terms when extracting variance components, all effects treated as random.

Family	H^2	
	Flowering Time	Height
RTx430 × SC979	0.89	0.93
RTx430 × BTx2752	0.93	0.86

H^2 = Broad-sense heritability

Genotypic Analysis of Height and Flowering Time

To further test the hypotheses on the genetic architecture of height and flowering time variation in the LT mapping population, a genome-wide association study (GWAS) was performed for each trait. DArTseq-LD SNP data was called by DArT for low density coverage of the genome. For sorghum chromosomes 1-10, 334, 321, 343, 306, 302, 204, 198, 247, 195, and 289 SNP markers were called respectively (Figure 2.4). Genotype (allele) frequencies (Figure 2.5) are 83.9% homozygous reference, 12.1% homozygous alternative, and 4.0% heterozygous.

Fifty-two putatively significant marker-trait associations (MTA) for height were identified. Putatively significant SNPs were detected near known height loci on chromosomes 6, 7, and 9 (Figure 2.6). For flowering time, 158 putatively significant MTAs were identified. Significant SNPs were detected near several known maturity loci on chromosomes 1, 6, and 10 (Figure 2.7).

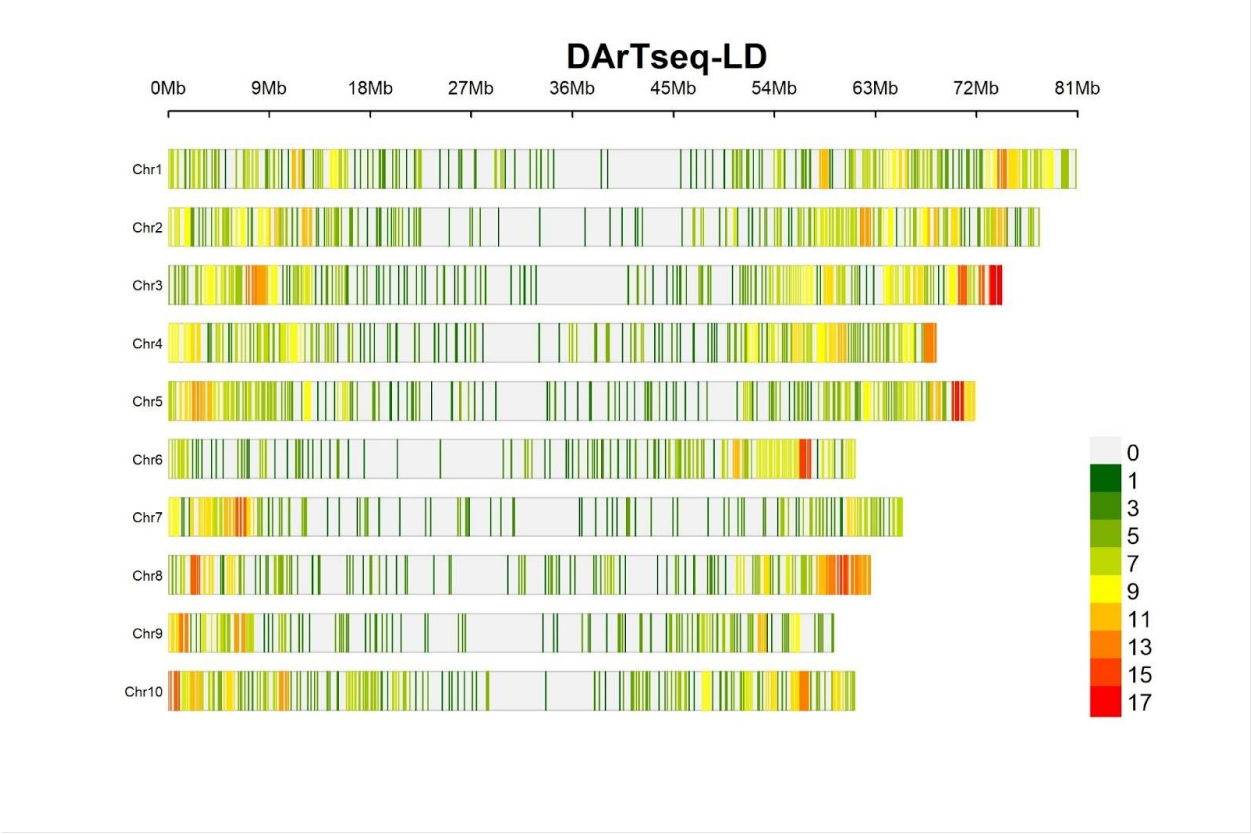


Figure 2.4. Per-chromosome single-nucleotide polymorphism (SNP) density of the LT mapping families using SNPs called with DArTseq-LD from 2021-2022 genotyping.

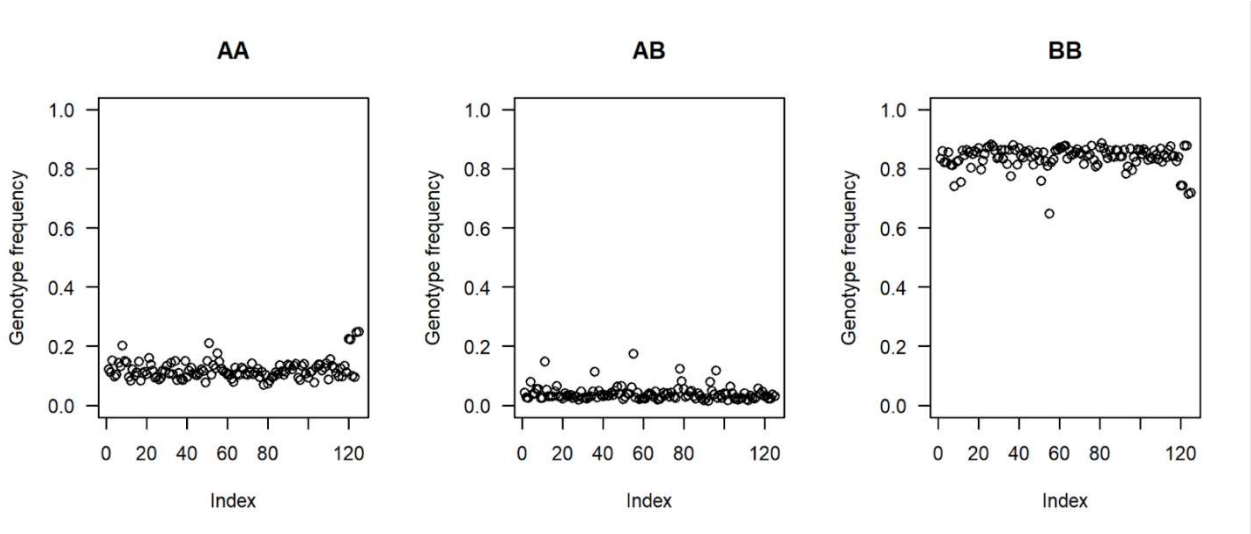


Figure 2.5. Genotype frequencies of imputed DArTseq-LT genotyping data from 2021-2022 genotyping.

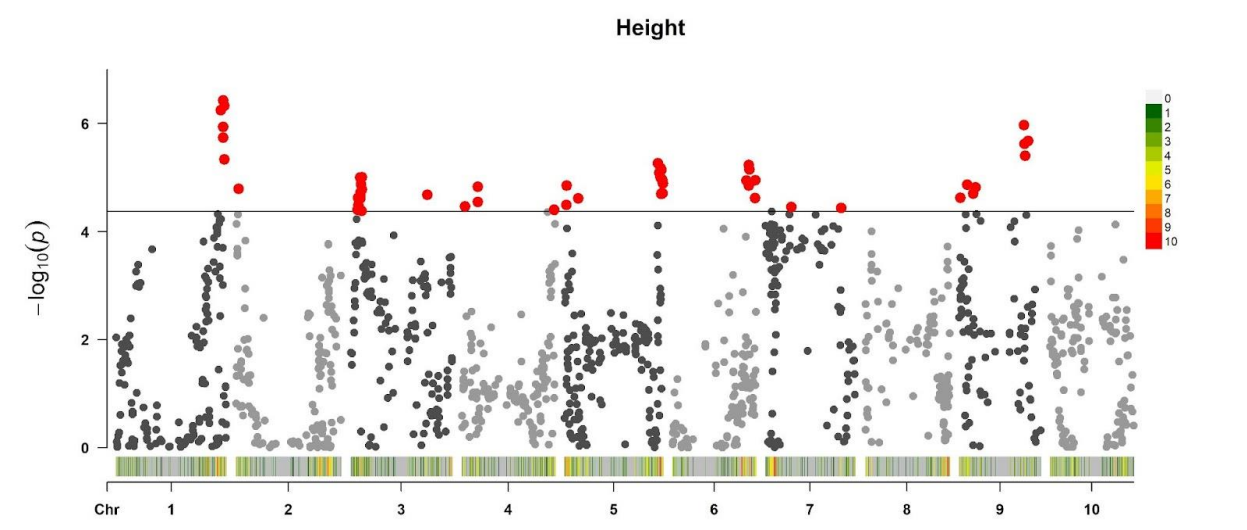


Figure 2.6. Manhattan plot for GWAS results showing associations of genetic markers and height BLUPs averaged across all locations. Red points indicate significant marker-trait associations. The black horizontal line indicates significance threshold for markers. The red to green scale indicates marker density.

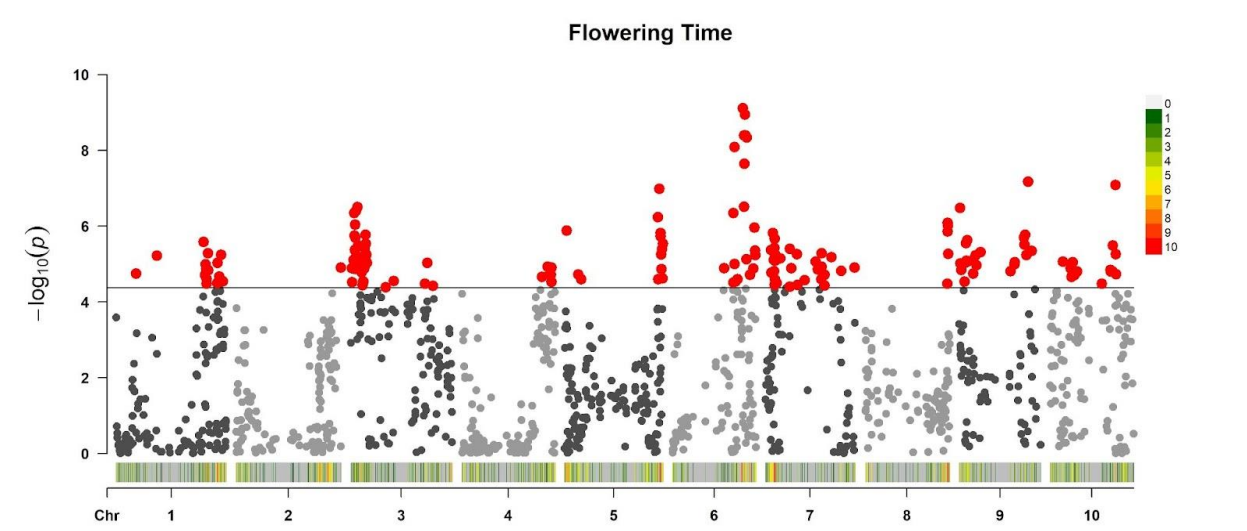


Figure 2.7. Manhattan plot for GWAS results showing associations of genetic markers and flowering time BLUPs averaged across all locations. Red points indicate significant marker-trait associations. The black horizontal line indicates significance threshold for markers. The red to green scale indicates marker density.

Agronomic Suitability is Comparable Between RIL Families

To test the hypotheses that the mapping families control or do not control for phenotypic covariates caused by heterogeneity in height when grown in a mapping population, the phenotype ranges were evaluated between RIL families in each location. The height of most

RILs is concentrated within a 60 cm window within all locations and a small proportion of outliers are responsible for the larger absolute ranges (Figure 2.8). In Colorado, 90% of the RTx430 × SC979 family and 93% of the RTx430 × BTx2752 family heights were within the 90 to 150 cm range. In western Kansas, 92% of the RTx430 × SC979 family and 93% of the RTx430 × BTx2752 family heights were within the 110 to 170 cm range. In eastern Kansas, 94% of the RTx430 × SC979 family and 91% of the RTx430 × BTx2752 were within the 105 to 165 cm range. The difference of the total range was 120 cm in Colorado, 128 cm in western Kansas, and 111 cm in eastern Kansas (Table 2.2). The height ranges between the families in the location were significantly different in Colorado ($p < 10^{-15}$) and western Kansas ($p < 10^{-15}$), but not in eastern Kansas ($p = 0.98$).

To test the hypotheses that the mapping families control or do not control for phenotypic covariates caused by heterogeneity in flowering time when grown in a mapping population, the phenotype ranges were evaluated between RIL families in both locations. Within each location, flowering is initiated at nearly the same number of days after emergence in both families. Both families in Colorado began flowering around 71 days after emergence and both families in eastern Kansas began flowering 58 days after emergence (Table 2.2, Figure 2.9). In Colorado, RILs in the RTx430 × BTx2752 family flowered over a range of 34 days and 31 days in the RTx430 × SC979 family. In eastern Kansas, the plots in the RTx430 × BTx2752 family flowered over the course of 34 days and 25 days in the RTx430 × SC979 family. The flowering time ranges of the families were significantly different in both Colorado ($p < 10^{-8}$) and eastern Kansas ($p < 10^{-15}$).

Table 2.2. Summary statistics (minimum, maximum, mean, median) for height and flowering time in mapping families by location in 2021.

Family	Height (cm)				Days to Flowering			
	Min	Max	Mean	Median	Min	Max	Mean	Median
Colorado								
RTx430 × BTx2752	70	148	109	107	76	110	90	89
RTx430 × SC979	78	190	121	119	71	102	83	82
Western Kansas								
RTx430 × BTx2752	82	185	130	131	-	-	-	-
RTx430 × SC979	103	210	142	141	-	-	-	-
Eastern Kansas								
RTx430 × BTx2752	97	232	136	134	58	92	71	70
RTx430 × SC979	88	181	135	135	58	83	68	70

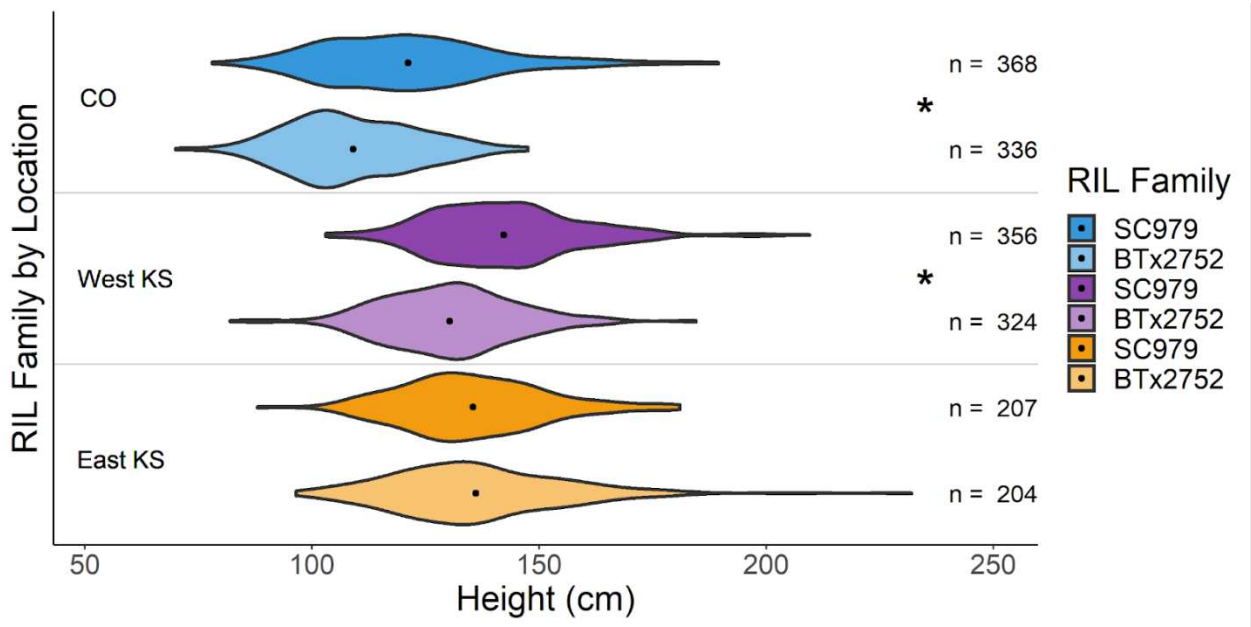


Figure 2.8. Height distributions by RIL family in each location. Asterisks indicate the families in that location are significantly different. Gray horizontal lines separate locations for visual clarity. Black dots indicate mean.

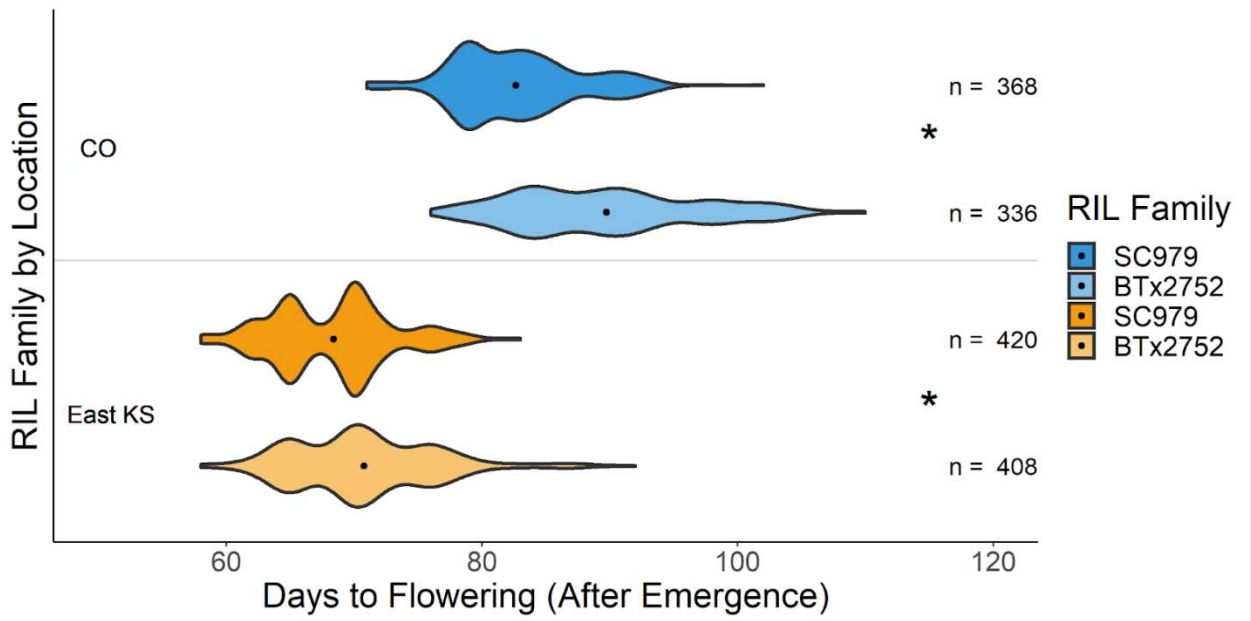


Figure 2.9. Flowering time distributions by RIL family in each 2021 location. Asterisks indicate the families in that location are significantly different. Gray horizontal lines separate locations for visual clarity. Black dots indicate mean.

DISCUSSION

Water-use traits, like limited transpiration, possess unique considerations for genetic mapping that sorghum mapping families developed previously do not address. Homogeneity within a mapping population for agronomic traits like height and flowering time is essential to increase power of the water-use trait phenotyping.

Haplotypes from Converted Germplasm Are Not Identical-by-State

Non-zero H^2 in our families supports the hypothesis that the LT mapping families contain genetic variation at large effect height and maturity QTL (Table 2.1) (Smith et al., 1998). Retrospective analyses of converted germplasm, which is present in the pedigrees of all parent lines in the LT mapping families, found that height and maturity haplotypes introgressed from elite donor BTx406 during the SCP were not identical-by-descent (Klein et al., 2008; Thurber et al., 2013). Therefore, genetic variation for large-effect height and maturity loci in the LT families may be residual from the conversions, potentially produced by haplotypes not identical-by-state

among the parental germplasm. High H^2 reflects that genetic variation is contributing a very large proportion of variance to the phenotypic distributions, attributed to a small number of moderate-effect loci consistent with the known architecture of height and maturity (Table 2.1, Figure 2.6, Figure 2.7) (Higgins et al., 2014). Potential heterogeneity of additive loci (Forsberg et al., 2015) or pervasive epistasis occurring in the background (Shao et al., 2008) may also contribute to putatively significant marker-trait associations. The presence of genetic variation for height and maturity requires evaluation of phenotype homogeneity to control for covariates.

Genetics and Genomics Reinforce Known Marker-Trait Associations

The unexpected segregation patterns of allele frequencies (Figure 2.5) in the outsourced genotyping data prompts concern for the reliability of the SNP calls (Lu et al., 2010). Patterns within markers across individuals indicate more effective recombinations than biologically expected for the generations of RILs (Bouchet et al., 2017). Communication has been ongoing to identify and resolve the concerns. Marker data was used as-is but all analyses will be repeated with corrected data, or genotype data from another service provider.

The GWAS for height and flowering time (Figure 2.6, Figure 2.7) provide insight into the genomic basis of variation explained by H^2. Colocalizations of significant MTA and known height and maturity loci (*Dw1* at ~57 Mb on chr 9, *Dw3* at ~58.557 Mb on chr 7, and *Ma1* at ~40.27 Mb on chr 6) (Brown et al., 2008; Casto et al., 2019; Li et al., 2015; Lopez et al., 2017; Quinby, 1966; Sukumaran et al., 2016) indicate there is variation among the mapping families for major-effect height and maturity genes. This also generates hypotheses on potentially novel smaller effect genes, or epistatic interactions contributing to phenotypic variance (Figure 2.8, Figure 2.9). Knowledge of height and flowering genetic variation in the RILs can be useful in

donor line development to understand how additive variation for those traits will combine with an elite background.

Height and Flowering Ranges Limit Phenotypic Covariates

A population perfectly homogeneous for height and flowering time would exist when, regardless of the underlying genotypes, the lines growing in the population all grow to the same height and all flower synchronously (Kammholz et al., 2001). The ranges of height and flowering time phenotypes collected on the LT mapping families show that there is not perfect homogeneity (Table 2.2). However, previous studies of UAS phenotyping suggest that the ranges will not substantially confound UAS-based thermal imaging. For instance, the height range of ~60 cm across most lines in the LT studies (Figure 2.8) leads us to conclude that it is insufficient height variation to shade neighbors in the thermal data extraction rows (Wang et al., 2018), and that the randomized complete block replication scheme prevents bias. While height alleles are not fixed within the families (Table 2.2), the combination of alleles at the four major-effect sorghum height loci and background interactions are producing comparable phenotypes between the RILs.

The flowering time range of ~30 days across the locations (Figure 2.9) has biological implications on the stage of canopy architecture between the early and late lines. We must consider the hypothesis that the differences in flowering time affect the timing of canopy closure and vegetative-to-reproductive transition, and therefore canopy temperature. Studies on growth stage dynamics and canopy traits (Liedtke et al., 2020; Umesh et al., 2022; Varela et al., 2021) have found that the horizontal and vertical canopy cover increases linearly pre-anthesis, with rate of gain highest in growth stages 30-50 days after planting. In both of our locations with flowering data, the largest concentration of plots flowered within a 12-day range after the

significant canopy growth period (Table 2.2). Biologically, this indicates that the LT mapping population plots are at a comparable canopy cover and architecture during the pre-anthesis thermal imaging window. The earlier flowering in the Kansas experiment (Figure 2.9) only restricts the thermal imaging data collection window. The LT canopy temperature data collection window is limited to the period of the growing season with high daily VPD (PRISM Climate Group, Oregon State University, https://prism.oregonstate.edu) after canopy closure and prior to anthesis. Ultimately, the mapping families are not precisely homogeneous for height and flowering time but appear to be acceptable to limit associated phenotypic covariates in LT phenotyping (Cockram & Mackay, 2018).

Replication Increases Power of Genetic Mapping

The power of this study is strengthened by multiple location data but limited by single year data. The acceptable (but not ideal) height and flowering background of the RIL families supports the need for replication (Singh et al., 2021; Thorp et al., 2018). Data from the 2022 field season are currently in collection, consolidation, and processing and will be included in analyses of this study when available. Outsourced genotype data obtained from DArT presents patterns of allele frequency segregation abnormal for RILs. When corrected SNP data is obtained, all analyses will be reassessed and completed again. With sufficient replication, the mapping families potentially provide utility for other water-use trait mapping studies, such as root traits, or differential irrigation studies to test physiology of water-use responses.

CONCLUSION

To contribute to the need for more water-efficient dryland cropping systems, two biparental RIL families were developed for the purpose of genetically characterizing a water-use trait,

limited transpiration, in sorghum. The use of canopy temperature captured using UAS thermal imaging as a high-throughput proxy necessitated selecting parents with similar agronomic characteristics like height and flowering time to control for associated phenotypic covariates. In this study, we attempt to validate the suitability of the LT mapping families in controlling for those covariates. The presence of germplasm developed through Sorghum Conversion Program in pedigrees of the parent lines would suggest an isogenic background for height and flowering, reducing heterogeneity of the field. When testing the hypothesis that the families are isogenic for these traits, we calculated high H^2 for both traits in both families, indicating that genetic variance is contributing to observed phenotypic distributions and therefore not isogenic. Variation at known large-effect height and maturity loci was observed from GWAS. However, the phenotypic ranges indicate that the populations in all locations are relatively homogeneous for height and maturity, and therefore acceptable in controlling for phenotypic covariates. These results suggest that, with adequate replication, the mapping families are suitable and powerful for genetic mapping of the limited transpiration trait and other water-use trait mapping studies.

REFERENCES

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bouchet, S., Olatoye, M. O., Marla, S. R., Perumal, R., Tesso, T., Yu, J., Tuinstra, M., & Morris, G. P. (2017). Increased Power To Dissect Adaptive Traits in Global Sorghum Diversity Using a Nested Association Mapping Population. *Genetics*, *206*(2), 573–585. https://doi.org/10.1534/genetics.116.198499

Boyles, R. E., Pfeiffer, B. K., Cooper, E. A., Zielinski, K. J., Myers, M. T., Rooney, W. L., & Kresovich, S. (2017). Quantitative Trait Loci Mapping of Agronomic and Yield Traits in Two Grain Sorghum Biparental Families. *Crop Science*, *57*(5), 2443–2456. https://doi.org/10.2135/cropsci2016.12.0988

Brown, P. J., Rooney, W. L., Franks, C., & Kresovich, S. (2008). Efficient Mapping of Plant Height Quantitative Trait Loci in a Sorghum Association Population With Introgressed Dwarfing Genes. *Genetics*, *180*(1), 629–637. https://doi.org/10.1534/genetics.108.092239

Browning, B. L., Tian, X., Zhou, Y., & Browning, S. R. (2021). Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics*, *108*(10), 1880–1890. https://doi.org/10.1016/j.ajhg.2021.08.005

Casto, A. L., Mattison, A. J., Olson, S. N., Thakran, M., Rooney, W. L., & Mullet, J. E. (2019). Maturity2, a novel regulator of flowering time in Sorghum bicolor, increases expression of SbPRR37 and SbCO in long days delaying flowering. *PLOS ONE*, *14*(4), e0212154. https://doi.org/10.1371/journal.pone.0212154

Chang, T.-G., Song, Q.-F., Zhao, H.-L., Chang, S., Xin, C., Qu, M., & Zhu, X.-G. (2020). An in situ approach to characterizing photosynthetic gas exchange of rice panicle. *Plant Methods*, *16*(1), 92. https://doi.org/10.1186/s13007-020-00633-1

Choudhary, S., Sinclair, T. R., Choudhary, S., & Sinclair, T. R. (2013). Hydraulic conductance differences among sorghum genotypes to explain variation in restricted transpiration rates. *Functional Plant Biology*, *41*(3), 270–275. https://doi.org/10.1071/FP13246

Cockram, J., & Mackay, I. (2018). Genetic Mapping Populations for Conducting High-Resolution Trait Mapping in Plants. In R. K. Varshney, M. K. Pandey, & A. Chitikineni (Eds.), *Plant Genetics and Molecular Biology* (pp. 109–138). Springer International Publishing. https://doi.org/10.1007/10_2017_48

Dai, A., Zhao, T., & Chen, J. (2018). Climate Change and Drought: A Precipitation and Evaporation Perspective. *Current Climate Change Reports*, *4*(3), 301–312. https://doi.org/10.1007/s40641-018-0101-6

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., James, R. A., Condon, A. G., Bovill, W. D., Hutchinson, P., Scarrow, J., Davy, R., & Furbank, R. T. (2016). Methodology for High-Throughput Field Phenotyping of Canopy Temperature Using Airborne Thermography. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01808

Devi, M. J., Sinclair, T. R., & Vadez, V. (2010). Genotypic Variation in Peanut for Transpiration Response to Vapor Pressure Deficit. *Crop Science*, *50*(1), 191–196. https://doi.org/10.2135/cropsci2009.04.0220

Dong, Z., Jiang, C., Chen, X., Zhang, T., Ding, L., Song, W., Luo, H., Lai, J., Chen, H., Liu, R., Zhang, X., & Jin, W. (2013). Maize LAZY1 Mediates Shoot Gravitropism and Inflorescence Development through Regulating Auxin Transport, Auxin Signaling, and

Light Response. *Plant Physiology*, *163*(3), 1306–1322. https://doi.org/10.1104/pp.113.227314

Duggal, P., Gillanders, E. M., Holmes, T. N., & Bailey-Wilson, J. E. (2008). Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics*, *9*(1), 516. https://doi.org/10.1186/1471-2164-9-516

Faye, J. M., Akata, E. A., Sine, B., Diatta, C., Cisse, N., Fonceka, D., & Morris, G. P. (2022). Quantitative and population genomics suggest a broad role of stay-green loci in the drought adaptation of sorghum. *The Plant Genome*, *15*(1), e20176. https://doi.org/10.1002/tpg2.20176

Forsberg, S. K. G., Andreatta, M. E., Huang, X.-Y., Danku, J., Salt, D. E., & Carlborg, Ö. (2015). The Multi-allelic Genetic Architecture of a Variance-Heterogeneity Locus for Molybdenum Concentration in Leaves Acts as a Source of Unexplained Additive Genetic Variance. *PLOS Genetics*, *11*(11), e1005648. https://doi.org/10.1371/journal.pgen.1005648

Fullana-Pericàs, M., Conesa, M. À., Gago, J., Ribas-Carbó, M., & Galmés, J. (2022). High-throughput phenotyping of a large tomato collection under water deficit: Combining UAVs' remote sensing with conventional leaf-level physiologic and agronomic measurements. *Agricultural Water Management*, *260*, 107283. https://doi.org/10.1016/j.agwat.2021.107283

Gholipoor, M., Choudhary, S., Sinclair, T., Messina, C., & Cooper, M. (2013). Transpiration Response of Maize Hybrids to Atmospheric Vapour Pressure Deficit. *Journal of Agronomy and Crop Science*, *199*. https://doi.org/10.1111/jac.12010

Gholipoor, M., Prasad, P. V. V., Mutava, R. N., & Sinclair, T. R. (2010). Genetic variability of

transpiration response to vapor pressure deficit among sorghum genotypes. *Field Crops Research*, *119*(1), 85–90. https://doi.org/10.1016/j.fcr.2010.06.018

Girma, F. S., & Krieg, D. R. (1992). Osmotic Adjustment in Sorghum: II. Relationship to Gas Exchange Rates. *Plant Physiology*, *99*(2), 583–588.

Gonzalez-Dugo, V., Lopez-Lopez, M., Espadafor, M., Orgaz, F., Testi, L., Zarco-Tejada, P., Lorite, I. J., & Fereres, E. (2019). Transpiration from canopy temperature: Implications for the assessment of crop yield in almond orchards. *European Journal of Agronomy*, *105*, 78–85. https://doi.org/10.1016/j.eja.2019.01.010

Higgins, R. H., Thurber, C. S., Assaranurak, I., & Brown, P. J. (2014). Multiparental Mapping of Plant Height and Flowering Time QTL in Partially Isogenic Sorghum Families. *G3 Genes|Genomes|Genetics*, *4*(9), 1593–1602. https://doi.org/10.1534/g3.114.013318

Hou, M., Tian, F., Ortega-Farias, S., Riveros-Burgos, C., Zhang, T., & Lin, A. (2021). Estimation of crop transpiration and its scale effect based on ground and UAV thermal infrared remote sensing images. *European Journal of Agronomy*, *131*, 126389. https://doi.org/10.1016/j.eja.2021.126389

Inoue, Y., Kimball, B. A., Jackson, R. D., Pinter, P. J., & Reginato, R. J. (1990). Remote estimation of leaf transpiration rate and stomatal resistance based on infrared thermometry. *Agricultural and Forest Meteorology*, *51*(1), 21–33. https://doi.org/10.1016/0168-1923(90)90039-9

Kammholz, S. J., Campbell, A. W., Sutherland, M. W., Hollamby, G. J., Martin, P. J., Eastwood, R. F., Barclay, I., Wilson, R. E., Brennan, P. S., & Sheppard, J. A. (2001). Establishment and characterisation of wheat genetic mapping populations. *Australian Journal of Agricultural Research*, *52*(12), 1079–1088. https://doi.org/10.1071/ar01043

Klein, R. R., Mullet, J. E., Jordan, D. R., Miller, F. R., Rooney, W. L., Menz, M. A., Franks, C. D., & Klein, P. E. (2008). The Effect of Tropical Sorghum Conversion and Inbred Development on Genome Diversity as Revealed by High-Resolution Genotyping. *Crop Science*, *48*(S1), S-12-S-26. https://doi.org/10.2135/cropsci2007.06.0319tpg

Li, X., Li, X., Fridman, E., Tesso, T. T., & Yu, J. (2015). Dissecting repulsion linkage in the dwarfing gene Dw3 region for sorghum plant height provides insights into heterosis. *Proceedings of the National Academy of Sciences*, *112*(38), 11823–11828. https://doi.org/10.1073/pnas.1509229112

Liedtke, J. D., Hunt, C. H., George-Jaeggli, B., Laws, K., Watson, J., Potgieter, A. B., Cruickshank, A., & Jordan, D. R. (2020). High-Throughput Phenotyping of Dynamic Canopy Traits Associated with Stay-Green in Grain Sorghum. *Plant Phenomics*, *2020*, 4635153. https://doi.org/10.34133/2020/4635153

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Lopez, J. R., Erickson, J. E., Munoz, P., Saballos, A., Felderhoff, T. J., & Vermerris, W. (2017). QTLs Associated with Crown Root Angle, Stomatal Conductance, and Maturity in Sorghum. *The Plant Genome*, *10*(2), plantgenome2016.04.0038. https://doi.org/10.3835/plantgenome2016.04.0038

Lu, Y., Zhang, S., Shah, T., Xie, C., Hao, Z., Li, X., Farkhari, M., Ribaut, J.-M., Cao, M., Rong, T., & Xu, Y. (2010). Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proceedings of the National Academy of Sciences of the United States of America*,

107(45), 19585–19590. https://doi.org/10.1073/pnas.1006105107

Maxted, N., Dulloo, M. E., & Ford-Lloyd, B. V. (2016). *Enhancing Crop Genepool Use: Capturing Wild Relative and Landrace Diversity for Crop Improvement*. CABI.

Murphy, R. L., Morishige, D. T., Brady, J. A., Rooney, W. L., Yang, S., Klein, P. E., & Mullet, J. E. (2014). Ghd7 (Ma6) Represses Sorghum Flowering in Long Days: Ghd7 Alleles Enhance Biomass Accumulation and Grain Production. *The Plant Genome*, *7*(2), plantgenome2013.11.0040. https://doi.org/10.3835/plantgenome2013.11.0040

Otto, S. P., & Jones, C. D. (2000). Detecting the Undetected: Estimating the Total Number of Loci Underlying a Quantitative Trait. *Genetics*, *156*(4), 2093–2107. https://doi.org/10.1093/genetics/156.4.2093

Perumal, R., Tesso, T. T., Morris, G. P., Jagadish, S. V. K., Little, C. R., Bean, S. R., Yu, J., Prasad, P. V. V., & Tuinstra, M. R. (2021). Registration of the sorghum nested association mapping (NAM) population in RTx430 background. *Journal of Plant Registrations*, *15*(2), 395–402. https://doi.org/10.1002/plr2.20110

Pignon, C. P., Fernandes, S. B., Valluru, R., Bandillo, N., Lozano, R., Buckler, E., Gore, M. A., Long, S. P., Brown, P. J., & Leakey, A. D. B. (2021). Phenotyping stomatal closure by thermal imaging for GWAS and TWAS of water use efficiency-related genes. *Plant Physiology*, *187*(4), 2544–2562. https://doi.org/10.1093/plphys/kiab395

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), Article 8. https://doi.org/10.1038/ng1847

Quinby, J. R. (1966). Fourth Maturity Gene Locus in Sorghum1. *Crop Science*, *6*(6), cropsci1966.0011183X000600060005x.

https://doi.org/10.2135/cropsci1966.0011183X000600060005x

Quinby, J. R. (1967). The Maturity Genes of Sorghum. In A. G. Norman (Ed.), *Advances in Agronomy* (Vol. 19, pp. 267–305). Academic Press. https://doi.org/10.1016/S0065-2113(08)60737-3

Quinby, J. R. (1974). The Genetic Control of Flowering and Growth in Sorghum. In N. C. Brady (Ed.), *Advances in Agronomy* (Vol. 25, pp. 125–162). Academic Press. https://doi.org/10.1016/S0065-2113(08)60780-4

Raymundo, R., Sexton-Bowser, S., Ciampitti, I. A., & Morris, G. P. (2021). Crop modeling defines opportunities and challenges for drought escape, water capture, and yield increase using chilling-tolerant sorghum. *Plant Direct*, *5*(9), e349. https://doi.org/10.1002/pld3.349

Rebetzke, G., Herwaarden, A., Jenkins, C., Weiss, M., Lewis, D., Ruuska, S., Tabe, L., Fettell, N., & Richards, R. (2008). Quantitative trait loci for water-soluble carbohydrates and associations with agronomic traits in wheat. *Australian Journal of Agricultural Research - AUST J AGR RES*, *59*. https://doi.org/10.1071/AR08067

Rebetzke, G. J., Condon, A. G., Farquhar, G. D., Appels, R., & Richards, R. A. (2008). Quantitative trait loci for carbon isotope discrimination are repeatable across environments and wheat mapping populations. *Theoretical and Applied Genetics*, *118*(1), 123–137. https://doi.org/10.1007/s00122-008-0882-4

Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., Condon, A. (Tony) G., Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., & Condon, A. (Tony) G. (2012). Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. *Functional Plant Biology*, *40*(1), 14–33.

https://doi.org/10.1071/FP12184

Riar, M. K., Sinclair, T. R., & Prasad, P. V. V. (2015). Persistence of limited-transpiration-rate trait in sorghum at high temperature. *Environmental and Experimental Botany*, *115*, 58–62. https://doi.org/10.1016/j.envexpbot.2015.02.007

Sadok, W., Lopez, J. R., & Smith, K. P. (2021). Transpiration increases under high-temperature stress: Potential mechanisms, trade-offs and prospects for crop resilience in a warming world. *Plant, Cell & Environment*, *44*(7), 2102–2116. https://doi.org/10.1111/pce.13970

Sahoo, L., Schmidt, J. J., Pedersen, J. F., Lee, D. J., & Lindquist, J. L. (2010). Growth and fitness components of wild × cultivated Sorghum bicolor (Poaceae) hybrids in Nebraska. *American Journal of Botany*, *97*(10), 1610–1617.

Schmidt, P., Hartung, J., Rath, J., & Piepho, H.-P. (2019). Estimating Broad-Sense Heritability with Unbalanced Data from Agricultural Cultivar Trials. *Crop Science*, *59*(2), 525–536. https://doi.org/10.2135/cropsci2018.06.0376

Shao, H., Burrage, L. C., Sinasac, D. S., Hill, A. E., Ernest, S. R., O'Brien, W., Courtland, H.-W., Jepsen, K. J., Kirby, A., Kulbokas, E. J., Daly, M. J., Broman, K. W., Lander, E. S., & Nadeau, J. H. (2008). Genetic architecture of complex traits: Large phenotypic effects and pervasive epistasis. *Proceedings of the National Academy of Sciences*, *105*(50), 19910–19914. https://doi.org/10.1073/pnas.0810388105

Sharma, S., Upadhyaya, H. D., Varshney, R. K., & Gowda, C. L. L. (2013). Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Frontiers in Plant Science*, *4*. https://doi.org/10.3389/fpls.2013.00309

Shekoofa, A., Balota, M., & Sinclair, T. R. (2014). Limited-transpiration trait evaluated in growth chamber and field for sorghum genotypes. *Environmental and Experimental*

Botany, *99*, 175–179. https://doi.org/10.1016/j.envexpbot.2013.11.018

Sinclair, T. R., Hammer, G. L., Oosterom, E. J. van, Sinclair, T. R., Hammer, G. L., & Oosterom, E. J. van. (2005). Potential yield and water-use efficiency benefits in sorghum from limited maximum transpiration rate. *Functional Plant Biology*, *32*(10), 945–952. https://doi.org/10.1071/FP05047

Singh, A., Jones, S., Ganapathysubramanian, B., Sarkar, S., Mueller, D., Sandhu, K., & Nagasubramanian, K. (2021). Challenges and Opportunities in Machine-Augmented Plant Stress Phenotyping. *Trends in Plant Science*, *26*(1), 53–69. https://doi.org/10.1016/j.tplants.2020.07.010

Smith, S. E., Kuehl, R. O., Ray, I. M., Hui, R., & Soleri, D. (1998). Evaluation of Simple Methods for Estimating Broad-Sense Heritability in Stands of Randomly Planted Genotypes. *Crop Science*, *38*(5), cropsci1998.0011183X003800050003x. https://doi.org/10.2135/cropsci1998.0011183X003800050003x

Stamenković, O. S., Siliveru, K., Veljković, V. B., Banković-Ilić, I. B., Tasić, M. B., Ciampitti, I. A., Đalović, I. G., Mitrović, P. M., Sikora, V. Š., & Prasad, P. V. V. (2020). Production of biofuels from sorghum. *Renewable and Sustainable Energy Reviews*, *124*, 109769. https://doi.org/10.1016/j.rser.2020.109769

Stephens, J. C., Miller, F. R., & Rosenow, D. T. (1967). Conversion of Alien Sorghums to Early Combine Genotypes1. *Crop Science*, *7*(4), cropsci1967.0011183X000700040036x. https://doi.org/10.2135/cropsci1967.0011183X000700040036x

Sukumaran, S., Li, X., Li, X., Zhu, C., Bai, G., Perumal, R., Tuinstra, M. R., Prasad, P. V. V., Mitchell, S. E., Tesso, T. T., & Yu, J. (2016). QTL Mapping for Grain Yield, Flowering Time, and Stay-Green Traits in Sorghum with Genotyping-by-Sequencing Markers. *Crop*

Science, *56*(4), 1429–1442. https://doi.org/10.2135/cropsci2015.02.0097

Takuno, S., Terauchi, R., & Innan, H. (2012). The power of QTL mapping with RILs. *PloS One*, *7*(10), e46545. https://doi.org/10.1371/journal.pone.0046545

Thorp, K. R., Thompson, A. L., Harders, S. J., French, A. N., & Ward, R. W. (2018). High-Throughput Phenotyping of Crop Water Use Efficiency via Multispectral Drone Imagery and a Daily Soil Water Balance Model. *Remote Sensing*, *10*(11), Article 11. https://doi.org/10.3390/rs10111682

Thurber, C. S., Ma, J. M., Higgins, R. H., & Brown, P. J. (2013). Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biology*, *14*(6), R68. https://doi.org/10.1186/gb-2013-14-6-r68

Turner, D., Lucieer, A., & Watson, C. (2012). An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds. *Remote Sensing*, *4*(5), Article 5. https://doi.org/10.3390/rs4051392

Umesh, M. R., Angadi, S., Begna, S., & Gowda, P. (2022). Planting Density and Geometry Effect on Canopy Development, Forage Yield and Nutritive Value of Sorghum and Annual Legumes Intercropping. *Sustainability*, *14*(8), Article 8. https://doi.org/10.3390/su14084517

Varela, S., Pederson, T., Bernacchi, C. J., & Leakey, A. D. B. (2021). Understanding Growth Dynamics and Yield Prediction of Sorghum Using High Temporal Resolution UAV Imagery Time Series and Machine Learning. *Remote Sensing*, *13*(9), Article 9. https://doi.org/10.3390/rs13091763

Wang, X., Singh, D., Marla, S., Morris, G., & Poland, J. (2018). Field-based high-throughput

phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods*, *14*(1), 53. https://doi.org/10.1186/s13007-018-0324-5

Yang, M., Hassan, M. A., Xu, K., Zheng, C., Rasheed, A., Zhang, Y., Jin, X., Xia, X., Xiao, Y., & He, Z. (2020). Assessment of Water and Nitrogen Use Efficiencies Through UAV-Based Multispectral Phenotyping in Winter Wheat. *Frontiers in Plant Science*, *11*. https://www.frontiersin.org/articles/10.3389/fpls.2020.00927

CHAPTER III: CHARACTERIZING GENETIC ARCHITECTURE AND MOLECULAR BASIS OF THE LIMITED TRANSPIRATION TRAIT IN SORGHUM

INTRODUCTION

Precipitation patterns are shifting because of climate change, spurring an increasing interest in breeding water-efficient crops for dryland agriculture (Dai et al., 2018). Trait discovery for the ecophysiological water-use 'limited transpiration (LT)' trait in sorghum (*Sorghum bicolor* (L.) Moench) offers potential for breeding more water-optimized varieties (Shekoofa et al., 2014). Plants with the LT genotype will reduce their transpiration rate in periods of annual and daily high vapor pressure deficit (VPD) (Gholipoor et al., 2010; Shekoofa et al., 2014) to conserve soil water for later use during grain fill. Hypotheses on the molecular basis of LT include influence of stomate density or stomatal regulation via aquaporins among other mechanisms (Heinen et al., 2009; Sinclair et al., 2017). Mapping populations with relatively homogeneous agronomic background can be phenotyped for LT using unoccupied aircraft system (UAS)-based canopy temperature as a high-throughput proxy (Belko et al., 2013; Deery et al., 2016). The lack of transpiration cooling effect in putative LT lines produces warmer canopy temperatures relative to putative non-LT lines (Gates, 1964). Understanding the architecture of genetic variation underlying the LT trait contributes to and development of selectable markers and an elite trait donor line as part of trait discovery.

Defined here as the number, location, and effect sizes of genomic regions conferring a specific phenotype (Holland, 2007), genetic architecture informs the strategy for donor line pre-breeding (Huynh et al., 2013). Genetic architectures can be divided broadly into two classes, qualitative and quantitative. Qualitative, or monogenic, variation is controlled by only one locus in the genome (or a very small number) with a large effect on the resulting phenotype (Rajon &

Plotkin, 2013). Quantitative traits are controlled complexly by a few genomic loci up to hundreds of loci, with effect sizes potentially ranging from large to miniscule. Knowledge of these genomic regions, called quantitative trait loci (QTL), is useful when making selections during breeding for ecophysiological traits that are difficult to phenotype (Zeng, 1994).

Quantitative architecture can be further subdivided into oligogenic or polygenic. Oligogenic traits are those controlled by a limited number of QTL of moderate effect size (Van Der Plank, 1966). The designation is subjective, and here we will use the range of 2-9 loci to indicate oligogenic architecture, reflecting the number of loci that can easily be tracked via outsourced marker genotyping (Thomson, 2014). Conversely, polygenic traits are those controlled by many loci of small effect size (Van Der Plank, 1966). Here, an architecture consisting of 10 or more loci would be considered polygenic. Highly polygenic architectures present difficulties for inclusion in targeted trait introgression and pre-breeding. Small effect sizes inhibit marker development due to the inability to statistically detect and map significant marker-trait associations (Kumar et al., 2019). Pre-breeders or trait introgression scientists must ensure that many necessary QTL are successfully transferred, and those introgressions are not disrupting other important haplotypes (Smith et al., 2010).

Linkage mapping and genome-wide association studies (GWAS) are the primary methods for evaluating trait genetic architecture. Linkage mapping most commonly uses biparental recombinant inbred line (RIL) families to leverage meiotic recombination, allowing pre-breeders to examine inheritance and identify haplotypes conserved across lines exhibiting the target phenotype (Liu et al., 2012). RIL mapping families provide sets of germplasm lines which inherit haplotypes from either parent in varying sizes and arrangements and are advanced to homozygosity at nearly all loci in the genome (Broman, 2005). These lines are closely related,

therefore constructing genetic linkage maps offers insight into regions of high recombination (Aguilar-Benitez et al., 2021; Bali et al., 2015). GWAS combines genotypes at various molecular markers in the genome with measured observations of a phenotype to find significant associations. GWAS leverages genetic diversity of lines in the sampling population (Gyawali et al., 2019). Additionally, estimating trait heritability provides insight into the influence of genetic variance on the phenotype. If trait mapping is not yielding significant associations due to many loci of small effect sizes, a non-zero heritability can validate that genetics are contributing to a phenotype and it is likely under polygenic control (Schmidt et al., 2019). After generating knowledge on the architecture of a trait, pre-breeders can then identify a breeding strategy for development of an elite donor line and trait transfer into varieties.

Many ecophysiological traits have quantitative variation due to environmental fluctuation and the integration of many physiological pathways within a plant (Casper et al., 2005; Edwards et al., 2011; Geber & Dawson, 1997). However, some physiological traits in crop plants are successfully conferred with CRISPR/Cas9 or targeted gene editing for a single gene, indicating qualitative control of LT may be possible (Zsögön et al., 2017). In quantitative traits, oligogenic architecture is more favorable than polygenic architecture for use in pre-breeding because fewer loci must be accounted for during targeted introgressions (Smith et al., 2010). Additionally, those loci will typically have effect sizes large enough to find significant QTL and develop markers for use in marker-assisted selection (Van Der Plank, 1966). Furthermore, detectable loci are useful in understanding the molecular basis of a trait. If an annotated reference genome is available, putative QTL detected in monogenic or oligogenic architectures allow pre-breeders to search associated regions for known genes. The goal of this study was to identify QTL associated with variation for the LT trait and understand the implications of their architecture on pre-breeding.

This knowledge will then be used to develop a donor line breeding strategy, including developing selectable markers if applicable.

MATERIALS AND METHODS

Field Design and Management

Field design and management is described in full detail in chapter 2. Briefly, two recombinant inbred line (RIL) mapping families RTx430 × SC979 and RTx430 × BTx2752 for the limited transpiration trait in sorghum were grown in three locations–northern Colorado, western Kansas, and eastern Kansas–in a single year. RTx430 is the putative non-LT parent while SC979 and BTx2752 are putative LT parents. The Colorado and western Kansas populations contained 161 RILs total from the two families while the eastern Kansas population contained 131 RILs. In each location, fields were planted according to a randomized complete-block design in early June following tilling, soil amending, and application of pre-emergent herbicide. The mapping populations were well-watered using supplemental irrigation to prevent drought stress.

UAS Data Collection

Weather stations were mounted at each mapping population location (eastern Kansas, western Kansas, northern Colorado). All flights in locations were completed between July 27th and September 1st, 2021, with the goal of at least 1 flight per week at each location. Flights were performed between the hours of 1200 and 1600 (the daily period of highest vapor pressure deficit based on 30-year PRISM data) (PRISM Climate Group, Oregon State University, https://prism.oregonstate.edu). Target temperature for flight days was >30°C and target relative humidity was <50%. Flights were only performed at timepoints with no cloud cover over the

field or casting shadows on the field. All field sites were equipped with ground control points (GCPs) at corners of the field and approximately evenly spaced within the field. No less than five GCPs were placed per location, typically with four at field corners and one or more distributed within the field. For western Kansas, the center-pivot quadrant necessitated placing GCPs apart in 100 m increments throughout the field, avoiding placing them in a straight line. GCPs allowed for high spatial accuracy and precise georeferencing within 2 cm (Martínez-Carricondo et al., 2018).

Equipment used in the Colorado data collection included the multi-rotor DJI Mavic Pro 2 with a 1" CMOS sensor for RGB, the FLIR Duo Pro R 13mm 640/512 sensor for thermal, and the MicaSense RedEdge-M sensor for multispectral imaging. Equipment used in the eastern and Western Kansas data collection included a multi-rotor DJI Mavic Pro 2, the FLIR Vue Pro R 13mm 640/512 sensor for thermal, and the MicaSense RedEdge3 sensor for multispectral imaging. Flights were completed at a speed of 3-4 m/s and altitude of 33-34 m. Forward overlap percentages were 75% and side overlap percentages were 75%. Flight missions in Colorado were completed by a different organization (Colorado State University Drone Center, Fort Collins, Colorado) than the flight missions in eastern and western Kansas (Kairos Geospatial, Abilene, Kansas).

Thermal calibration was completed using metal panels and a handheld infrared thermometer. At least 30 minutes before the flight mission, three rectangular metal panels painted white, gray, and black respectively were placed directly next to the field, unobstructed and unshaded, where they would be visible in the UAS imagery. A FLIR TG54 Spot IR Thermometer with default settings (emissivity = 0.95) was used to measure the metal panels directly before and after flights. Holding the handheld thermometer 2 ft from the panel, the

trigger was held down for 5 seconds to collect a single reading. Four readings were collected on each panel, one in each quadrant. The readings were recorded and used in thermal data calibration. MicaSense calibration was completed using the same metal panels as thermal calibration. The downwelling light sensor was used to measure irradiance before and after the flight mission in the same quadrant sampling method described for thermal calibration.

Imagery Data Processing

Agisoft Metashape Professional (version 1.7.4 build 13028 (64 bit)) was used to stitch the photos captured by the sensors into workable orthomosaics for data filtering and extraction. The orthomosaic processing described here was completed for each flight in each location. For both MicaSense RedEdge-M sensor and FLIR sensor data, photos were loaded as a multi-camera system and photos from ascending and descending were removed so that only calibration and grid-mission photos remained. Reflectance was calibrated using reflectance panels and sun sensor activated for the MicaSense data. Next, photos were aligned with settings active for high accuracy, generic preselection, reference preselection (source), key point limit set to 40,000, tie point limit set to 4,000, and no masks applied.

For the MicaSense sensor data, WGS 84 (EPSG::4326) coordinate system was used. Markers were detected using auto-detection and false detections were deleted. Georeferenced markers from the GPS file were imported and accuracy changed to 2 cm, where any markers missed in the auto-detection were manually marked. The bounding box was set for the field and the point cloud was trimmed, optimization was completed, and the error of control points was validated to be <10 cm before progressing. Next, the digital elevation model (DEM) was built with settings for a dense cloud and extrapolation. The orthomosaic was then built using the DEM and average values during blending to reduce temporal variability. This means that the values

from all photos that contribute to a particular pixel are used to calculate the average for that pixel in the resulting orthomosaic. Finally, the MicaSense orthomosaic was exported.

For the FLIR sensor data, any photos that did not initially align were reset and re-aligned. Using the WGS 84 (EPSG::4326) coordinate system, georeferenced markers were imported and accuracy changed to 5 cm for eastern and western Kansas maps and 7 cm for Colorado maps. While the GPS accuracy is 2 cm, the ground sample distance is calculated to be ~5 cm and ~7 cm respectively for Kansas and Colorado locations hence the adjustment. Next, the photos were filtered by tie points and markers were placed on aerial targets in a minimum of five images per flight. The bounding box was set for the field and the point cloud was trimmed, optimization was completed, and the error of control points was validated to be <10 cm before progressing. The DEM was built with settings for a sparse cloud and extrapolation, as dense cloud settings often produce large gaps when processing FLIR data. The orthomosaic was then built using the DEM and average values during blending to reduce temporal variability. Finally, the FLIR orthomosaic was exported.

Imagery-Based Trait Extraction

To reduce noise caused by edge effects of the plots, canopy temperature data was only extracted from the center two rows of each four-row plot in the mapping populations. The use of GCPs in the field sites provided high spatial accuracy (within 2 cm) to georeference the center two rows of each plot, ensuring the same sample area is extracted from each plot. Using ArcGIS Pro, a layer was created where an equal-sized rectangular polygon was fit to the center two rows with ~2 feet buffered from each side of every four-row plot using the MicaSense/RGB images as reference. Each plot polygon was named with a unique ID corresponding to the range and row

location of the plot in the field. The layer was then exported as a shapefile for use in data extraction.

ArcGIS Pro (version 2.9.3, Esri Inc., Redlands, United States) was used to extract canopy temperature data from mapping population orthomosaics. The MicaSense orthomosaic raster, thermal orthomosaic raster, and shapefile of plot polygons were loaded into the map interface. First, the MicaSense orthomosaic was used to generate a Normalized Difference Vegetation Index (NDVI) layer with a scientific output using the analysis raster functions. Next, in the Symbology tab, the NDVI layer was classified into five classes using the 'Natural Breaks (Jenks)' setting. After generating the histogram of values, the value of the lower boundary of the fifth bin of classified values was noted as the threshold between plant and soil pixels. Next, the 'Raster Calculator (Spatial Analyst Tools)' function was used to classify soil versus plant pixels. The NDVI layer was used as the raster calculator input and all pixels greater than or equal to the lower boundary of the fifth bin were sorted into one class and all pixels less than the threshold were sorted into a second class in the resulting new layer. The pixels above the threshold value (class 1) were recognized as plant material in the orthomosaic. Once the two classes (soil, plant) were classified, the soil pixels (class 0) were removed in the Symbology tab. This created a layer of only pixels assigned to plant material from the MicaSense raster.

The thermal orthomosaic raster was then clipped with the 'Clip' imagery raster functions tool using the plant material layer to remove all temperature pixels outside of the plant material. This generated a raster layer with temperature values for only plant pixels. Finally, this clipped temperature raster layer was used to pull zonal statistics. The 'Zonal Statistics as Table (Spatial Analyst Tools)' function was used with the plot polygon shapefile as the input feature zone data, the zone field as the plot name from the shapefile, and the plant pixel-clipped temperature raster

as the input value raster. All statistic types were generated, including count, area, minimum, maximum, range, mean, standard deviation, sum, median, and 90th percentile. Missing data was ignored in the statistics calculation. The final table containing all statistics for each plot was exported. The general thermal data extraction workflow is shown in Figure 3.1.

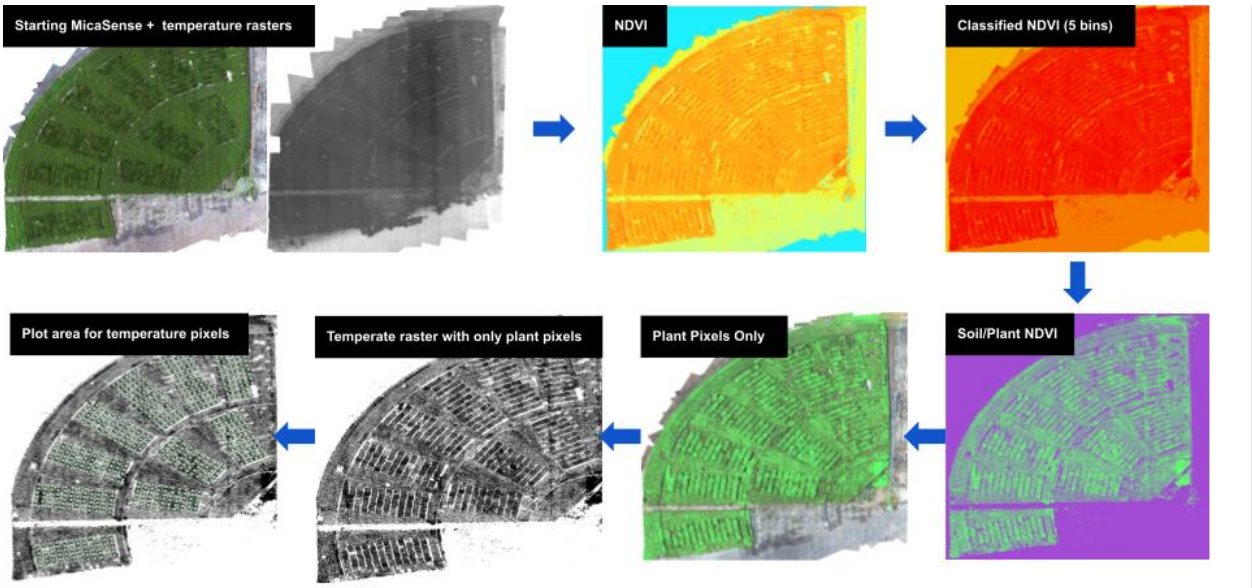


Figure 3.1. ArcGIS Pro thermal data extraction workflow using MicaSense and thermal calibrated orthomosaic raster images. The workflow is shown on the western Kansas field site.

Phenotype Extraction and Analysis

The canopy temperature data extracted from a mapping population flight is treated as an independent location-timepoint (Figure 3.2). The Colorado flights produced 4 datasets (08/11/21, 08/13/21, 08/18/21, 08/20/21), the western Kansas flights produced 1 dataset (08/08/21), and the eastern Kansas flights produced 4 datasets (08/06/21, 08/09/21, 08/25/21, 09/01/21) from the 2021 season. Each location-timepoint dataset was assessed and filtered to eliminate plots with poor stand using the 'COUNT' value extracted with the zonal statistics. Plots in Colorado were filtered to remove those with less than 350 pixels and plots in western and eastern Kansas were

filtered to remove those with less than 400 pixels. The 90th percentile temperature value ('PCT90' from the extracted zonal statistics) was used as the limited transpiration phenotype. Temperature values were standardized by centering around the mean with a unit standard deviation for each location-timepoint. Standardized temperature values were plotted for each location-timepoint to visualize distributions. Pearson's product-moment correlation coefficient was calculated for both canopy temperature and plant height, and canopy temperature and flowering time to assess spatial artifacts associated with heigh and flowering.

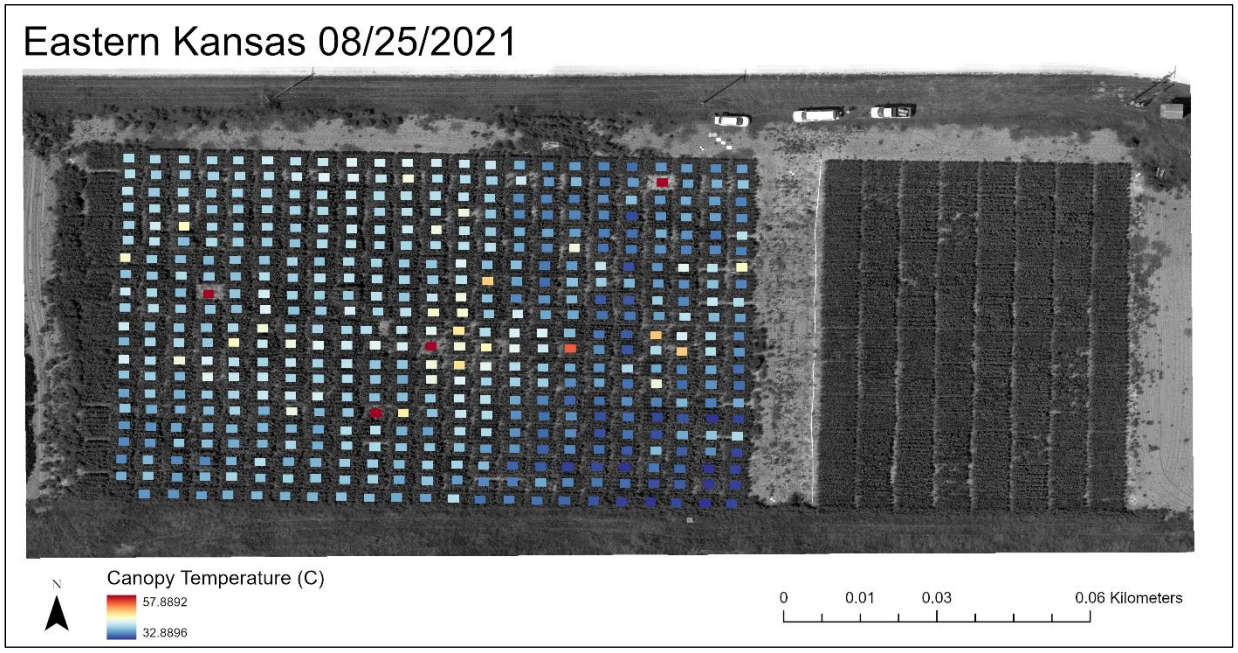


Figure 3.2. Example of spatial visualization of raw 90th percentile canopy temperature values extracted from zonal statistics in the ArcGIS Pro thermal data extraction pipeline for all plots in the 08/25/2021 eastern Kansas location-timepoint. The spatial distribution of canopy temperatures offers visual quality control assessments, such as noting that the hottest plots are those with poor stand and therefore greater soil temperature captured. Plots with low stand were filtered out before use in analysis to limit associated artifacts of falsely inflated canopy temperatures. All other location-timepoint visualizations are located in the Chapter 3 Supplementary Figures.

Estimation of Broad-Sense Heritability

Broad-sense heritability (H^2) was calculated for canopy temperature in each location-timepoint as well as across all locations. Variance components (σ^2) for the canopy temperature were estimated using the lmer() function from the lme4 package (Bates et al., 2015) in R (R Core

Development Team, 2020). For H^2 of individual location-timepoints, terms in the lmer() model included genotype and replication (block). For H^2 across location-timepoints, terms in the model included genotype and location-timepoint.

$$H^2 = \sigma^2_G/\sigma^2_P$$

The H^2 for canopy temperature in each location-timepoint (total and within-family) was estimated using the calculated variance components and the Cullis method. This method accounts for the unbalanced number of genotypes included across locations and unbalanced number of reps within locations after filtering for stand (Schmidt et al., 2019). The Cullis method uses the mean variance of pairwise differences of genotypic best linear unbiased predictors (BLUPs) using the formula:

$$H^2_{\text{Cullis}} = 1 - \frac{\bar{v}^{\text{BLUP}}_{\Delta}}{2\sigma^2_g}$$

Where σ^2_g is the genotypic variance and $\bar{v}^{BLUP}_{\Delta}$ is the mean variance of the difference of two genotypic best linear unbiased predictions (BLUPs).

Genome-Wide Association Study and Candidate Gene Analysis

The DArTseq-LD genotype data processing described in chapter 2 was used for mapping marker-trait associations between maximum canopy temperature and SNPs. Briefly, a set of 2,738 SNPs across 366 lines were generated after filtering to remove individuals with a call rate < 0.5 and SNPs that were monomorphic or non-biallelic, had a call rate < 0.5, a reproducibility index of < 0.9.

BLUPs for canopy temperature were calculated for all lines in each location-timepoint and used as phenotypes. Genotype and block were used as terms in the model and all effects were treated as random. The genotype input file was filtered to only include lines with

phenotypes for each analysis. The Genome Association and Prediction Integrated Tool (GAPIT) package for R (Lipka et al., 2012) was used to conduct a GWAS for each location-timepoint. A general linear model (GLM) was fit (Price et al., 2006) with zero principal components minimal population structure present in highly-related RILs. In the GLM model, all individuals are treated as a single group. A minor allele frequency filter was set to 5%. A nominal threshold for significant SNPs was calculated using the Bonferroni method and effective number of SNPs (Duggal et al., 2008).

Known aquaporins were identified in the model plant *Arabidposis thaliana* using The Arabidopsis Information Resource (Berardini et al., 2015). Using Phytozome (Goodstein et al., 2012), the *Arabidopsis thaliana TAIR10* genome was used to search the identified aquaporins and find sorghum protein homologs (orthologs and paralogs) in the *Sorghum bicolor v3.1.1* genome (McCormick et al., 2018). Sorghum homologs were recorded with name, position, and similarity and plotted on the GWAS Manhattan plots to visualize putative LT and aquaporin colocalizations.

RESULTS

Non-Zero H^2 Indicates Genetic Contribution to Canopy Temperature Variance

To test the hypothesis that there is a genetic contribution to the observed phenotypic variance for maximum canopy temperature, broad-sense heritability (H^2) was calculated in each location-timepoint. A non-zero H^2 was calculated in all location-timepoints except one (Table 3.1, Figure 3.4), indicating there is genetic variance underlying canopy temperature across nearly all location-timepoints.

The hypothesis that the LT trait is environmentally dependent can be tested by plotting H^2 with annual VPD weather data. Under the hypothesis that the trait is environmentally dependent on high VPD, we would predict days with higher VPD to produce higher H^2 (attributing more genetic variance to the phenotypic variance on those days). In Colorado, H^2 is relatively high in two location-timepoints (08/13/2021 and 08/18/2021) with a ~2 kPa difference of maximum VPD between those location-timepoints (Figure 3.5). Of the remaining two lower H^2 Colorado location-timepoints, the 08/11/2021 VPD is comparable to the 08/18/2021 VPD while the 08/20/2021 VPD is comparable to the 08/13/2021 VPD. In eastern Kansas, VPD is relatively low in all location-timepoints compared to Colorado and western Kansas. The eastern Kansas location-timepoints with higher H^2 (08/09/2021 and 08/25/2021) occur on days with lower VPDs compared to the lower H^2 location-timepoints (Figure 3.5). Weather station data was not available past 08/31/2021 in the eastern Kansas location, therefore the 09/01/2021 location timepoint cannot be compared with daily VPD.

Continuous Phenotype Distributions Inform Qualitative vs. Quantitative Architecture

To evaluate the hypotheses that the LT trait is qualitative (monogenic) versus quantitative, the distributions of phenotypic observations were assessed. Under a qualitative model, we would predict a bimodal distribution in each location-timepoint. The observations (Figure 3.3) instead form a continuous distribution in each location-timepoint. Standardized phenotype values range from < -2 to > 2 across the data.

Additionally, the Pearson correlation test was run for height and flowering time each against canopy temperature to evaluate the influence of agronomic background on temperature. Only height in two location-timepoints was significantly correlated ($p < 0.05$) with canopy

temperature (Table 3.2). All other height and flowering time correlations were not statistically significant.

Marker-Trait Associations Reveal Regions of Interest

To test hypotheses on oligogenic versus polygenic architecture, GWAS was used to associate regions of the genome with the LT phenotype. No significant marker-trait associations (MTA) were observed across all location-timepoints when using the Bonferroni-adjusted threshold. While not significant, regions of MTA were consistent within and across locations (Table 3.3). A repeated pattern of a small peak on the end of chromosome 1 is present in the Colorado GWAS data (Figure 3.8). In three of four eastern Kansas datasets, there is an association peak near the end of chromosome 6 (Figure 3.6). The end of chromosome 8 in the 08/08/2021 western Kansas location-timepoint shows a small peak (Figure 3.7) present also in the 08/25/2021 eastern Kansas location-timepoint. Across all locations, a small peak on the end of chromosome 4 is observed.

To begin testing the hypothesis that aquaporins are underlying variation for LT, the regions of interest were plotted with locations of known aquaporin homologs in sorghum (Table 3.4). The MTA peak on the end of chromosome 1 visible in the Colorado data is colocalized with SIP2;1 aquaporin (Sobic.001G389900.1) (Figure 3.11). The peak on the end of chromosome 6 is colocalized with PIP2;3 (Sobic.006G150100.1) (Figure 3.9). The peak on the end of chromosome 4 is colocalized with PIP3 (Sobic.004G222000.1) (Figure 3.9, Figure 3.10, Figure 3.11). The peak on the end of chromosome 8 is not colocalized with known Arabidopsis aquaporin homologs in sorghum (Figure 3.10) identified for this study.

Table 3.1. Broad-sense heritability (H^2) of maximum temperature estimated using the Cullis method at a population level and family level in each location-timepoint in the 2021 season.

Location-Timepoint		H^2
		Population-Level (Both Families)
Eastern KS		
	08/06/21	0
	08/09/21	0.25
	08/25/21	0.28
	09/01/21	0.14
Western KS		
	08/08/21	0.25
Colorado		
	08/11/21	0.19
	08/13/21	0.40
	08/18/21	0.48
	08/20/21	0.12

H^2= Broad-sense heritability

Table 3.2. Pearson correlation between height and canopy temperature, and between flowering time and canopy temperature in mapping population in each 2021 location-timepoint.

Location-Timepoint	Correlation			
	Height & Canopy Temp.	P-value	Flowering Time & Canopy Temp.	P-value
Eastern KS				
08/06/21	-0.31	0.0004 *	0.06	0.07
08/09/21	-0.17	0.05	-0.04	0.7
08/25/21	-0.16	0.06	0.06	0.6
09/01/21	-0.04	0.7	-0.12	0.5
Western KS				
08/08/21	-0.09	0.3	-	-
Colorado				
08/11/21	-0.24	0.002 *	-0.07	0.4
08/13/21	-0.11	0.2	-0.13	0.2
08/18/21	-0.09	0.26	0.08	0.3
08/20/21	-0.0015	0.99	0.02	0.8

* = significant

Table 3.3. LT marker-trait associations of interest in each 2021 location-timepoint. MTAs of interest are based on the most highly significant associations and visible peaks in Manhattan plots.

Location-Timepoint	SNP ID	Location	MAF	P-value
East Kansas 08/09/2021				
	SNP_1947303	6:60851889	0.47	$< 10^{-4}$
	SNP_34775236	5:19861937	0.27	0.0026
	SNP_2050839	9:5019093	0.31	0.0088
	SNP_28193619	1:29603323	0.14	0.0093
East Kansas 08/25/2021				
	SNP_34779343	8:21106018	0.19	$< 10^{-4}$
	SNP_15054236	8:46780288	0.3	0.00050
	SNP_2186040	6:61114540	0.48	0.00083
	SNP_15048023	4:57209399	0.22	0.0012
	SNP_15039125	3:74059839	0.5	0.0015
	SNP_2218023	10:55572007	0.06	0.0017
East Kansas 09/01/2021				
	SNP_2218023	10:55572007	0.07	0.0018
	SNP_15055950	3:69573801	0.05	0.0046
	SNP_1938396	4:68341693	0.3	0.0056
	SNP_24775441	6:44628662	0.4	0.0058
West Kansas 08/08/2021				
	SNP_34779136	10:1484372	0.3	0.00014
	SNP_1936229	8:58972170	0.25	0.00035
	SNP_15039196	6:1657119	0.26	0.00078
	SNP_15048023	4:57209399	0.22	0.0012
	SNP_2706658	1:78412303	0.42	0.0036
Colorado 08/11/2021				
	SNP_34779357	5:8898517	0.22	0.0060
Colorado 08/13/2021				
	SNP_1941743	1:75827323	0.31	0.0015

	SNP_34778082	3:43166453	0.2	0.0029
	SNP_15038396	5:4180764	0.2	0.0035
	SNP_1960273	5:64657564	0.15	0.0042
Colorado 08/18/2021				
	SNP_1918643	4:7916572	0.12	0.00067
	SNP_1887698	9:47977415	0.35	0.0025
	SNP_34779086	3:1479177	0.22	0.0032
	SNP_34779191	1:65279181	0.21	0.0046
Colorado 08/20/2021				
	SNP_15048868	1:66413583	0.23	0.00014
	SNP_34778986	4:62731172	0.26	0.00030
	SNP_28192827	2:64314914	0.24	0.0020

MAF = minor allele frequency
Location = chromosome #: bp location on chromosome

Table 3.4. Known sorghum homologs of *Arabidopsis thaliana* aquaporin loci, including sorghum gene name, physical position, and similarity to the *Arabidopsis* gene.

Arabidopsis Reference ID	Arabidopsis Name	Sorghum Gene	Sorghum Position	Similarity (%)
AT1G52180	-	Sobic.001G505100.1	1:77324937-77327995	78.8
	-	Sobic.004G295100.1	4:63501043-63502618	76
	-	Sobic.006G170600.1	6:52722391-52723580	77.1
	-	Sobic.010G146100.1	10:41392270-41394011	77.5
AT2G29870	-	Sobic.004G102200.1	4:9450178-9453286	73.5
AT2G37180	PIP2;3	Sobic.002G124700.1	2:16844699-16848362	90.5
	PIP2;3	Sobic.002G125200.1	2:16897835-16899264	86.9
	PIP2;3	Sobic.004G222000.1	4:57220819-57224296	90.5
AT3G04090	SIP1;1	Sobic.005G091600.1	5:13565973-13569467	74.2
	SIP1;1	Sobic.009G131500.1	9:48499290-48503602	79.1
AT3G06100	NIP7;1	Sobic.001G195800.1	1:17588587-17593923	60.9
AT3G56950	SIP2;1	Sobic.001G389900.1	1:67642856-67645670	69.4
AT4G18910	-	Sobic.003G026400.1	3:2231971-2234369	84.2
	-	Sobic.009G075900.1	9:9904508-9909435	78.5

AT4G35100	PIP3	Sobic.002G125000.1	2:16883368-16884816	85.3
	PIP3	Sobic.006G150100.1	6:51145122-51147729	90.1
AT4G38220	AQI	Sobic.007G208300.1	7:63740857-63745071	81.1
	AQI	Sobic.010G080300.1	10:6840554-6846224	74.9

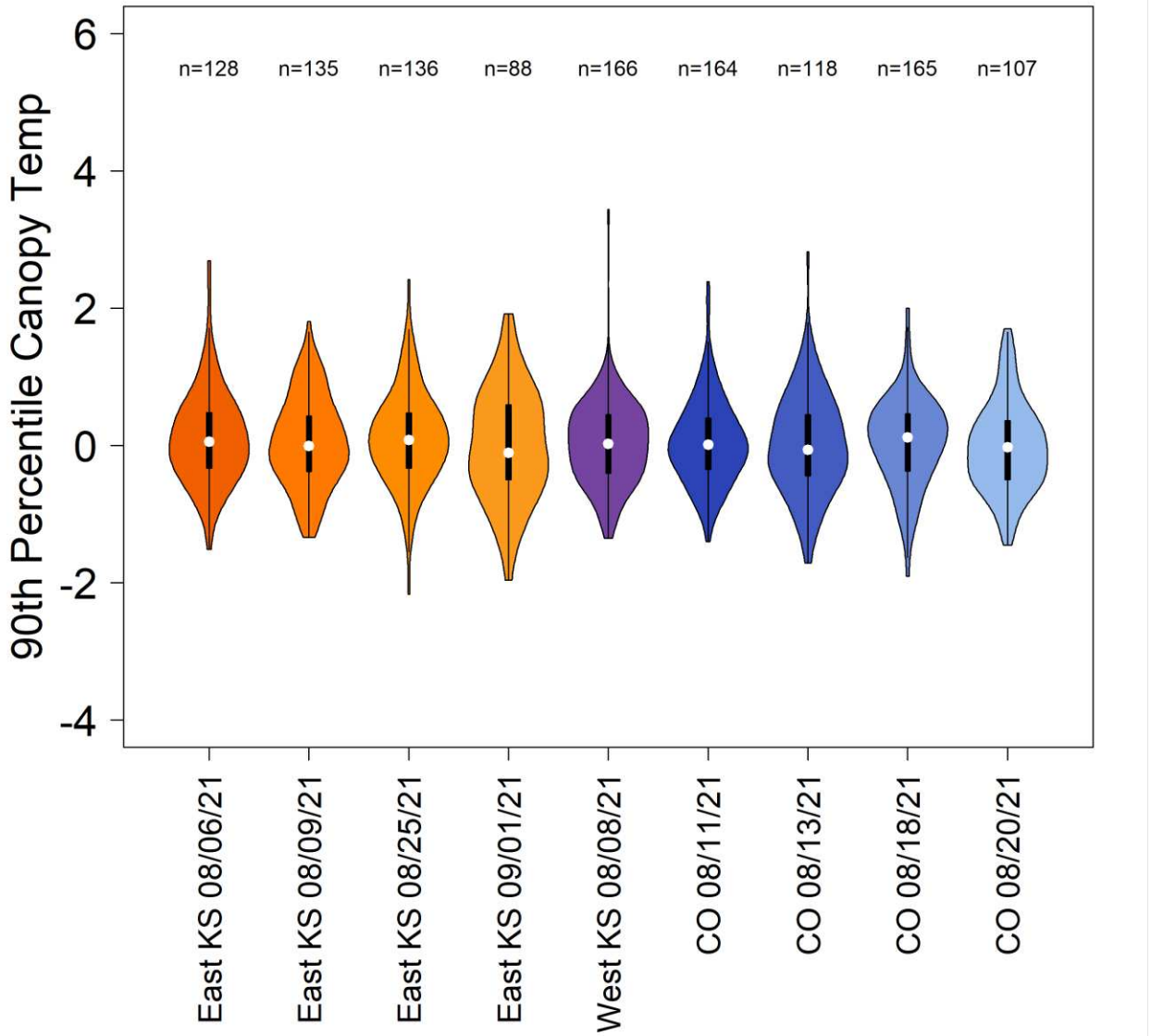


Figure 3.3. Phenotypic distributions of average 90th percentile canopy temperature (standardized) for each RIL in the population (after filtering for stand count) in each location-timepoint from the 2021 season. White dots indicate mean, centered at zero due to the standardization. Population size is denoted by "n =".

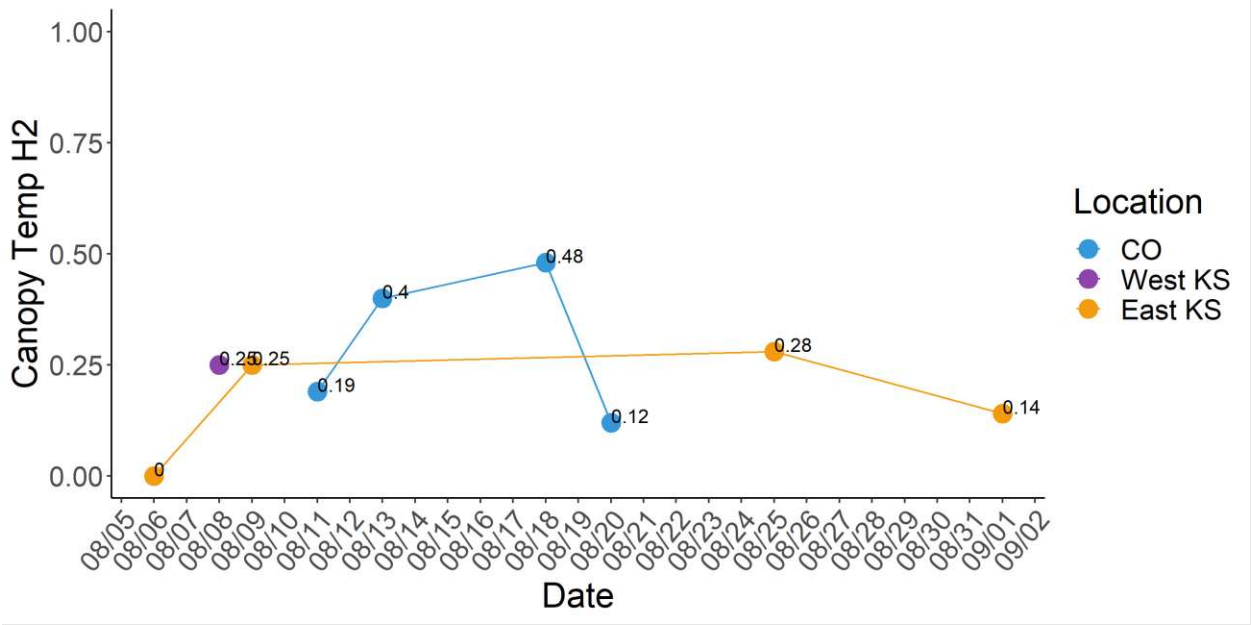


Figure 3.4. Broad-sense heritability (H^2) of 90th percentile canopy temperature estimated using the Cullis method across location-timepoints from the 2021 season.

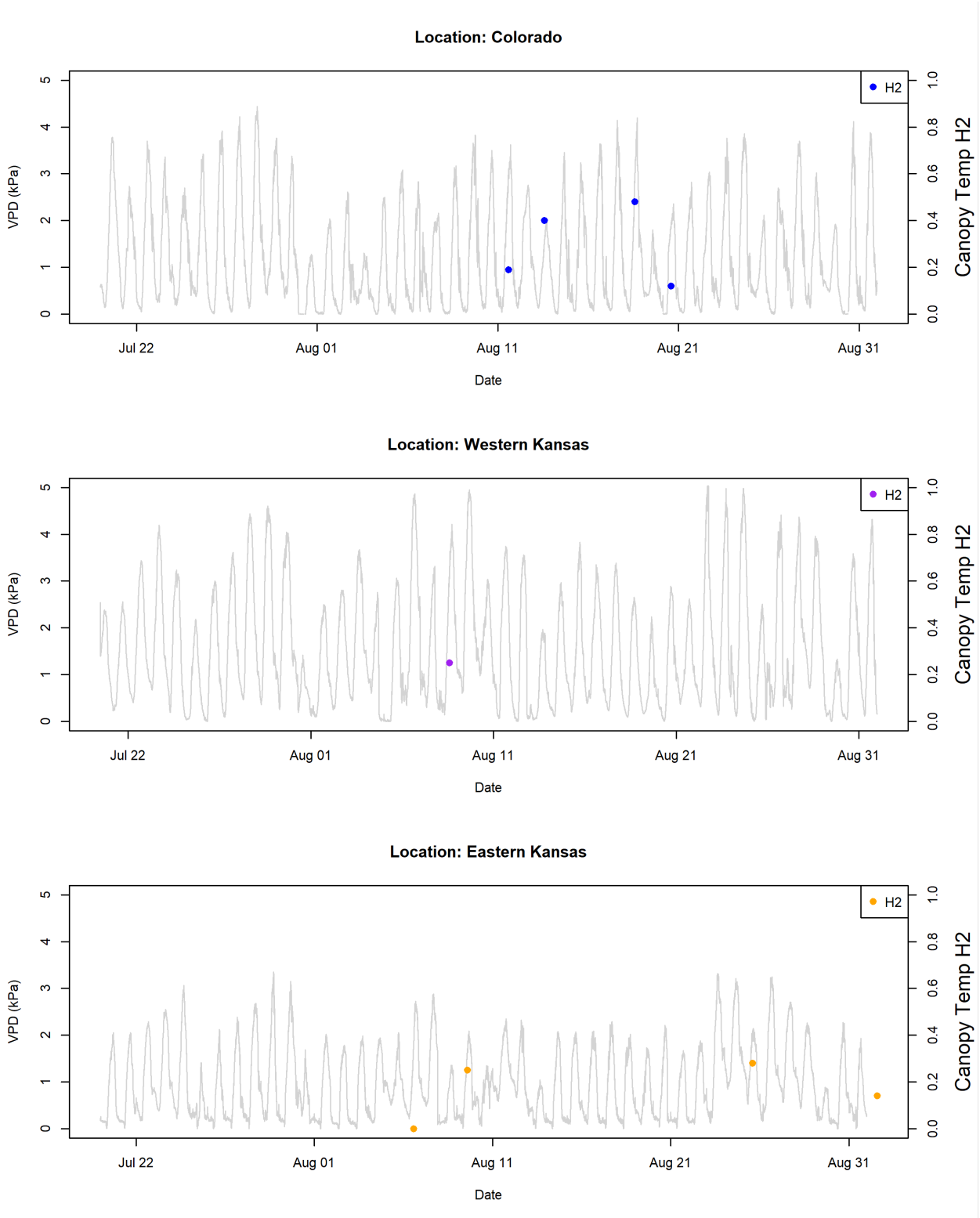


Figure 3.5. Broad-sense heritability (H^2) (right y-axis) of 90th percentile canopy temperature for each 2021 location-timepoint overlaid on weather station vapor pressure deficit (VPD) data (left y-axis, light gray line) collected every five minutes for each field site. Comparing the H^2 to the maximum daily VPD (light gray peaks) can reveal a VPD threshold for the environmental dependence of the LT trait.

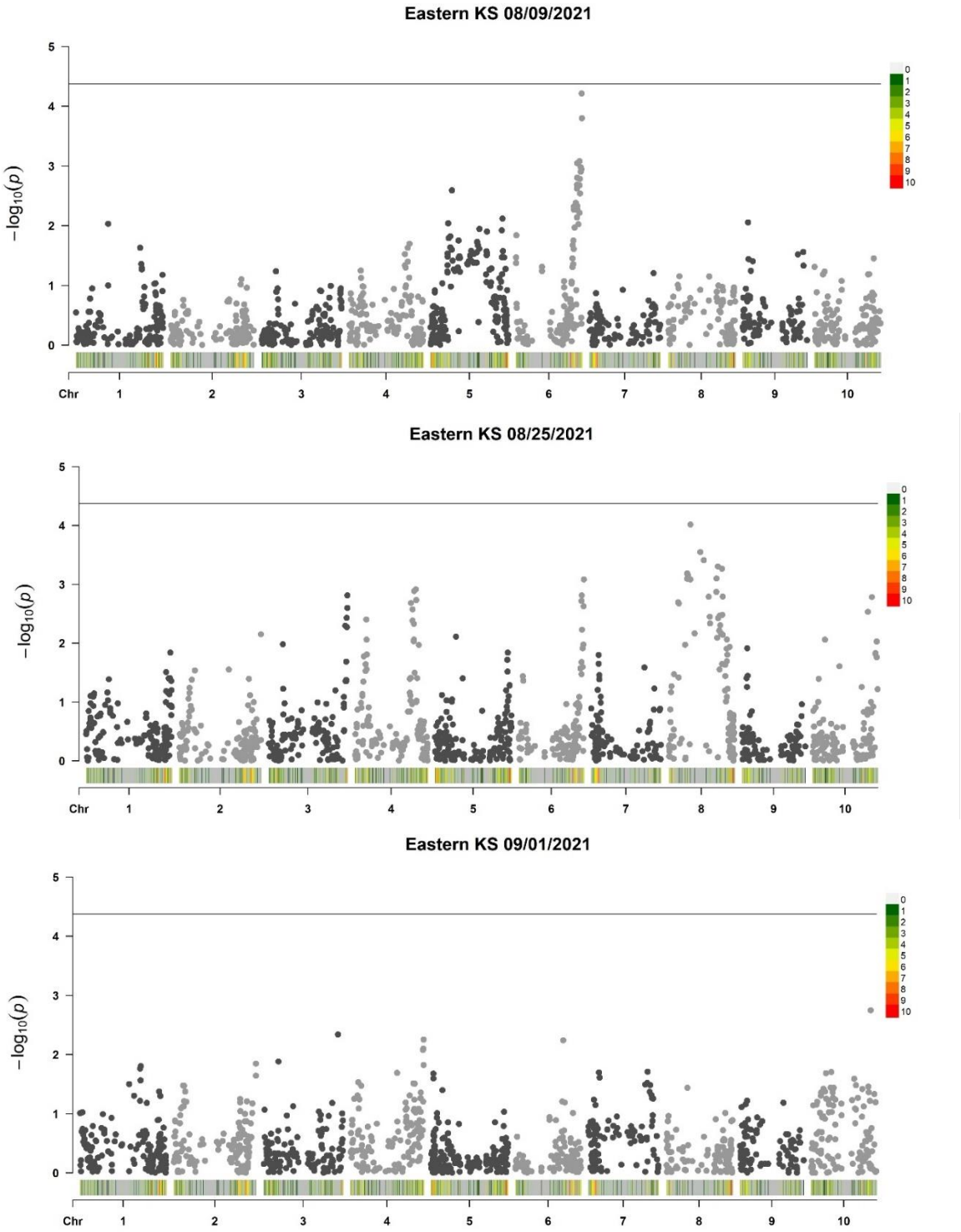


Figure 3.6. Manhattan plots for GWAS results showing associations of genetic markers and 90th percentile canopy temperature BLUPs in the eastern Kansas 2021 location-timepoints. The black horizontal line indicates Bonferroni-adjusted significance threshold for markers. The red to green scale indicates marker density.

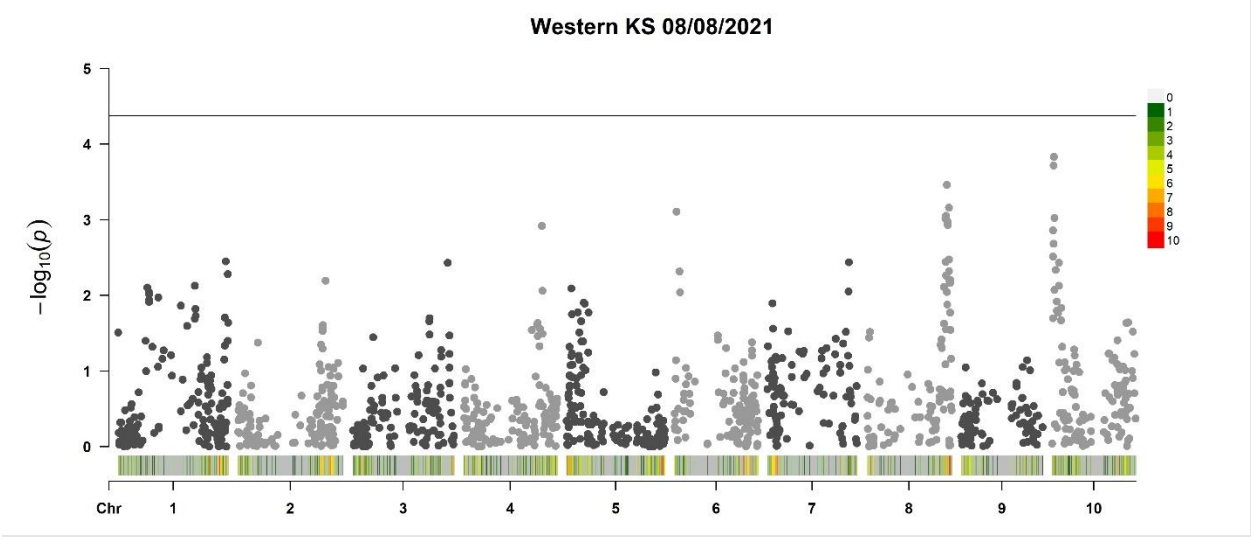
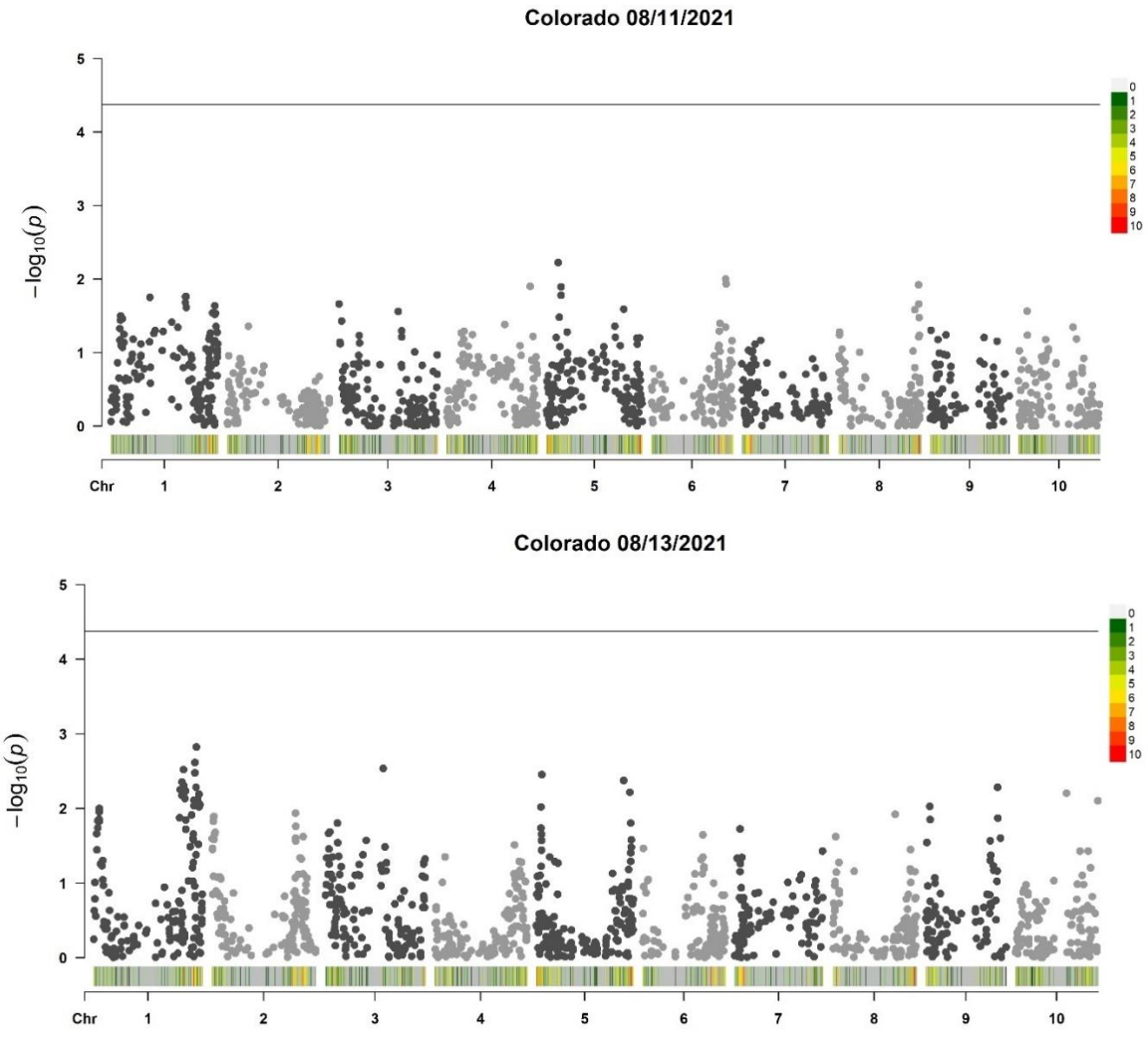


Figure 3.7. Manhattan plot for GWAS results showing associations of genetic markers and 90th percentile canopy temperature BLUPs in the western Kansas 2021 location-timepoint (08/08/2021). The black horizontal line indicates Bonferroni-adjusted significance threshold for markers. The red to green scale indicates marker density.


Colorado 08/11/2021
$-\log_{10}(p)$
Chr
Colorado 08/13/2021
$-\log_{10}(p)$
Chr

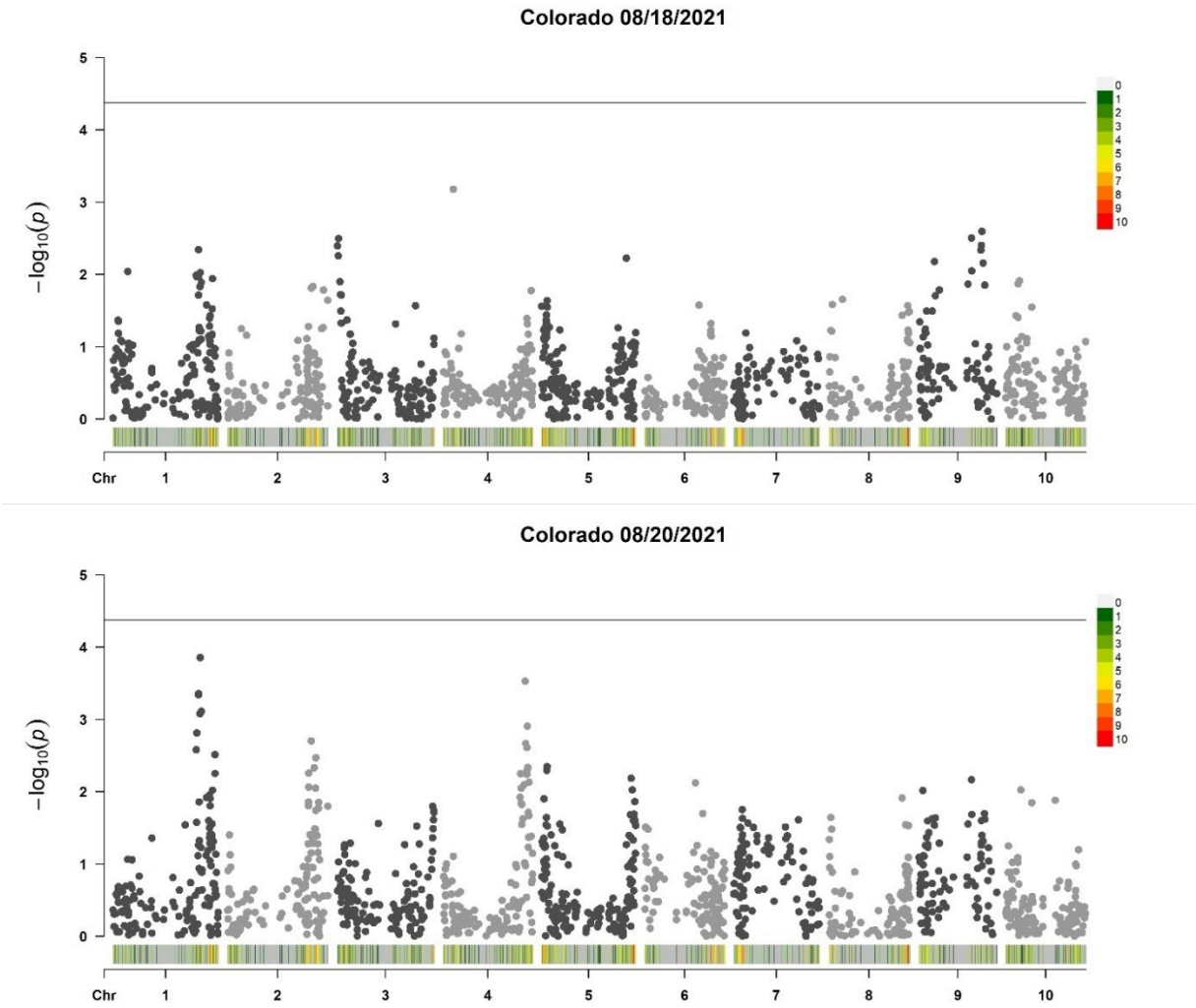


Figure 3.8. Manhattan plots for GWAS results showing associations of genetic markers and 90th percentile canopy temperature BLUPs in the Colorado 2021 location-timepoints. The black horizontal line indicates Bonferroni-adjusted significance threshold for markers. The red to green scale indicates marker density.

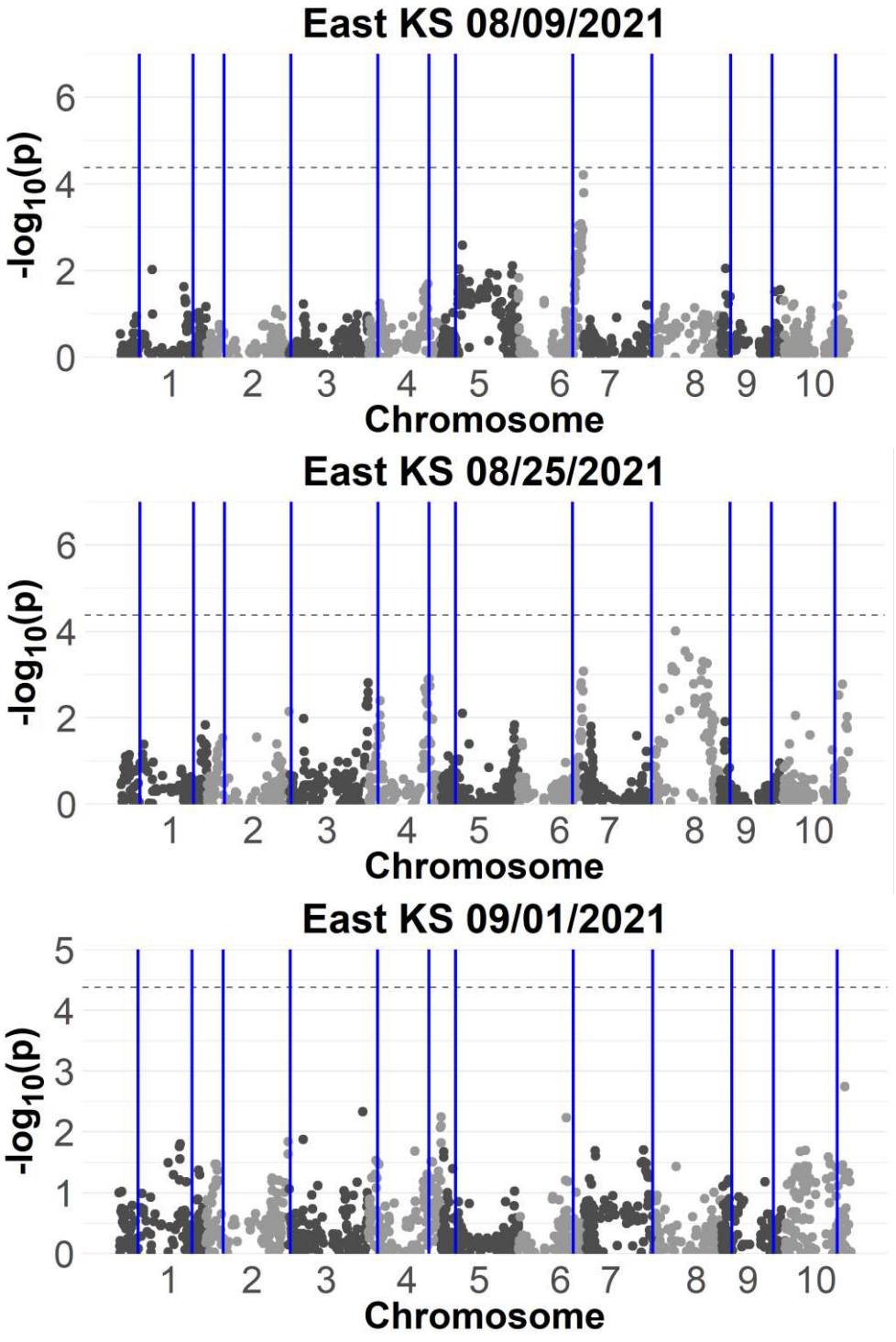


Figure 3.9. Manhattan plots for GWAS results showing associations of genetic markers and 90th percentile canopy temperature BLUPs in the Eastern Kansas 2021 location-timepoints with loci corresponding to known *Arabidopsis thaliana* aquaporin homologs in sorghum marked (blue vertical lines). Horizontal dashed gray line indicates Bonferroni adjusted significance threshold.

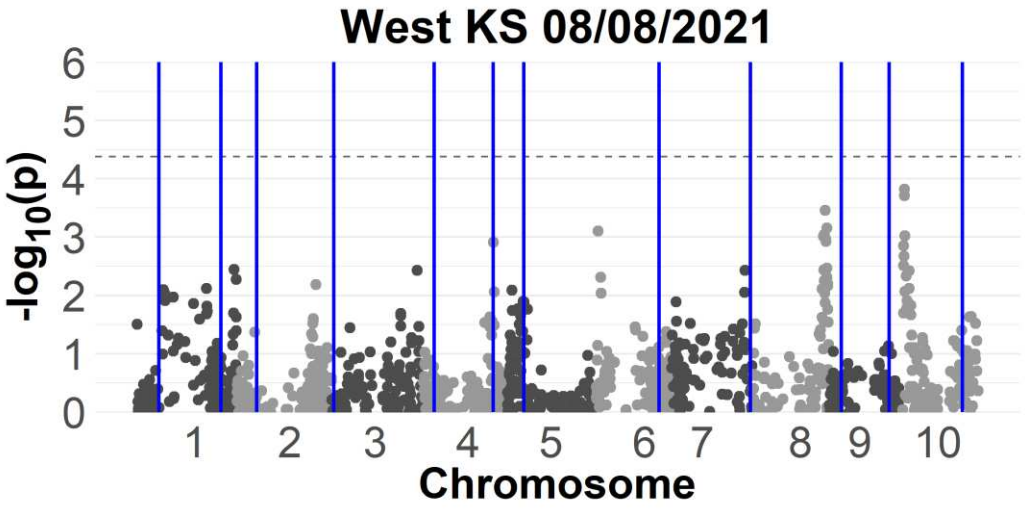
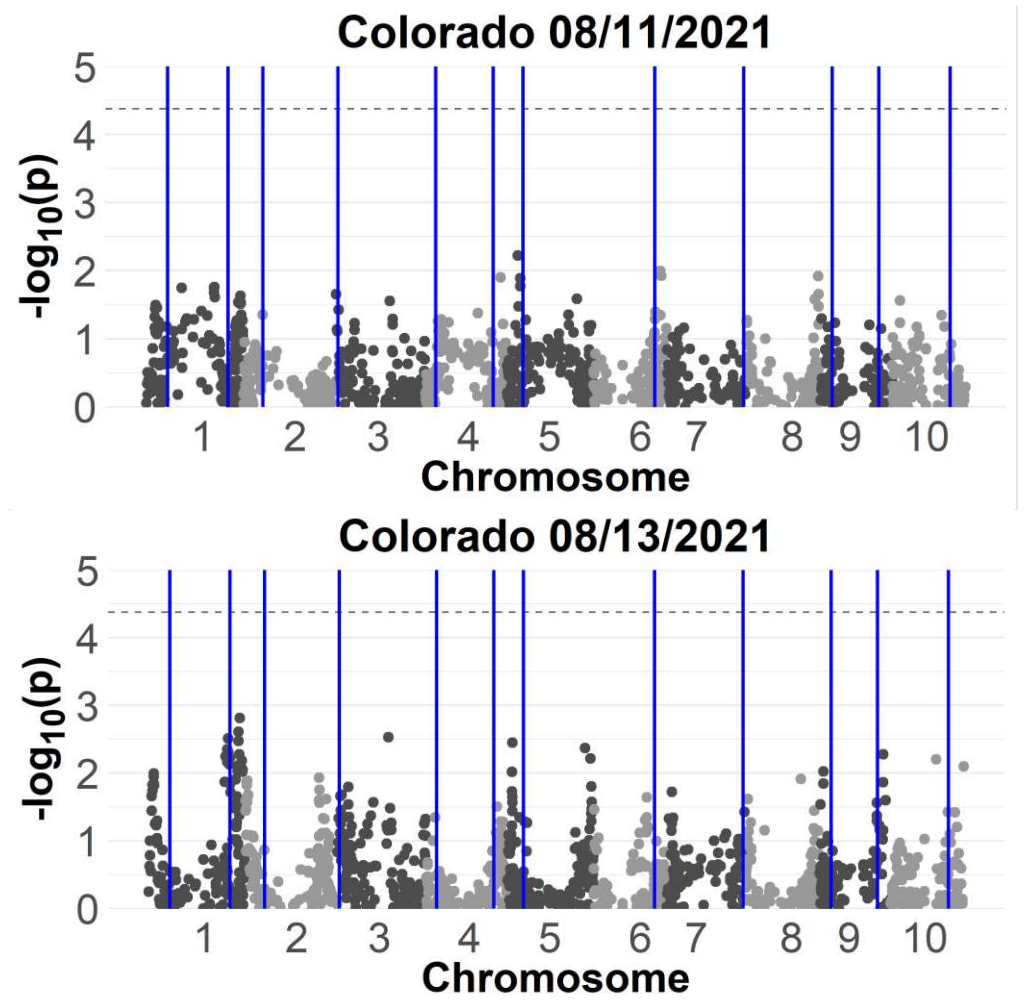


Figure 3.10. Manhattan plots for GWAS results showing associations of genetic markers and 90th percentile canopy temperature BLUPs in the Western Kansas 2021 location-timepoint with loci corresponding to known *Arabidopsis thaliana* aquaporin homologs in sorghum marked (blue vertical lines). Horizontal dashed gray line indicates Bonferroni adjusted significance threshold.



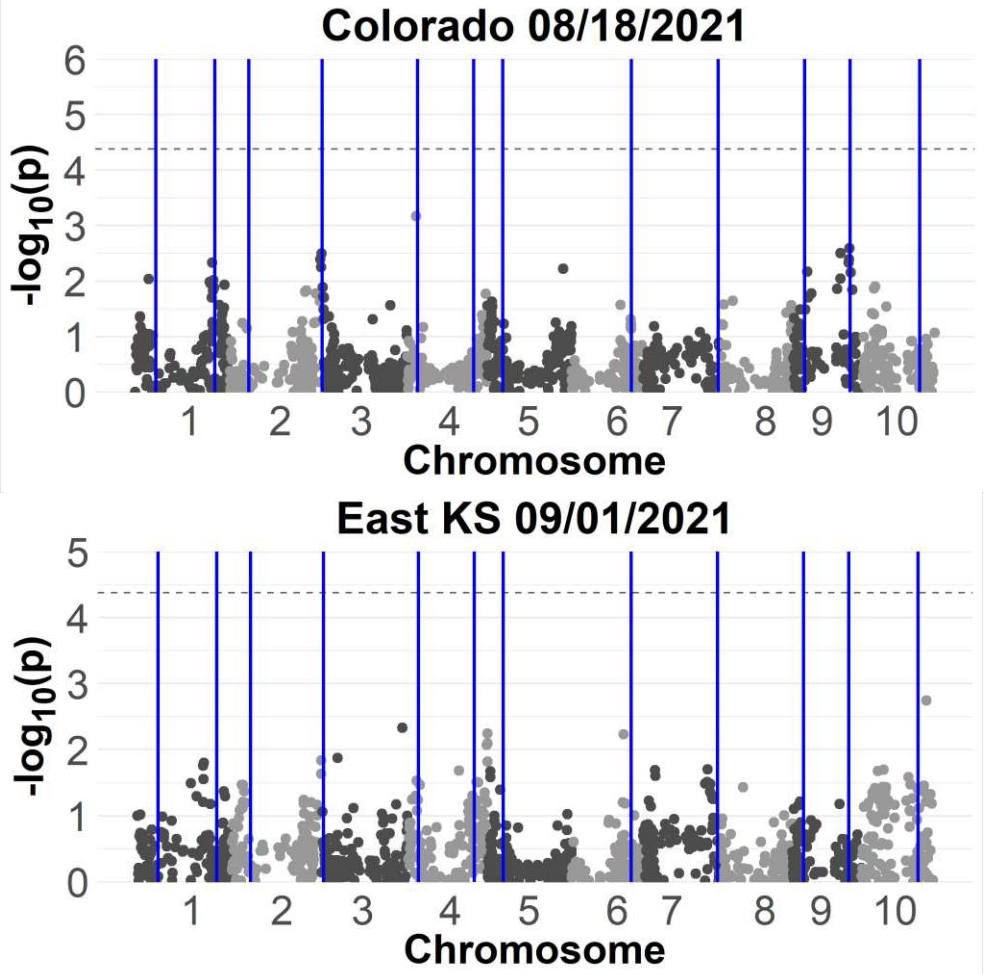


Figure 3.11. Manhattan plots for GWAS results showing associations of genetic markers and 90th percentile canopy temperature BLUPs in the Colorado 2021 location-timepoints with loci corresponding to known *Arabidopsis thaliana* aquaporin homologs in sorghum marked (blue vertical lines). Horizontal dashed gray line indicates Bonferroni adjusted significance threshold.

DISCUSSION

Knowledge of genetic variation underlying a trait of interest is integral to developing a pre-breeding strategy (Bernardo, 2008). A structured hypothesis testing framework allows us to exclude certain architectures and generate knowledge on trait molecular basis that can best inform breeding decisions for a trait of interest.

H^2 and Phenotypic Distribution of Canopy Temperature Establishes the Presence of Quantitative Genetic Variation for LT

Characterizing LT genetic architecture begins by confirming there is genetic variation for the trait in the mapping population. The temporal variability of canopy temperature may result in a large environmental effect on the phenotype (Figure 3.3) when using UAS-based phenotype data. One study evaluating the use of UAS to capture thermal data in wheat found that the H^2 of canopy temperature ranges between 0.36 to 0.74 (Perich et al., 2020), while another study comparing canopy temperature and stomatal conductance in wheat using similar UAS methodology found H^2 up to 0.75 (Rebetzke et al., 2012). The LT mapping populations in this study produced ranges of non-zero H^2 across all but one location-timepoint (Table 3.1, Figure 3.4), supporting the hypothesis that there is a genetic contribution to the phenotypic variance. However, it is important to note that genetic variation for traits other than LT may be contributing to the canopy temperature phenotype, in this study and previous studies (Ries et al., 2012; Saint Pierre et al., 2010). The LT mapping families specifically sought to account for covariates associated with morphological traits like height, flowering time, and general canopy architecture (Raymundo et al., in preparation). Pearson tests found that these traits were not correlated with canopy temperature in most location-timepoints, indicating sufficient control of covariates. However, leaf size and leaf angle were not measured and may affect soil irradiance and canopy temperature interference captured by the UAS sensing to produce artifacts (Ferguson et al., 1973; Fuchs et al., 1967).

Understanding the dependence of the LT trait on high VPD conditions could elucidate a hypothetical VPD threshold and narrow target environments for the trait. H^2 across timepoints in each location did not show the predicted trend when compared with VPD data under the

hypothesis that expression of the LT trait is triggered by high VPD (Gholipoor et al., 2013). Days with high VPD relative to days with lower VPD did not show consistently higher H^2 (Figure 3.5). Certain location-timepoints followed the predicted trend, but half did not. Further work to characterize the VPD-dependence of the LT trait using daily flight data could provide further insight into this environment-gene relationship.

Traits controlled by a single Mendelian factor will produce bimodal distributions of phenotypes (Sameri et al., 2009; Zhang et al., 2012). The opposite is true of quantitative traits, which produce continuous distributions as the result of the contribution of more genetic factors with smaller additive effects (Lande, 1982). Studies across a multitude of plant species have shown that ecophysiological water-use traits are typically quantitative (Brendel et al., 2008; Chen et al., 2011; Tharanya et al., 2018). The phenotypes for canopy temperature of the LT mapping populations (Figure 3.3) display a continuous range and suggest quantitative control.

Moderate Effect QTL May Be Conferring LT

Oligogenic architecture, characterized by a modest number of moderate-effect loci, is more favorable for targeted introgressions using marker-assisted selection (MAS) than polygenic architecture (DeWitt et al., 2021). Again, the quantitative control of many physiological and water-use traits is polygenic (Faralli et al., 2019). The LT mapping families produced no putative significant MTA across all location-timepoints (Figure 3.6, Figure 3.7, Figure 3.8). However, GWAS revealed smaller association peaks, some of which were consistent across timepoints within location or across locations. Notable MTA on chromosome 6 near SNP_1947303 (MAF = 0.47), chromosome 8 near SNP_ 34779343 (MAF = 0.19), chromosome 1 near SNP_15048868 (MAF = 0.23), and chromosome 4 near SNP_1918643 (MAF = 0.12) demonstrated such patterns (Table 3.3). We therefore infer that the LT trait is under oligogenic control, and that those

regions with consistent MTA are probable for at least partial control of LT variation. Further GWAS studies using denser marker coverage and more effective recombinations may be advantageous to further elucidate LT trait architecture. Additional data from the 2022 season is under collection and consolidation and will be used for additional analyses of genetic architecture of the LT trait in sorghum.

The hypothesis that aquaporins are underlying the LT trait is generated by the rich body of work describing the relationship of aquaporins and transpiration rate (Heinen et al., 2009; Maurel et al., 2016). Aquaporins facilitate water, CO_2, and uncharged solute transfer across membranes. In turn, the turgor pressure of guard cells that regulate stomata opening and closing are partially a function of aquaporins (Mosehlion et al., 2015). While the findings on the molecular basis of LT are preliminary, the colocalizations of association peaks with known aquaporin loci (Figure 3.9, Figure 3.10, Figure 3.11) generate new hypotheses on the contribution of particular aquaporin genes to LT variation. Notably, *SIP2;1* (Sobic.001G389900.1) and SNP_15048868 on chromosome 1 are approximately 500 kb apart. Sobic.006G170600.1 and SNP_1947303 are in close proximity on chromosome 6, and *PIP2;3* (Sobic.004G222000.1) and SNP_1918643 are within 12,000 basepairs of one another (Table 3.2, Table 3.3). Colocalizations are observed at aquaporin loci representing several subgroups including small basic intrinsic proteins and plasma membrane intrinsic proteins. These findings suggest variation of LT may be attributed to a variety of water transport functions related to aquaporins within the plant.

Implications for LT Donor Line Development

Pre-breeding with oligogenic traits is facilitated through marker-assisted selection. Developing selectable markers is possible due to statistically detectable variation present in

oligogenic traits. Breeding for polygenic traits is complicated by the large number of undetectable loci contributing to the phenotype (Scott et al., 2021). Breeding programs build elite yield haplotypes that can span most of a linkage group. Transferring a trait with tens or hundreds of underlying genes may break up those haplotypes, lowering the performance of the progeny or reduced heterosis (Smith et al., 2010). LT growth chamber studies originally found 17 genotypes with transpiration rate breakpoint responses (Gholipoor et al., 2013), indicating genetic variation for LT may already be present in elite breeding materials. Therefore, a more complex polygenic architecture would suggest using genomic selection (GS) to identify and advance lines with high LT performance in target environment breeding trials (Meuwissen et al., 2001). The oligogenic architecture of LT proposed by this study guides development of selectable markers for use in pre-breeding a trait donor line.

CONCLUSION

The limited transpiration trait holds promise for increasing water-use efficiency in dryland sorghum cropping systems. Development of an elite donor line for use in plant breeding is stalled by the need for effective molecular markers to select on. Multi-environment trials of an LT mapping population using canopy temperature as a phenotype proxy found that genetic variation for LT is present in the mapping families, and small but repeated peaks across timepoints and locations indicate oligogenic architecture. Known *Arabidopsis thaliana* aquaporin homologs in sorghum colocalize with several of the peaks, providing some insights into the potential molecular basis underlying LT. The implications of oligogenic architecture for donor line development support identifying selectable markers for marker-assisted selection. Further repetition across years and environments will increase power and resolution of the putative QTL for the LT trait in sorghum.

REFERENCES

Aguilar-Benitez, D., Casimiro-Soriguer, I., Maalouf, F., & Torres, A. M. (2021). Linkage mapping and QTL analysis of flowering time in faba bean. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-92680-4

Bali, S., Mamgain, A., Raina, S. N., Yadava, S. K., Bhat, V., Das, S., Pradhan, A. K., & Goel, S. (2015). Construction of a genetic linkage map and mapping of drought tolerance trait in Indian beveragial tea. *Molecular Breeding*, *35*(5), 112. https://doi.org/10.1007/s11032-015-0306-5

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Belko, N., Zaman-Allah, M., Diop, N. n., Cisse, N., Zombre, G., Ehlers, J. d., & Vadez, V. (2013). Restriction of transpiration rate under high vapour pressure deficit and non-limiting water conditions is important for terminal drought tolerance in cowpea. *Plant Biology*, *15*(2), 304–316. https://doi.org/10.1111/j.1438-8677.2012.00642.x

Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., & Huala, E. (2015). The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. *Genesis*, *53*(8), 474–485. https://doi.org/10.1002/dvg.22877

Bernardo, R. (2008). Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. *Crop Science*, *48*(5), 1649–1664. https://doi.org/10.2135/cropsci2008.03.0131

Brendel, O., Le Thiec, D., Scotti-Saintagne, C., Bodénès, C., Kremer, A., & Guehl, J.-M. (2008).

Quantitative trait loci controlling water use efficiency and related traits in Quercus robur L. *Tree Genetics & Genomes*, *4*(2), 263–278. https://doi.org/10.1007/s11295-007-0107-z

Broman, K. W. (2005). The Genomes of Recombinant Inbred Lines. *Genetics*, *169*(2), 1133–1146. https://doi.org/10.1534/genetics.104.035212

Casper, B. B., Forseth, I. N., & Wait, D. A. (2005). Variation in Carbon Isotope Discrimination in Relation to Plant Performance in a Natural Population of Cryptantha flava. *Oecologia*, *145*(4), 541–548.

Chen, J., Chang, S. X., & Anyia, A. O. (2011). Gene discovery in cereals through quantitative trait loci and expression analysis in water-use efficiency measured by carbon isotope discrimination. *Plant, Cell & Environment*, *34*(12), 2009–2023. https://doi.org/10.1111/j.1365-3040.2011.02397.x

Dai, A., Zhao, T., & Chen, J. (2018). Climate Change and Drought: A Precipitation and Evaporation Perspective. *Current Climate Change Reports*, *4*(3), 301–312. https://doi.org/10.1007/s40641-018-0101-6

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., James, R. A., Condon, A. G., Bovill, W. D., Hutchinson, P., Scarrow, J., Davy, R., & Furbank, R. T. (2016). Methodology for High-Throughput Field Phenotyping of Canopy Temperature Using Airborne Thermography. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01808

DeWitt, N., Guedira, M., Lauer, E., Murphy, J. P., Marshall, D., Mergoum, M., Johnson, J., Holland, J. B., & Brown-Guedira, G. (2021). Characterizing the oligogenic architecture of plant growth phenotypes informs genomic selection approaches in a common wheat population. *BMC Genomics*, *22*(1), 402. https://doi.org/10.1186/s12864-021-07574-6

Duggal, P., Gillanders, E. M., Holmes, T. N., & Bailey-Wilson, J. E. (2008). Establishing an

adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics*, *9*(1), 516. https://doi.org/10.1186/1471-2164-9-516

Edwards, C. E., Ewers, B. E., Williams, D. G., Xie, Q., Lou, P., Xu, X., McClung, C. R., & Weinig, C. (2011). The Genetic Architecture of Ecophysiological and Circadian Traits in Brassica rapa. *Genetics*, *189*(1), 375–390. https://doi.org/10.1534/genetics.110.125112

Faralli, M., Matthews, J., & Lawson, T. (2019). Exploiting natural variation and genetic manipulation of stomatal conductance for crop improvement. *Current Opinion in Plant Biology*, *49*, 1–7. https://doi.org/10.1016/j.pbi.2019.01.003

Ferguson, H., Eslick, R. F., & Aase, J. K. (1973). Canopy Temperatures of Barley as Influenced by Morphological Characteristics1. *Agronomy Journal*, *65*(3), 425–428. https://doi.org/10.2134/agronj1973.00021962006500030021x

Fuchs, M., Kanemasu, E. T., Kerr, J. P., & Tanner, C. B. (1967). Effect of Viewing Angle on Canopy Temperature Measurements with Infrared Thermometers1. *Agronomy Journal*, *59*(5), 494–496. https://doi.org/10.2134/agronj1967.00021962005900050040x

Geber, M. A., & Dawson, T. E. (1997). Genetic Variation in Stomatal and Biochemical Limitations to Photosynthesis in the Annual Plant, Polygonum arenastrum. *Oecologia*, *109*(4), 535–546.

Gholipoor, M., Prasad, P. V. V., Mutava, R. N., & Sinclair, T. R. (2010). Genetic variability of transpiration response to vapor pressure deficit among sorghum genotypes. *Field Crops Research*, *119*(1), 85–90. https://doi.org/10.1016/j.fcr.2010.06.018

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, *40*(D1), D1178–D1186.

https://doi.org/10.1093/nar/gkr944

Gorjanc, G., Jenko, J., Hearne, S., & Hickey, J. (2016). Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. *BMC Genomics*, *17*, 30. https://doi.org/10.1186/s12864-015-2345-z

Gyawali, A., Shrestha, V., Guill, K. E., Flint-Garcia, S., & Beissinger, T. M. (2019). Single-plant GWAS coupled with bulk segregant analysis allows rapid identification and corroboration of plant-height candidate SNPs. *BMC Plant Biology*, *19*(1), 412. https://doi.org/10.1186/s12870-019-2000-y

Heinen, R. B., Ye, Q., & Chaumont, F. (2009). Role of aquaporins in leaf physiology. *Journal of Experimental Botany*, *60*(11), 2971–2985. https://doi.org/10.1093/jxb/erp171

Holland, J. B. (2007). Genetic architecture of complex traits in plants. *Current Opinion in Plant Biology*, *10*(2), 156–161. https://doi.org/10.1016/j.pbi.2007.01.003

Huynh, B.-L., Close, T. J., Roberts, P. A., Hu, Z., Wanamaker, S., Lucas, M. R., Chiulele, R., Cissé, N., David, A., Hearne, S., Fatokun, C., Diop, N. N., & Ehlers, J. D. (2013). Gene Pools and the Genetic Architecture of Domesticated Cowpea. *The Plant Genome*, *6*(3), plantgenome2013.03.0005. https://doi.org/10.3835/plantgenome2013.03.0005

Kumar, S., Kirk, C., Deng, C. H., Shirtliff, A., Wiedow, C., Qin, M., Wu, J., & Brewer, L. (2019). Marker-trait associations and genomic predictions of interspecific pear (Pyrus) fruit characteristics. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-45618-w

Lande, R. (1982). A Quantitative Genetic Theory of Life History Evolution. *Ecology*, *63*(3), 607–615. https://doi.org/10.2307/1936778

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S.,

& Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Liu, W., Reif, J. C., Ranc, N., Porta, G. D., & Würschum, T. (2012). Comparison of biometrical approaches for QTL detection in multiple segregating families. *Theoretical and Applied Genetics*, *125*(5), 987–998. https://doi.org/10.1007/s00122-012-1889-4

Martínez-Carricondo, P., Agüera-Vega, F., Carvajal-Ramírez, F., Mesas-Carrascosa, F.-J., García-Ferrer, A., & Pérez-Porras, F.-J. (2018). Assessment of UAV-photogrammetric mapping accuracy based on variation of ground control points. *International Journal of Applied Earth Observation and Geoinformation*, *72*, 1–10. https://doi.org/10.1016/j.jag.2018.05.015

Maurel, C., Verdoucq, L., & Rodrigues, O. (2016). Aquaporins and plant transpiration. *Plant, Cell & Environment*, *39*(11), 2580–2587. https://doi.org/10.1111/pce.12814

McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., Kennedy, M., Amirebrahimi, M., Weers, B. D., McKinley, B., Mattison, A., Morishige, D. T., Grimwood, J., Schmutz, J., & Mullet, J. E. (2018). The Sorghum bicolor reference genome: Improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *The Plant Journal*, *93*(2), 338–354. https://doi.org/10.1111/tpj.13781

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

Moshelion, M., Halperin, O., Wallach, R., Oren, R., & Way, D. A. (2015). Role of aquaporins in determining transpiration and photosynthesis in water-stressed plants: Crop water-use efficiency, growth and yield. *Plant, Cell & Environment*, *38*(9), 1785–1793.

https://doi.org/10.1111/pce.12410

Perich, G., Hund, A., Anderegg, J., Roth, L., Boer, M. P., Walter, A., Liebisch, F., & Aasen, H. (2020). Assessment of Multi-Image Unmanned Aerial Vehicle Based High-Throughput Field Phenotyping of Canopy Temperature. *Frontiers in Plant Science*, *11*. https://www.frontiersin.org/articles/10.3389/fpls.2020.00150

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), Article 8. https://doi.org/10.1038/ng1847

Rajon, E., & Plotkin, J. B. (2013). The evolution of genetic architectures underlying quantitative traits. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1769), 20131552. https://doi.org/10.1098/rspb.2013.1552

Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., Condon, A. (Tony) G., Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., & Condon, A. (Tony) G. (2012). Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. *Functional Plant Biology*, *40*(1), 14–33. https://doi.org/10.1071/FP12184

Ries, L. L., Purcell, L. C., Carter Jr., T. E., Edwards, J. T., & King, C. A. (2012). Physiological Traits Contributing to Differential Canopy Wilting in Soybean under Drought. *Crop Science*, *52*(1), 272–281. https://doi.org/10.2135/cropsci2011.05.0278

Saint Pierre, C., Crossa, J., Manes, Y., & Reynolds, M. P. (2010). Gene action of canopy temperature in bread wheat under diverse environments. *Theoretical and Applied Genetics*, *120*(6), 1107–1117. https://doi.org/10.1007/s00122-009-1238-4

Sameri, M., Nakamura, S., Nair, S. K., Takeda, K., & Komatsuda, T. (2009). A quantitative trait

locus for reduced culm internode length in barley segregates as a Mendelian gene. *Theoretical and Applied Genetics*, *118*(4), 643–652. https://doi.org/10.1007/s00122-008-0926-9

Schmidt, P., Hartung, J., Rath, J., & Piepho, H.-P. (2019). Estimating Broad-Sense Heritability with Unbalanced Data from Agricultural Cultivar Trials. *Crop Science*, *59*(2), 525–536. https://doi.org/10.2135/cropsci2018.06.0376

Scott, M. F., Fradgley, N., Bentley, A. R., Brabbs, T., Corke, F., Gardner, K. A., Horsnell, R., Howell, P., Ladejobi, O., Mackay, I. J., Mott, R., & Cockram, J. (2021). Limited haplotype diversity underlies polygenic trait architecture across 70 years of wheat breeding. *Genome Biology*, *22*(1), 137. https://doi.org/10.1186/s13059-021-02354-7

Shekoofa, A., Balota, M., & Sinclair, T. R. (2014). Limited-transpiration trait evaluated in growth chamber and field for sorghum genotypes. *Environmental and Experimental Botany*, *99*, 175–179. https://doi.org/10.1016/j.envexpbot.2013.11.018

Sinclair, T. R., Devi, J., Shekoofa, A., Choudhary, S., Sadok, W., Vadez, V., Riar, M., & Rufty, T. (2017). Limited-transpiration response to high vapor pressure deficit in crop species. *Plant Science*, *260*, 109–118. https://doi.org/10.1016/j.plantsci.2017.04.007

Smith, S., Primomo, V., Monk, R., Nelson, B., Jones, E., & Porter, K. (2010). Genetic Diversity of Widely Used U.S. Sorghum Hybrids 1980–2008. *Crop Science*, *50*(5), 1664–1673. https://doi.org/10.2135/cropsci2009.10.0619

Tharanya, M., Kholova, J., Sivasakthi, K., Seghal, D., Hash, C. T., Raj, B., Srivastava, R. K., Baddam, R., Thirunalasundari, T., Yadav, R., & Vadez, V. (2018). Quantitative trait loci (QTLs) for water use and crop production traits co-locate with major QTL for tolerance to water deficit in a fine-mapping population of pearl millet (Pennisetum glaucum L.

R.Br.). *Theoretical and Applied Genetics*, *131*(7), 1509–1529. https://doi.org/10.1007/s00122-018-3094-6

Thomson, M. J. (2014). High-Throughput SNP Genotyping to Accelerate Crop Improvement. *Plant Breeding and Biotechnology*, *2*(3), 195–212. https://doi.org/10.9787/PBB.2014.2.3.195

Van Der Plank, J. E. (1966). Horizontal (polygenic) and vertical (oligogenic) resistance against blight. *American Potato Journal*, *43*(2), 43–52. https://doi.org/10.1007/BF02871406

Varshney, R. K., Barmukh, R., Roorkiwal, M., Qi, Y., Kholova, J., Tuberosa, R., Reynolds, M. P., Tardieu, F., & Siddique, K. H. M. (2021). Breeding custom-designed crops for improved drought adaptation. *Advanced Genetics*, *2*(3), e202100017. https://doi.org/10.1002/ggn2.202100017

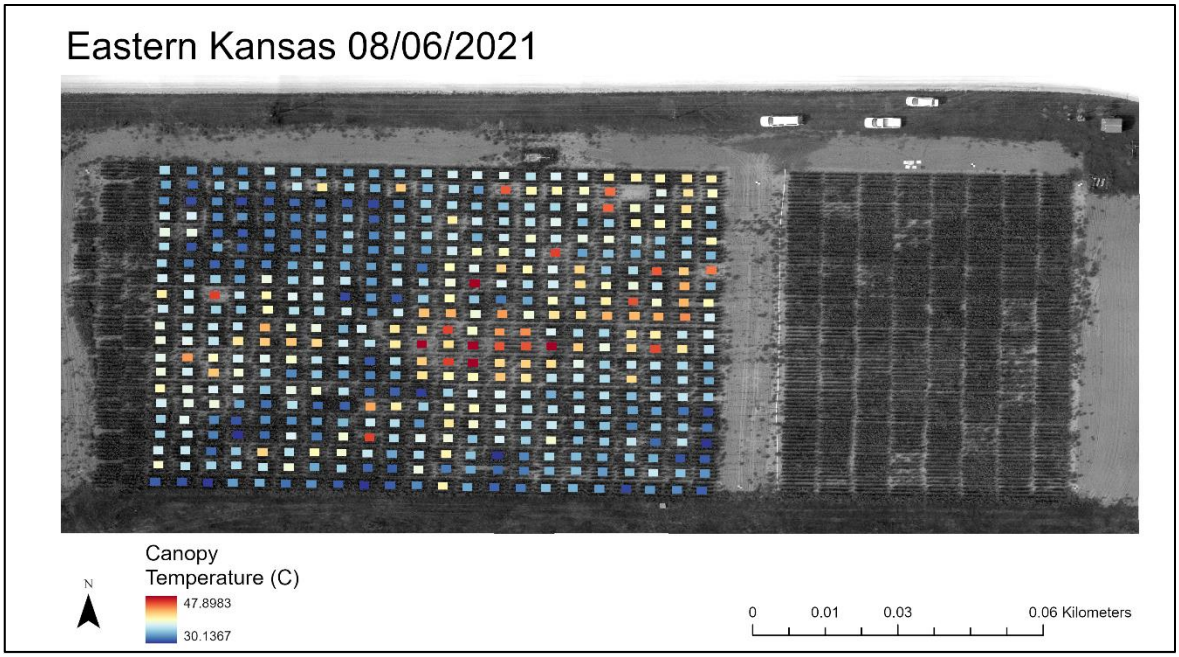
Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. *Genetics*, *136*(4), 1457–1468. https://doi.org/10.1093/genetics/136.4.1457

Zhang, L., Li, S., Chen, L., & Yang, G. (2012). Identification and mapping of a major dominant quantitative trait locus controlling seeds per silique as a single Mendelian factor in Brassica napus L. *Theoretical and Applied Genetics*, *125*(4), 695–705. https://doi.org/10.1007/s00122-012-1861-3

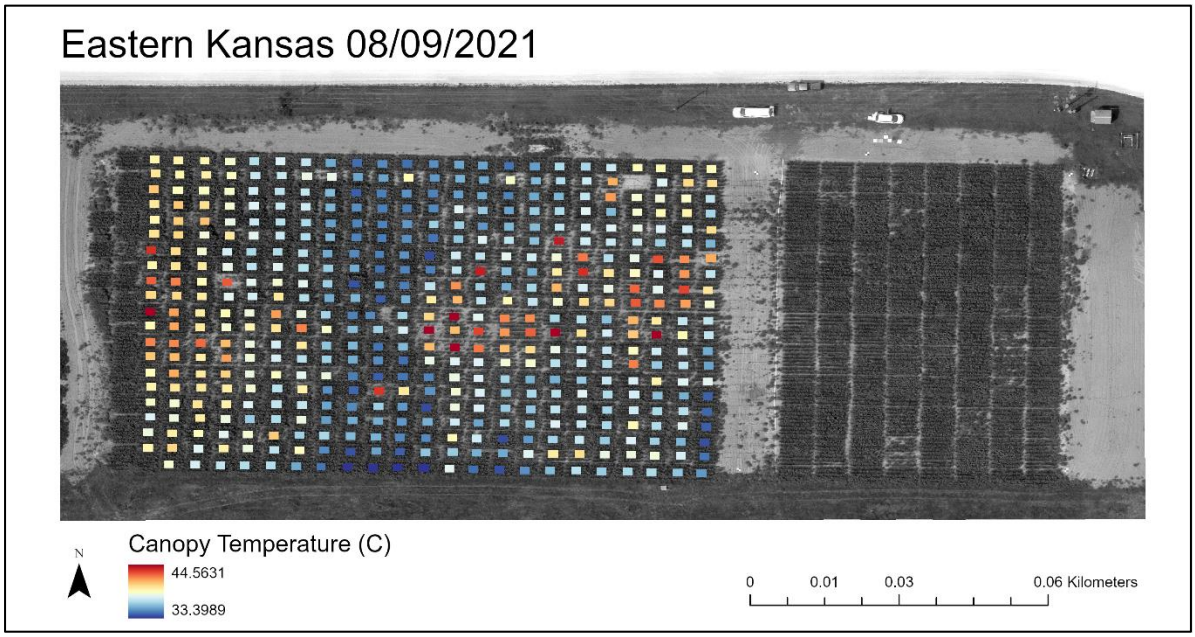
Zsögön, A., Cermak, T., Voytas, D., & Peres, L. E. P. (2017). Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: Case study in tomato. *Plant Science*, *256*, 120–130. https://doi.org/10.1016/j.plantsci.2016.12.012

APPENDIX

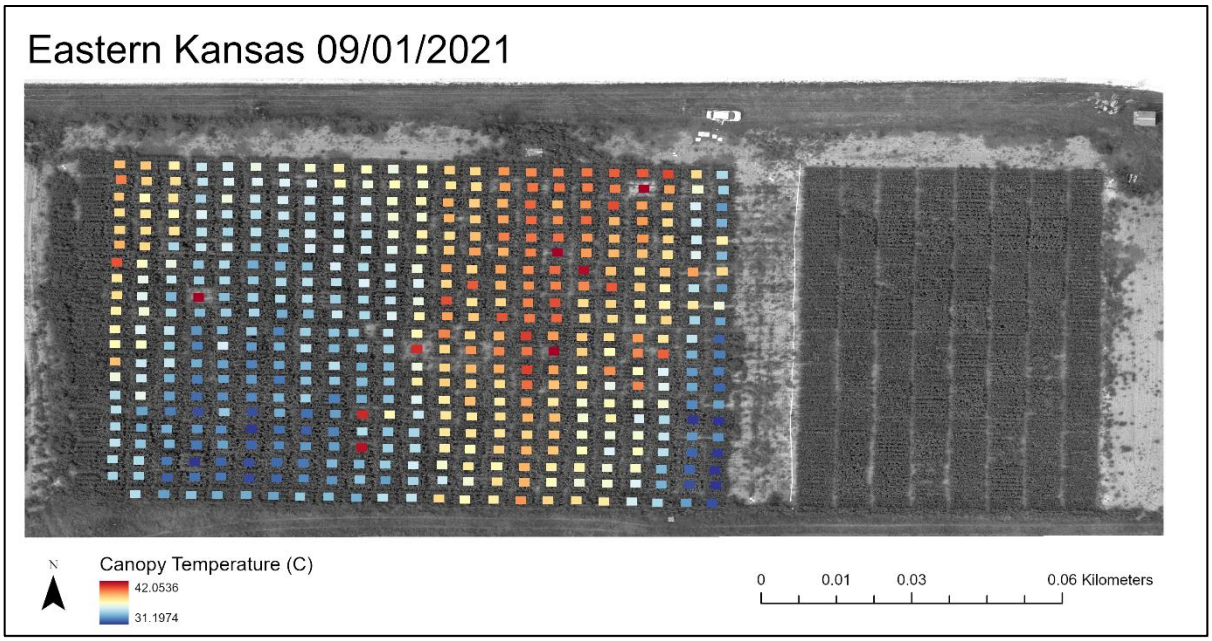
Chapter III Supplemental



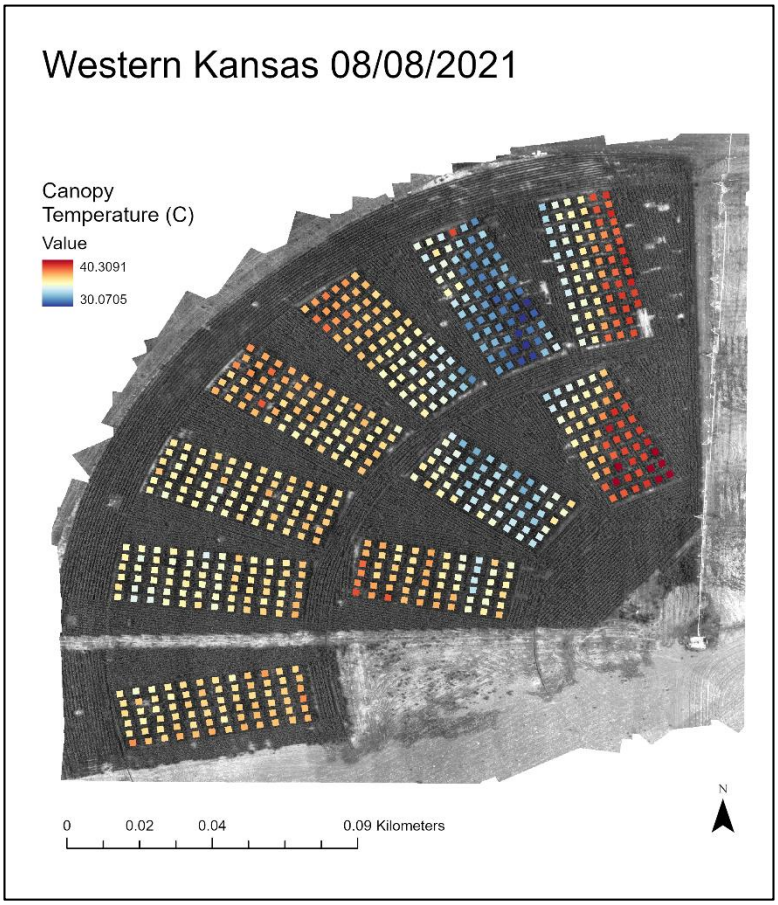
Supplementary Figure 1. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from eastern Kansas 08/06/2021 flight data.



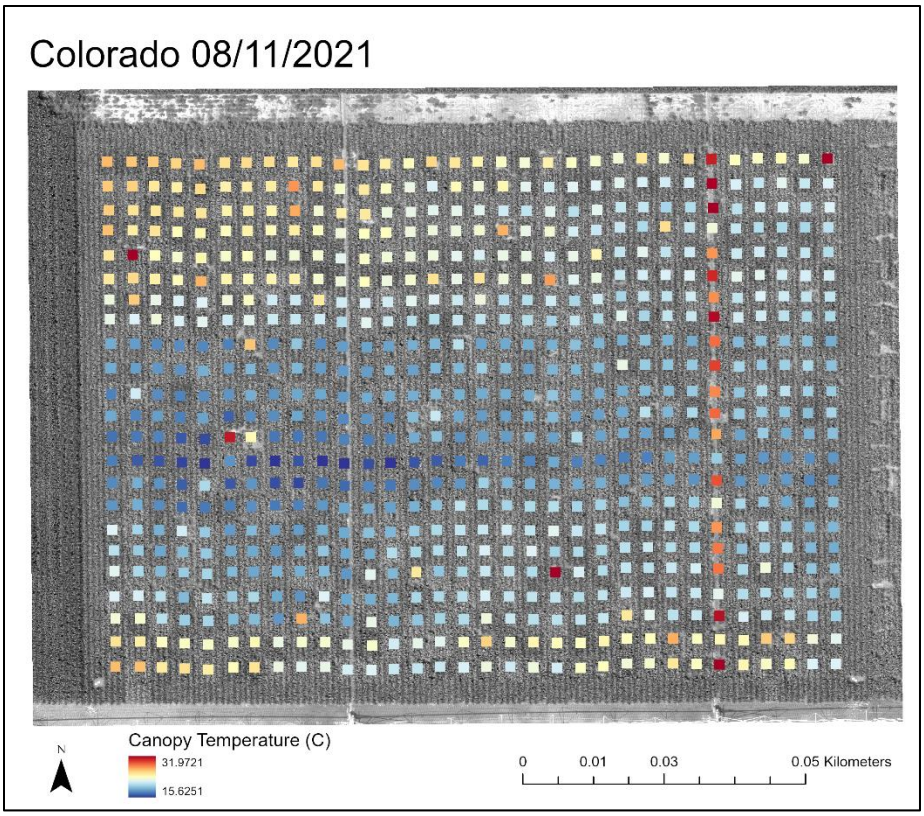
Supplementary Figure 2. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from eastern Kansas 08/09/2021 flight data.



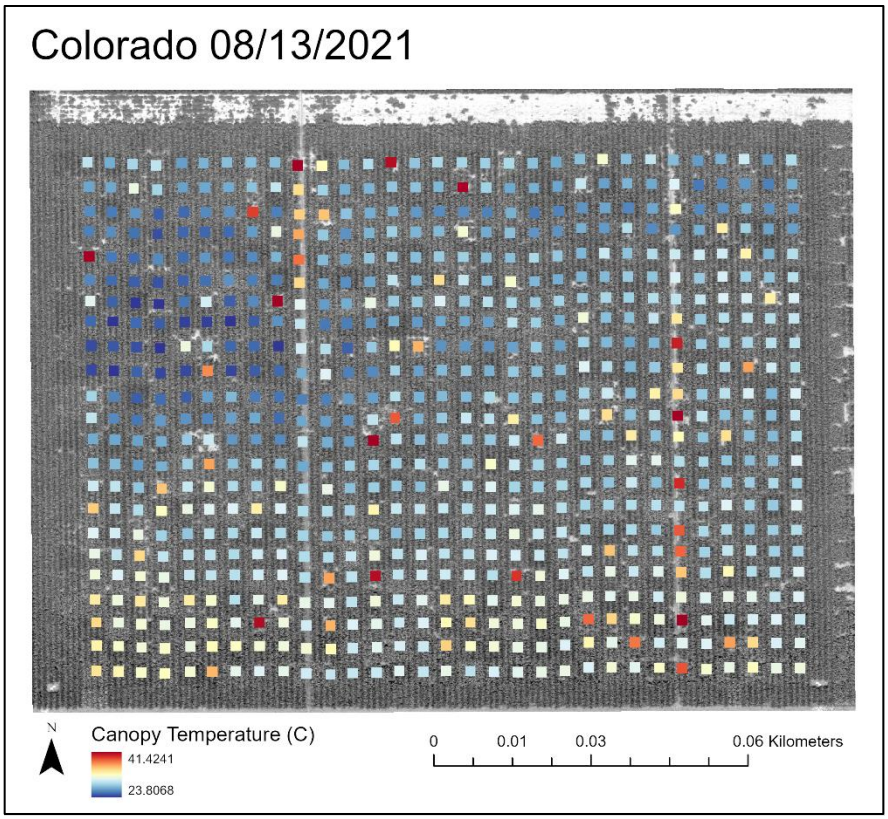
Supplementary Figure 3. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from eastern Kansas 09/01/2021 flight data.



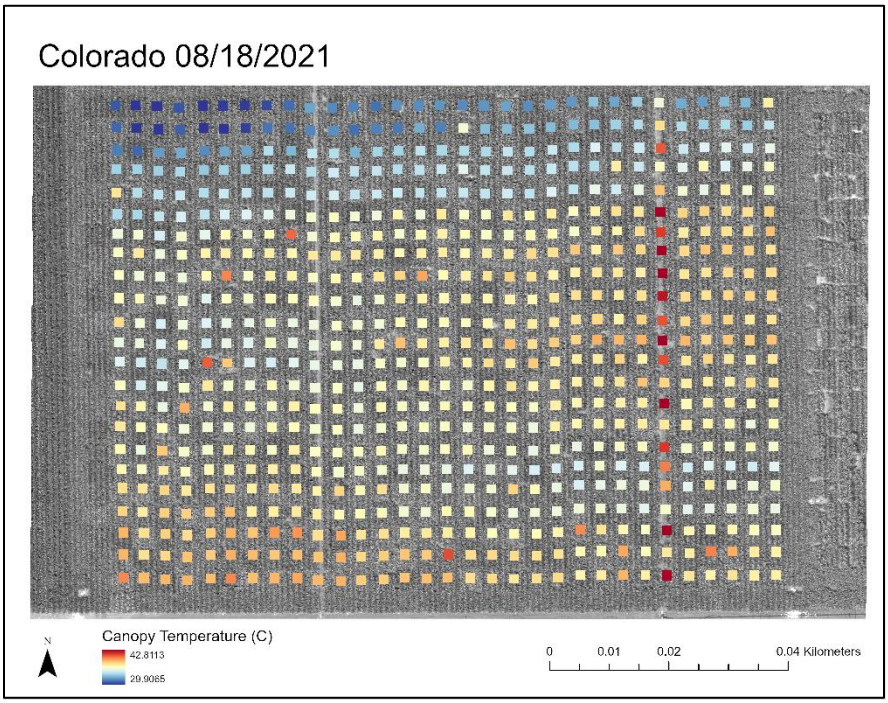
Supplementary Figure 4. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from western Kansas 08/08/2021 flight data.



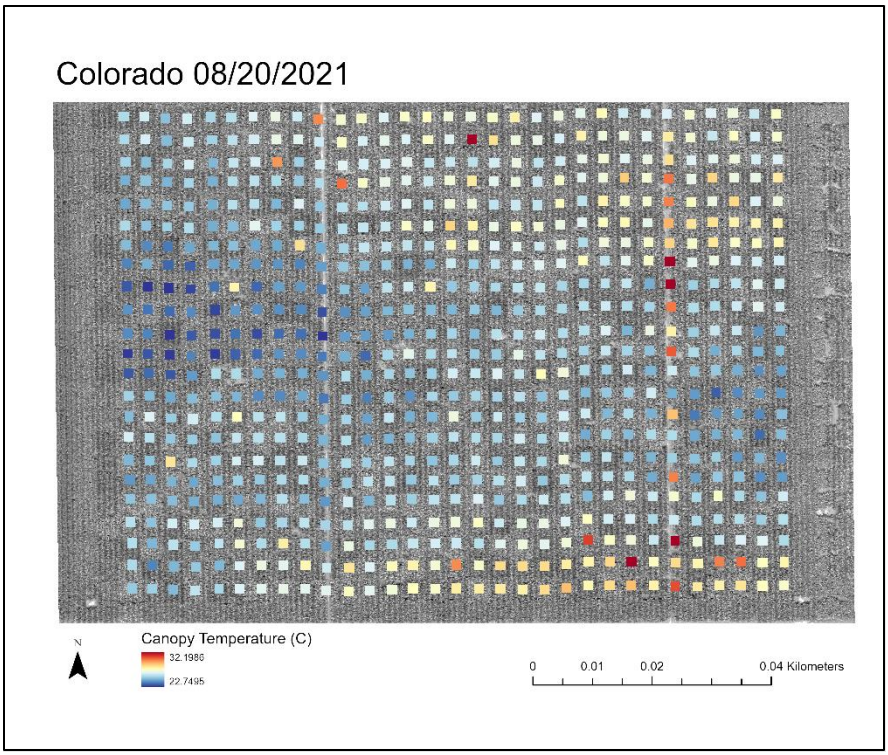
Supplementary Figure 5. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from Colorado 08/11/2021 flight data.



Supplementary Figure 6. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from Colorado 08/13/2021 flight data.



Supplementary Figure 7. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from Colorado 08/18/2021 flight data.



Supplementary Figure 8. Spatial visualization of 90th percentile canopy temperature (degrees C) for each plot extracted using zonal statistics from Colorado 08/20/2021 flight data.